

Measuring Product Type with Sequential Dynamics of Online Product Reviews: Theory and Applications

Completed Unpublished Paper (word count: 9,982)

Kevin Yili Hong
Temple University

Pei-yu Chen
Arizona State University

Lorin M. Hitt
University of Pennsylvania

Abstract

The distinction between experience products and search products (“product type”) is increasingly important in online commerce due to the way products of different types interact with Internet-based technologies such as online retailing, search engines and online reviews. One challenge in understanding the role of product type in online commerce is that it is difficult to measure objectively. We propose an approach where product type can be inferred based on the statistical analysis of online product reviews. Our theoretical analyses indicate that for pure search products, the variance of the mean rating should decrease as more consumers rate the product; for experience products the variance of the mean rating may remain constant or increase depending on the importance of the experience attributes in consumer utility. We demonstrate the use of this classification approach at category, product and attribute level using archival data from Amazon.com, Yelp.com and Ctrip.com. Implications of this analytical tool and empirical findings for research, theory and managerial practice are discussed.

Keywords: Online Product Reviews, Product Type, Statistical Analysis, Information Content, Product Quality

1. Introduction

This decade has seen huge amount of user generated content (UGC) on the World Wide Web, which experienced unprecedented growth with the support of social media technologies. One aspect of UGC that has attracted significant research attention is online product reviews (e.g., Dellarocas 2003, Chevalier and Mayzlin 2006, Chen and Xie 2008, Li and Hitt 2010, Archak et al. 2011, Sun 2012, Godes and Silva 2012, Gu and Ye 2012) due to the important effect these technologies can have on online and offline commerce. To a great extent, online product reviews are a natural fit with social media, allowing the extension of “word of mouth” (WOM), which is known to be important for product marketing in the offline world, to online commerce. In addition, it is these systems that have facilitated reductions in consumer search costs in recent years which is critical for understanding how information systems affect organizations and markets (Malone et al. 1987, Bakos 2007). Online review systems also provide a wealth of data about consumer perceptions and the way these perceptions are communicated, providing data on behaviors that are difficult or impossible to observe on a large scale in the offline world. This characteristic has led to the development of a substantial literature that has connected measurable characteristics of online reviews such as review valence and volume (Chevalier and Mayzlin 2006, Dellarocas et al. 2007, Duan et al. 2008b, Zhu and Zhang 2010), text comments (Archak et al. 2011, Ghose 2010), and review variance (Clemons et al. 2006, Sun 2012) on consumer purchase behavior, product market share or other economic outcomes. More recently, however, researchers have begun to use this information to better understand the characteristics of products purchased online (Ghose and Ipeirotis 2010).

In this study we utilize online review data to make inferences about the relative proportion of search attributes versus experience attributes in products sold online which we refer to as “product type”, consistent with the existing literature. Search attributes are characteristics that can be unambiguously described and therefore amenable to identification through consumer search. Experience attributes are characteristics that cannot be perceived without a consumer actually consuming the product.

The identification and effects of product type are part of a longer stream of work on product

differentiation and its effect on market behavior (e.g. Bain 1993, Chamberlin 1950). The foundations for this distinction can be traced to the Nobel winning work of Stigler (1961) who noted that the relative cost of obtaining product information could have a significant impact on how consumers search and the type of price equilibrium that will be observed. Nelson (1970) made a pioneering effort in classifying search and experience products, arguing that for a rational consumer, experience will only be used when search becomes too expensive. He focused specifically on repair expenditure¹ as an experience attribute, since it is difficult to quantify at time of purchase. Figure 1 shows the basis for his proposed classification scheme as applied to a variety of products. Efforts to empirically distinguish experience attributes from search attributes has considerable developed over time (Nelson 1970, 1974, 1981, Smith 1990, Laband 1991, Weathers et al. 2007) and a parallel literature has developed linking these characteristics to marketing strategy, consumer behavior, or marketplace outcomes (see e.g., Klein 1998, Senecal and Nantel 2004, Villas-Boas 2004, Zhu and Zhang 2010, Mudambi and Schuff 2010, Huang et al. 2009). These concepts are sufficiently developed that they can be directly applied in the business school marketing curriculum (see e.g., Clarkson and Miller 1982, Greer 1984, Martin 1988).

ESTIMATES OF THE RATIO OF REPAIR EXPENDITURES TO SALES BY GOOD, 1963		
Goods	Nonmerchandise Receipts*	Total Repair†
Experience goods:		
Jewelry	12.6	17.3
Typewriters	8.8	...
Radio, television	8.1	24.3‡
Tire, battery	7.5	...
Aircraft, boats, motorcycles	7.0	...
Heating and plumbing	6.4	...
Bicycles	6.0	...
Automobiles	5.7	30.0§
Music instruments	4.9	10.7
Appliances	4.4	14.0
Search goods:		
Floor covering	3.8	...
Garden implements	3.5	...
Sporting goods	2.8	...
Household trailers	2.5	...
Cameras	2.4	...
Furniture	2.1	7.6
Paint and mirrors	1.9	...
China, glassware	1.9	...
Hardware	1.6	...
Hobbies, games	0.8	...

* The percentage of nonmerchandise receipts to sales of stores specializing in the sale of the specified good.
 † To the percentage of nonmerchandise receipts, we add the ratio of the receipts of service specialists to the sales of the good for those goods whose service specialists are specified in the U.S. Bureau of the Census (1966b).
 ‡ The U.S. Commerce Department estimates this percentage as roughly 29% (U.S. Department of Commerce 1964).
 § Commerce estimates for this percentage is roughly 23%.
 || Furniture stores have a percentage of nonmerchandise receipts of 2.7. But they also sell television and appliances. The tabulated figure adjusts for the proportion of sales in the three categories.

¹ The data Nelson (1970) used for measuring repair expenditure is nonmerchandise receipts value.

² For instance, we do not consider forum manipulation (e.g. Dellarocas 2006). While these effects are likely present, they play less of role in our analysis since we focus on making statistical inferences from large numbers of reviews where the fluctuations 3

Figure 1. Original Classification by Nelson (1970, page 319)

Since Nelson (1974), critics have attacked the dichotomous classification of products as either “search” or “experience” as being insufficiently rich (see e.g., Caves and Williamson 1985) which has encourage a shift to classification systems where products exist on a continuum between search and experience (Nelson 1981, Sheffet 1983). However, even with this more fine-graining characterization, it is common to label products as either search or experience based on their dominant attribute (Klein 1998) and even the modern empirical literature has continued to use dichotomous classifications (Senecal and Nantel 2004, Huang et al. 2009, Mudambi and Schuff 2010) and many still use the original classification from Nelson (1970). In this paper we will refer to product type as the position on this continuum, and our goal is to use fine-grained review data to provide an objective measure of this position. Thus our primary research question is:

- *How can online product reviews enable a precise measurement of product type?*

Our theoretical approach is based on consumer utility theory, adopting a similar approach to prior work that has examined the dynamics of online reviews (e.g. Li and Hitt 2008, Hu et al. 2009, Sun et al. 2012, Godes and Silva 2012). Consumers consider purchasing products that are a bundle of search and experience attributes to maximize their utility. Consumers have priors regarding the values of these attributes, and can seek further information about these attributes through a search process. The efficacy of this process depends on the relative proportion of attributes for which information can be easily transferred among consumers (search vs. experience attributes). Consumers then make purchase decisions and provide reviews that reflect their product satisfaction. The statistical properties of these reviews can then be used to determine how effective consumers are in resolving product uncertainty (Dimoka et al. 2012), which is directly affected by product type. In particular, the relationship between review variance and number of reviews should be directly related to the proportion of experience attributes – greater levels of experience attributes should lead to an increase in review variance as the number of reviews increases; products that are comprised of primarily search attributes should see lower variance as the number of

reviews increases. Consistent with the prior literature (Li and Hitt 2008, Sun 2012, Godes and Silva 2012), we adopt the approach that consumers are making rational product choices but may be imperfectly informed about product characteristics, and we assume that reviews are a truthful reflection of their product satisfaction.²

We apply our theoretical model to measuring product type at three levels: category, product, and attribute. We analyze 15 categories of products (4,817 products) from Amazon.com with online product review data over the time period July 2004 to February 2012, and re-categorized the product type for these product categories on Amazon. Additionally, we demonstrate empirical applications for product level and attribute level analyses with online product review data from Yelp.com and Ctrip.com. This paper contributes to theory and practice by providing an empirical framework grounded with theory support that can be directly employed by academic researchers and practitioners to understand their product type, not only at the category level, but also at the product and attribute level.

2. Literature Review

2.1 Product Type

As noted in the introduction, the concept of product type (search versus experience product) has been studied extensively in the literature. For any product, the consumer has a choice between searching and experimenting to obtain information about the product's attributes (Nelson 1970, p. 317). In this theory, rational consumers will use “experience” or sampling when the expectation of costs associated with information search becomes too high. Most of the empirical studies that involve the construct of “product type” base their rationale on Nelson’s original classification (Nelson 1970, 1974). In studying consumer behavior for search and experience products, Huang et al. (2009) used Nelson's original classification³,

² For instance, we do not consider forum manipulation (e.g. Dellarocas 2006). While these effects are likely present, they play less of a role in our analysis since we focus on making statistical inferences from large numbers of reviews where the fluctuations due to these effects are at least partially averaged out. Empirically we take forum manipulation and self-selection influences on early reviews into consideration by robustness checks (such as dropping the first 5 and 10 reviews and repeat the analysis).

³ In Huang et al. (2009), shoes, furniture and garden were categorized as search goods; while automotive, health and camera were categorized as experience goods (p. 59).

even though they acknowledged that the Internet is likely to *change* the traditional relationship between search and experience products. Similarly, Mudambi and Schuff (2010) used Nelson's original classifications and noted that “for products outside of Nelsons original list of products, researchers have *disagreed* on their categorizations” (p. 191, emphasis added). One approach used in extant studies to avoid this criticism is to sample at the extremes. For instance products, Senecal and Nantel (2004) contrasted calculators (search) with wine (experience) in studying the effect of online product recommendations on consumer choice. An alternative approach was taken by Weathers et al. (2007) who argued that product type depends on consumer perceptions and measured product type with surveys. A product is an experience product if consumers feel the need to directly use the product to make an informed decision and these perceptions are directly shaped by how easily product information can be transferred among consumers (or between firms and consumers). Although this approach may work at individual level, the sample size tends to be small, survey respondent recruitment is costly, and obtained samples tend to be biased. Other attempts to empirically test Nelson’s framework of information search and product type include Smith (1990)⁴ and Laband (1991)⁵. Until recently, however, large-scale perceptual data was challenging to obtain. The theoretical literature has evolved from dichotomous classifications to a search-experience continuum (Klein 1998, Nelson 1981) but, perhaps for practical convenience, much of the empirical literature still relies on a dichotomous distinction between “search products” and “experience products” using the measure based on offline store non-merchandise receipts data (back in 1964). Altogether, these studies suggest that there is considerable value in developing an easily implementable continuous measure of product type.

2.2 Information Content of Online Product Reviews

A substantial literature has emerged tying product reviews to market outcomes such as price or sales, for a variety of products sold online. For example, Chevalier and Mayzlin (2006) found a significant effect of

⁴ Smith (1990) tested Nelson (1974)’s framework on the effect of advertised attribute on consumer’s use of information.

⁵ Laband (1991) proposes the use of product price as a measure of product type because price reflects the expected benefits to consumers of acquiring information about product and vendor performance prior to purchase.

review valence on book sales while Liu (2006) argues that the explanatory power of review on sales mainly comes from volume of reviews instead of the valence of reviews. Other studies suggest that variance of ratings may also affect sales for some products, such as books (Sun 2012) and beer (Clemons et al. 2008). Further research has suggested that the identity or past performance of the reviewer can be important in determining how product reviews affect sales (Forman et al. 2008). While most studies were careful to control for endogeneity – better products will have more sales and better reviews regardless of any review influence – Duan et al. (2008a) suggested that at least for movies, reviews reflect rather than cause increased sales. Taken collectively, there is sufficient evidence to suggest that consumers find information in product reviews to be a valuable in making product decisions.

A related stream of literature examines the extent to which reviews actually reflect true consumer utility. Dellarocas (2006) analyzes the potential for forum manipulation (firms pay reviewers to influence ratings), albeit concludes that in equilibrium this type of manipulation will have a limited effect on review accuracy. An alternative argument for why reviews may not reflect utility is consumer heterogeneity – consumers may be truthful, but not reflective of the “average” because reviews reflect both product attributes and consumer tastes (Li and Hitt 2008) or are shaped by the interaction with other reviewers (Hu et al. 2009, Moe and Schweidel 2012). These problems may be resolved by increasing the amount of metadata on reviews such as evaluating reviewer characteristics (Forman et al. 2008) or extracting subjective attributes from the textual content of reviews (Ghose and Ipeirotsis 2010). However, many of these bias effects are most prevalent at early stages of product introduction where the marginal effect of a single review is large, and therefore can at least be partially avoided by looking at long-run ratings.

2.3 Views on Product Quality

Product quality is a heavily debated concept, for which different scholars have viewed it from different vantage points. Tuchman (1980) observed that quality is inherent: “...a condition of excellence implying fine quality from poor quality...Quality is achieving or reaching for the highest standard as against being satisfied with the sloppy or fraudulent” (p.38). On the other hand, scholars like Kuehn and Day (1962)

argued “the quality of a product depends on how well it fits patterns of consumer preferences” (p. 101). More explicitly, Juran et al. (1999) simply ascribed quality to be “fitness for use” (p. 242). In an attempt to reconcile these perspectives by reviewing the extant literature, Garvin (1984), proposed a separation of quality assessment into a “product-based approach” (objective product quality) and a “user-based approach” (subjective consumer preference) of quality. It is argued that the observable attributes of a product determine its quality. This distinction lends a vertical or hierarchical dimension to quality. For example, a memory card with faster read speed and more reliability is considered of higher quality than one that reads more slowly and has higher failure rates. In this case, quality is objective, and utility experienced by one consumer is easily transferrable to another. On the other hand, preference sees quality to be “in the eyes of the beholder”. When consumers have different preferences, the product that best fits their preferences are regarded as having the highest quality. The literature has also discussed aggregation for product quality. Since objective product quality is based on evaluation of objective indices such as product performance, reliability, conformance, durability and serviceability (Garvin 1984), it is plausible to aggregate the objective quality metric. Naturally, the mean of the quality measure reported by different consumers would be an unbiased estimator of product quality. However, aggregating or averaging consumer preferences can be problematic. Product-consumer fit is based on consumer-specific subjective quality indices such as experience attributes, features, aesthetics and perceived quality. In this respect, Theil (1971) noted the difficulty of devising an unbiased statistical procedure for aggregating wildly varying preferences of consumers. Since quality is a multi-dimensional concept (Gavin 1984), and as Archak, Ghose and Ipeirotis (2011) observed, “by compressing a complex review to a single number, we implicitly assume that the product quality is one-dimensional”, the one-dimensional view of online product reviews may be problematic when consumer tastes vary (Rosen 1974) and attributes offer idiosyncratic utility (as opposed to common utility) to different consumers (Nelson 1981). This implicit multi-dimensional nature of online product reviews, serves as our basis for devising the analytical tool to categorize products.

3. Theoretical Model

3.1 Connect Product Ratings to Product Type

In our analysis, products are defined as bundles of attributes; each attribute is either a search attribute or an experience attribute consistent with the “goods-attributes” approach to product differentiation (Tirole 1988). Based on the historical definitions proposed by previous researchers (Nelson 1970, 1981; Klein 1998; Weathers et al. 2007). We define type of product attribute from an “information transfer”

perspective. A search attribute is defined as one for which *information can be transferred (from the seller or other previous consumers to the current consumer) without ambiguity and for which a consumer can form rational expectation about the utility prior to consumption based on the information given to her*⁶.

Examples of search attributes for a camera lens would be focal length, maximum aperture, size, weight, or price. In other words, search attributes offer objective information from which consumers can form reasonable expectation of their utility from consuming the attributes. An experience attribute is defined as one for which *a consumer's utility depends on the degree of match between himself and the attribute, and experience attributes offer idiosyncratic utility (Nelson 1981), which is difficult to form a precise expectation based on the readily available information*. For a camera lens experience attributes might be durability, sharpness, or bokeh. A product is usually on the continuum between a pure search good and a pure experience good.

Our major observation is that for search attributes, customer experience should reflect the underlying attributes since they are based on common information and preferences (at least among consumers who have chosen the product). Consumers’ ratings would be comprised of a common component, reflecting the true quality, and a random (across consumers and products) idiosyncratic component reflecting their tastes. Therefore an average of a large number of truthful reviews for products that are mostly search attributes should converge to a reflection of true quality, with most people rate at or closer to the “mean rating”, or “true quality”. This type of convergence will not happen as quickly or at all for goods that are

⁶ In this paper, we use “her” to refer to a consumer.

largely comprised of experience attributes since the common component is a much smaller proportion of total variance. It is this distinction that forms the basis of our model, which we formalize below.

3.2 The Model

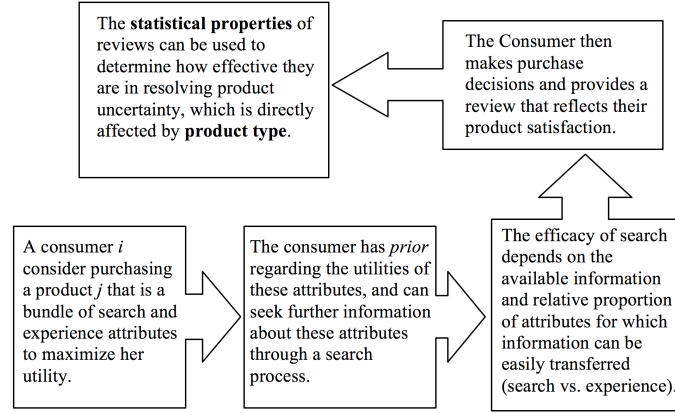


Figure 2. Consumer Product Search and Rating

We consider the following setting for a product. A product (indexed by j) may have one or more search or experience attributes. In general, vertically differentiated attributes are search attributes, while horizontally differentiated attributes, which depend on consumer tastes, may be either search or experience attributes. Those horizontally differentiated attributes, such as color, which information can be easily transferred, will fall into search attributes, while those rely on “experience” to know true utility, such as flavor, will be experience goods. When product uncertainty comes from different consumer preferences, the distribution of ratings will be related to the distance the consumers’ taste over the experience attributes. Therefore, using the same set up as Sun (2012) and a variant of Hotelling’s transportation model (1929), we model consumers taste space as a circle space with radius of I on which the component of experience attributes of a product is located at the center point. Consumers are uniformly distributed in the circle space: A consumer’s location represents her ideal type of attribute for this product in the taste space. Suppose a product has τ experience attributes (τ refers to the relative weight of experience attributes as opposed to search attributes in a product). If a consumer (indexed by i) with distance x_i from the product made the purchase, her utility derived from that experience attributes

would be $\tau * (1 - x_i) * t_j, x_i \in [0, 1]$. Since our analysis focuses on one market at a time, we normalize $t_j \equiv 1$. Suppose the consumer i , after information search for the needed product, purchased product j in an online market at price p^7 and submitted a rating r_{ij} after consuming the product. The rating would be determined by two sources of utility the consumer receives from the product: common utility q_{ij} that reflects the product's quality⁸; and idiosyncratic utility $\tau_j * (1 - x_{ij})$ that reflects her taste (Nelson 1981). Relating to the product quality literature (Garvin 1984), since objective quality is common knowledge, while consumer-specific quality is related to consumer preference and product-consumer fit, we relate search attributes to product objective quality, and experience attributes to consumer-specific subjective product quality. Similar to Li and Hitt (2008), total utility derived from the two dimensions of quality of product j is written as:

$$U(q_{ij}, t_{ij}) = q_{ij} + \tau_j * (1 - x_{ij}) - p_j \quad [1]$$

The consumer would post a rating based on her satisfaction, which, in accordance with the expectation-disconfirmation framework (Rust et al. 1999, Anderson and Sullivan 1993), is a function of the consumer's net utility received from the product $U(q_{ij}, t_{ij})$, and the degree to which her expectation of utility $E(U)^{prior}$ was disconfirmed, assuming a linear equation, the rating of consumer i for product j can be written as:

$$r_{ij} = U(q_{ij}, t_{ij}) + (U(q_{ij}, t_{ij}) - E(U)^{prior}) - p_j \quad [2]$$

Note that we are assuming in this framework that the reviewer truthfully reports their experience as

⁷ In this model, we consider p to be fixed.

⁸ Note that objective quality of different copies of the same product may vary for different consumers due to uncertainty in production or distribution process. For example, it is very common to find that for the same camera lens (used by the same consumer), some copies may produce sharp images and while some do not; besides, shipping process may also have an effect on product objective quality.

described by Equation [2].⁹ Since consumers form their expectations of utility from the product based on *prior* ratings, usually the valence of the average rating, and the variance of prior ratings¹⁰, therefore,

$E(U)^{prior} = \sum_{i=1}^{n-1} r_i = \bar{r}_{i-1}$, to consolidate the equations, we obtain:

$$r_{ij} = 2 * [q_{ij} + \tau_j * (1 - x_{ij})] - \bar{r}_{i-1,j} - p_j \quad [3]$$

Therefore, with a sample of reviews we obtain a distribution, assuming the utility expected from prior mean quality does not affect the utility of a subsequent consumer, we obtain the variance of first i ratings sample as:

$$Var(r_i) = 4 * (\sigma_q^2 + \tau^2 * \sigma_x^2) - 8 * \tau * Cov(q, x) + Var(\bar{r}_{i-1}) \quad [4]$$

Therefore, we propose that a product's review rating variance to come from two sources of product uncertainty. First, σ_q^2 represents product quality uncertainty (due to variations in objective product quality of the search attribute), which results from imperfect production/distribution/shipping process, and it should remain stable over time. Second, σ_x^2 represents product-consumer fit uncertainty (due to consumer preference variations for the experience attribute). Third, $Var(\bar{r}_{i-1})$ is the variance of the sample mean of the first i ratings. We assume that search and experience attributes are independent dimensions of a product (Nelson 1981), so $Cov(q, x) = 0$. This is consistent with our definitions¹¹. Under the assumption that the search attribute and experience attribute are independent, Equation [4] can be simplified as:

$$Var(r_i) = 4 * (\sigma_q^2 + \tau^2 * \sigma_x^2) + Var(\bar{r}_{i-1}) \quad [5]$$

3.2.1 Dynamics of the Variances of Sample Mean

⁹ This assumption would be violated with reviewer self-selection or significant forum manipulation. We therefore focus our empirical analysis of long-run data where individual reviews or reviewers have limited influence, yet the connection between rating variance and number of reviewers is still observable.

¹⁰ The distributional information of ratings is usually visible and clear on the product page, in websites such as Amazon, eBay, Yelp, Ctrip, Dianping, etc., especially the average rating of the previous consumers. For example, <http://tinyurl.com/8dbq2ub>.

¹¹ If experience attributes were correlated with search attributes we could project the experience attribute onto the set of search attributes to get a new experience attribute that was orthogonal to the search attributes.

When consumers have same preferences (or the product has no experience attributes), $\sigma_x^2 = 0$. In this case, while the first term is constant, but the second term, which involves the average of a true value plus a customer-specific noise term which is independent. Since it is reasonable to assume the mean and variance of ratings are bounded, the conditions for the law of large numbers is satisfied and the term indicating the variance of the sample mean $Var(\bar{r}_{i-1})$ will see a decrease as number of ratings increase.¹²

Since if a statement is true, its contrapositive is always true, we come to our first proposition:

Proposition 1. If, as number of reviews increases, the variance (or standard deviation) of the ratings does not decrease, a product has at least one experience attribute.

3.2.2 Consumer Preferences

There are two ways in which variance can be maintained as the number of reviews increase. The first is consumer preferences. For products where experience attributes are dominant, tastes may vary widely and a larger number of reviews suggest participation by more diverse groups. For example, by examining the motion pictures, Holbrook (1999) suggests that ordinary consumers and professional critics do emphasize different criteria in the formation of their tastes. The variance may increase because as more people submitting product reviews, it is likely that the sample is drawn from more different groups leading to the possibility that the variance of the distribution of an individual review could increase or remain the same. Although mean rating would converge regardless, variance will only stop increasing once no more heterogeneous groups are included (number of heterogeneous groups are based on number/weight of experience attributes τ_j). However, variance should not go down, it only stabilizes when number of reviews reaches a threshold. Therefore for products that are dominated by experience attributes, increase in number of reviews may not reduce variance for the mean, instead, it may increase it for certain period of time.

3.2.3 Search Efficacy

¹² Specifically, consider online product review as a discrete random variable whose value r_1, r_2, \dots, r_n are drawn from an independent rating process, with expected value μ , and finite variance. Let $S_n = r_1 + r_2 + \dots + r_n$, then for any ε , based on the law of large numbers (L.L.N), we have $P(|S_n/n - \mu| < \varepsilon) \rightarrow 1$, as $n \rightarrow +\infty$.

The second way to explain dynamics of variances is consumer search efficacy. As we argued earlier, consumers have priors regarding the values of product attributes, and can seek further information about these attributes through a search process. The efficacy of this process depends on the relative proportion of attributes for which information can be easily transferred among consumers (search vs. experience attributes). Since search attributes are more transferrable from one consumer to another, it is likely that later reviewers benefit from the early reviews and make a better choice (therefore submitting a review that is less likely to deviate from the mean). Below we try to model this dynamic process.

The theoretical argument here is based on our search efficacy logic. The first consumer on the market only has information on the product from sources such as manufacturer or website descriptions. Later consumers obtain information from prior consumers' reviews and other sources. Assuming a simple scenario where a consumer writes a comprehensive review of the product's attributes after consumption. And this review may be searched by late consumers for further information about the product. In a simple set-up, each review contributes v and $\tau \cdot v$ information, respectively for search attributes and experience attributes. After a searching process, consumer will be able to gather information on search attributes, forming a more precise estimate of search qualities by incorporating the information v . According to the *L.L.N.*, the mean of the first N ratings will converge to the true quality, as N increases to infinity.

Therefore, late consumers of experience goods are less likely to make purchasing errors because an additional review helps them form a more precise expectation of product quality. It is possible that consumers are also less likely to see any "pleasant surprises" due to the increase in precision of expected quality. And as rating is a perceived utility (satisfaction) measure, it is plausible to argue that when the precision of quality expectation increases, consumer expectations are more likely to be confirmed, and the rating she posts will be less likely to deviate from her *prior* expectation (average rating she observes).

Hence σ_q^2 will decrease.

However, since experience attributes are not easily transferred from one consumer to another by definition, $\tau \cdot v$ may not be as informative to the late consumers as compared to the case of search

attributes, and this piece of information may even be misleading since “one man’s meat may be another man’s poison”. The opposite argument about search goods would apply here that since consumers’ expectation of utility from the product will be less precise, which is likely to lead them to reach uncertain decisions and even purchasing errors. We show the utility view as follows from the search efficacy perspective.

$$U(q_{ij}, t_{ij}) - E_{ij}(U) = v * (1 - \tau_j) * i_j \quad [6]$$

Consolidate Equation [6] with Equation [3], we have:

$$\sigma_{xij}^2 \propto \text{abs}(r_{ij} - \bar{r}_{i-1,j}) \propto v * (\tau_j - 1) * i_j \quad [7]$$

Based on Equation [7], we find $\frac{d\sigma_{xij}^2}{di} \propto \tau_j$, which indicates, as experience attributes increase, the next rating will be more likely to deviate from the average; and therefore the more likely the variance of the first i ratings will increase.

Proposition 2. As number of reviews increases, the more the variance (or standard deviation) of rating increases, the more likely experience attributes dominate this product.

4. Empirical Applications

We present our empirical applications in this section. Our empirical task is to classify products based on the weight of experience/search attributes using the relationships derived earlier. The key challenge in the empirical application is to translate the propositions to the empirical data and construct appropriate test statistics to rank product type based on the dominant attributes (experience vs. search) of a product. We consider applications at three levels of aggregation: category (e.g., clothing, computers, etc.), product/merchant (e.g., Nikon D800e digital SLR, Red Lobster Restaurant in Philadelphia, etc.) and attribute (e.g., location, ease of use, taste, etc.).

4.1 Data

We collect data from three sources: Amazon.com (covering July 2004 to February 2012), Yelp.com

(covering October 2006 to October 2011) and Ctrip.com (covering May 2003 to May 2010).¹³ Our sample therefore includes products as well as services, and the Ctrip data also allows us to look at attributes in addition to the product as whole. For each review, we collect the time stamp (date), sequence (sorted by the website), and the star rating (an integer between 1 and 5).

Amazon.com

We collect online product reviews data from Amazon.com, which is one of the largest and generalist online retailers in the US. We obtained all the reviews for 15 product categories pre-defined by Amazon and often sampled by researchers, including *video games, TV, music instruments, digital cameras, laptop, keyboards, camera lens, android apps, beauty products, books, health products, software, memory card, hard drive, and GPS*. Amazon data is considered because Amazon provides a relatively fine-grained category; therefore we will observe less heterogeneity within each category. Table 1 summarizes the number of products and observations in each category in our Amazon data set.

Table 1. Summary Statistics for Amazon Product Samples

#	Category	Number of Products	Total Observations
1	Android Apps	880	48400
2	Videogame	648	28751
3	TV	552	24918
4	Digital Camera	473	20285
5	Music Instrument	395	16575
6	Books	384	15889
7	Laptop	275	8089
8	DVD	176	9153
9	Camera Lens	152	7146
10	Memory Card	151	6726
11	Keyboard	137	6900
12	Software	133	6367
13	GPS	98	4945
14	Hard Drive	85	3891
15	Beauty Products	70	2742

¹³ We hired a professional software programming company to produce two customized web data crawlers that allowed us to download and auto-parse data from Amazon.com and Yelp.com. Ctrip data was crawled with the help of a PhD student.

Yelp.com

Yelp.com provides consumer reviews for local services and has extensive coverage of restaurants. We sample the local services reviews of restaurants in the 50 largest cities¹⁴ in the U.S. to obtain an adequate number of reviews from each restaurant. We collected the reviews for 5,557 restaurants with a total of 249,199 observations of ratings.

Ctrip.com

Ctrip.com is a large online intermediary for travel in China, which holds ratings for reputable hotels in China. One reason we include this data set in our paper is that Ctrip is from another country (China vs. the U.S.) and culture (collectivist vs. Individualist) (Hofstede 1983), and therefore making our results more generalizable. The Ctrip data also include three aspects of rating information of a hotel: hotel overall rating, location rating, and service rating, which allows us to perform the analysis at the attribute level. The Ctrip data set is comprised of 5,439 large hotels in China with a total of 901,482 observations of ratings on the above three dimensions of hotel quality.

For each product in our sample we extract the entire review history (rating, time, text) and order them by time stamps. To avoid potential unwanted effects of review bias or self-selection, we retain products that have at least 25 reviews, which is a common threshold for defining a “large” sample in a statistical sense (Stock and Watson 2003), therefore we retain products that have more than 25 reviews. Using time stamps,¹⁵ we number each review in sequence that is captured in the variable SEQ_i , which is order of the i th review (Godes and Silva 2012). We then calculate two additional variables that hold the rolling variance and standard deviation for all reviews up to the i^{th} review ($i=1, 2, \dots, n$). For the purposes of

¹⁴ The choices of cities are based on the size using population as selection criteria. We referred to the US census bureau for the population estimates. Source: U.S. Census Bureau. Web: www.census.gov.

¹⁵ We were not able to identify the order of reviews that arrived on the same day, but the marketplace/portal allows sequential ordering (sort by date function), and we follow the marketplace/portal’s ordering. An alternative approach to deal with potential “ties” in the sequence for reviews that arrive on the same date is provided by Godes and Silva (2012), where authors defined the variable $ORDER$ as: $ORDER(d') = \sum_{d < d'} N(\Delta_d) + 1$, where $N(S)$ is the cardinality of set S . The shortcoming of this method is that, econometrically speaking, it is not a true dynamic panel. We employed same approach as an robustness check, and there is no change in coefficient estimates, in terms of magnitude and significance.

exposition, we retain the original product categorization from each review site, although we recognize that this could introduce some noise in our category level analysis to the extent that this deviates from the ways consumers categorize products. These groupings are not an issue for our product or attribute level analyses.

4.2 Model Specification

Our primary model relates variance of ratings (capture as the rating standard deviation up to the i^{th} rating: STD_i) to the position of the review in the review sequence (variable $SEQ=i$). This model can be estimated at a category level, or at an attribute-level. Since the data set pools multiple products (or product-attribute contributions), we also include fixed effects (α_j) for the cross-sectional unit in the analysis. Standard deviation of the first i ratings is defined and measured as follows:

$$STD_{ij} = \left\{ \sum_{t=1}^i (r_{ij} - \mu_{ij}) \right\}^{1/2} \quad [8]$$

Where r_{ij} is the consumer i 's posted rating for product j of the rating, μ_{ij} is the mean of first i ratings for product j .

For a dataset including j products ($j=1 \dots J$) in a category, the estimating equation is therefore:

$$STD_{ij} = \beta_1 * SEQ_{ij} + \alpha_j + \varepsilon_{ij} \quad [9]$$

Note that the unit of observation is the review so the sequence number is just the position of the review ($SEQ_{ij} = i$). β_1 captures the estimated effect of rating sequence on the variance of the sample of first i ratings, ε_{it} captures individual random errors. Although the dependent variable has serial correlation and is not independently distributed, the estimates are consistent and unbiased. Estimates become efficient when SEQ is large. According to Hamilton (1994), in the presence of only one time trend parameter, any standard OLS χ^2 test of the null hypothesis can be calculated and interpreted in the usual way. And the test is asymptotically valid for any hypothesis about any subset of the parameters (p. 472).

4.3 Category Level Analysis (Amazon Data)

Using same approach as Li and Hitt (2008), we dropped the first 5 reviews to mitigate the influence of strategic review forum manipulation and self-selection bias (Dellarocas 2006).¹⁶ This approach will be used in all subsequent analyses. In Table 2, we report estimates of Equation [9] for each Amazon category. We consider ordering the categories from largest to smallest in terms of the coefficient estimate (β_1) of the variable *SEQ*. A larger value implies more experience attributes, while a smaller (especially negative) coefficient suggests the products are dominated by search attributes. The ordering seems intuitively reasonable at least on the boundaries with memory cards being the most “search” related and beauty products being the most experience related. Due to the exploratory nature of the empirical analysis, we only suggest three cut-off points for the 15 products we obtained based on the estimated coefficients. First cut-off point is what we label as dominant experience products, which exhibit strongly significant increase in variance over the increase of sequence of reviews: *Beauty products, TV, Music Instrument, Video Game, Camera Lens, Keyboard, Book, Android App and Software*. We categorize the second batch of products as experience products, which exhibit increase in variance over the increase of sequence of reviews (but the coefficient estimate of *SEQ* is smaller than 0.002): *DVD and Digital Camera*. The third category of products does not show significant change (either increase or decrease) over the increase of sequence of reviews. We categorize them as search products: *Hard Drive and Laptop*. The last group of products exhibits significant decrease over the sequence of reviews: *Portable GPS and Memory Card*. We categorize them as dominant search products. Also, we acknowledge that although product level fixed heterogeneities in variance over time is controlled in the estimation, due to the coarse pre-defined categories, there might be cross-products unobserved heterogeneities. A potential way of mitigating this issue is to sample product reviews for the most fine-grained categories (use digital point-and-shoot cameras instead of a broad “digital cameras” category) to enhance the precision of the coefficient estimates and standard errors. We discuss this in limitations and future research section.

¹⁶ We repeated all of our analyses with the full data and find that while these effects do change the coefficients, the ordering of our results remains unchanged.

Table 2. Estimation Results in Amazon Data (DV= STD_{ij})

Category	Coefficient	Standard Error	Within R^2	F Statistic	Classification
Beauty Product	0.0043***	0.0016	0.07	7.32	Dominant Experience
TV	0.0038***	0.0004	0.1	88.73	Dominant Experience
Music Instrument	0.0037***	0.0005	0.09	53.17	Dominant Experience
Video Game	0.0034***	0.0003	0.11	113.3	Dominant Experience
Camera Lens	0.0033***	0.0006	0.09	24.03	Dominant Experience
Keyboard	0.0032***	0.0005	0.12	30.12	Dominant Experience
Book	0.0026***	0.00049	0.05	27.28	Dominant Experience
Android App	0.0023***	0.00023	0.06	98.9	Dominant Experience
Software	0.0022***	0.0007	0.05	9.52	Dominant Experience
DVD	0.0016***	0.0004	0.04	14.57	Experience
Digital Camera	0.0013***	0.0003	0.02	16.61	Experience
Hard Drive	0.0004	0.001	0.002	0.12	Search
Laptop	-0.0002	0.0005	0.002	0.06	Search
Portable GPS	-0.0011***	0.0003	0.04	9.35	Dominant Search
Memory Card	-0.0013*	0.0007	0.02	2.94	Dominant Search

Note: Product fixed effects estimation employed;

Cluster robust standard errors in parentheses, standard errors adjusted for all clusters (Groups);

Coefficients significant at *** $p < 0.001$, ** $p < 0.01$, * $p < 0.05$

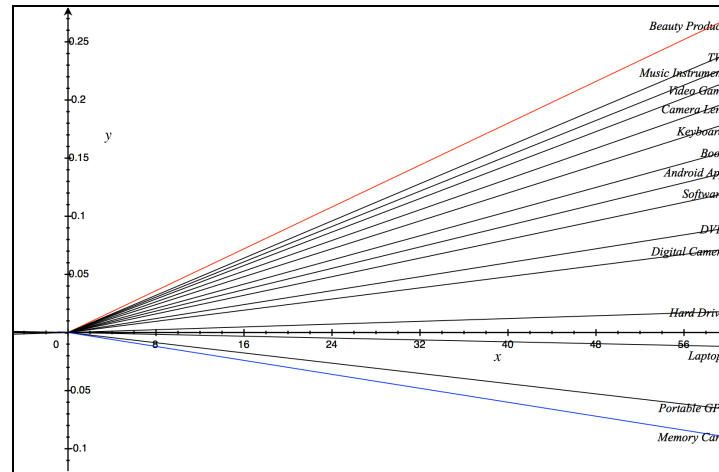


Figure 3. Empirical Observations of the Regression Line for Different Hotel Attributes

4.4 Product Level Analysis (Yelp Data)

Our Yelp data is entirely comprised of restaurants that are likely to be heavy on experience attributes.

This is confirmed (within the context of our model) by Figure 4, which shows that in aggregate, variance is increasing with sample size. This is confirmed with an aggregate regression shown in Table 3 – there is

a positive correlation between review variance and number of ratings, consistent with our model and conjecture that restaurants have substantial experience attributes. Nonetheless, it would be useful to be able to make finer grain distinctions between different products within categories, which is objectively possible with our method, but challenging with product-based classification schemes. Even within restaurants there are services that are essentially standardized by design (e.g., McDonalds), restaurants that have a strong similarity but allow for some individual variance (e.g., pizza, ethnic restaurants or sports bars), and those that are likely to be highly idiosyncratic (e.g., fine dining). The same model applied at the category level can be applied at lower levels (groupings within category) or even at the product level (individual restaurant).

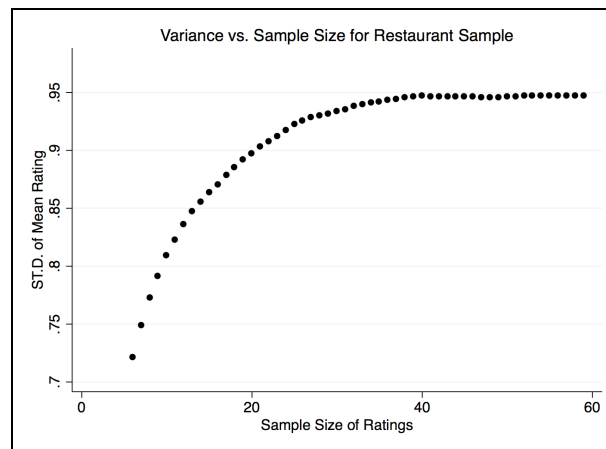


Figure 4. A Plot of the Data Set for All Restaurants in the Sample

Notes. Panels (a) and (b) depict the a scatter plot of “sample size of ratings” and “variance of mean rating”, each point in the panel represents the average variance of mean ratings across all restaurants and reviews at the value of *SIZE* on the *x* axis (starting from 6 reviews). Thus, for example, the first point is the average variance of mean ratings for the first 6 reviews of all the restaurants. This scatter plot has not controlled for other effects, such as the lagged period’s variance on this period.

Table 3. Fixed Effects Estimations in Yelp Data (DV=*STD*, Restaurant Sample)

	OLS	FE
VARIABLES	<i>STD</i>	<i>STD</i>
<i>SEQ</i>	0.00318*** (4.70e-05)	0.00318*** (8.78e-05)
Constant	0.813*** (0.00161)	0.813*** (0.00263)
Observations	170,586	170,586
R-squared	0.029	0.130
Number of Merchants		4,374

Note: For the Fixed Effects Model -Cluster robust standard errors in parentheses, standard errors adjusted for all groups

Coefficients significant at *** $p < 0.001$, ** $p < 0.01$, * $p < 0.05$

Table 4. Estimations for Individual Restaurants at Yelp.com

	Fastfood ¹⁷	Dinner Restaurant ¹⁸	Pizza ¹⁹
VARIABLES	<i>STD</i>	<i>STD</i>	<i>STD</i>
<i>SEQ</i>	-0.000263*** (7.76e-06)	0.000309*** (6.37e-06)	-2.07e-05 (3.46e-05)
Constant	1.314*** (0.00526)	0.643*** (0.00472)	1.110*** (0.0147)
Observations	1,007	1,151	629
R-squared	0.760	0.811	0.001

Note: Robust standard errors in parentheses

*** $p < 0.01$, ** $p < 0.05$, * $p < 0.1$

In Table 4, we report results from different subgroups of the restaurant category. Consistent with our prior the coefficient is strongly negative for fast food (mostly search), moderate for pizza (a mix of search and experience) and positive for the “dinner restaurant” category (mostly experience). These results can be extended down to the individual facility. For instance, we find similar results restricting our analysis to the 123-Eatery Restaurant in Minneapolis. However, even within these finer grained categories, there is still heterogeneity – for instance, the Bistro Bar in Miami has a positive coefficient in our model,

¹⁷ The Yelp page is: <http://www.yelp.com/biz/cin%C3%A9bistro-miami-2>

¹⁸ The Yelp page is: <http://www.yelp.com/biz/112-eatery-minneapolis>

¹⁹ The Yelp page is: <http://www.yelp.com/biz/galactic-pizza-minneapolis>

suggesting that reviews begin to better reflect preferences as more reviews are received consistent with this being a search product.

4.4 Attribute Level Analysis (Ctrip Data)

The Ctrip data has detailed information about three attributes of hotel: location, service and perceived overall quality. Clearly location is objective and therefore a search attribute, while service is likely to be increasingly influenced by experience attributes. Perceived quality is an overall assessment so is likely to be a blend of both. In Table 5, we utilize the same models as before applied to the data on each of the three attributes. Overall, we find that the results are consistent with our expectations – the coefficient is positive for service, negative for location and essentially zero for the aggregate.

Table 5. Fixed Effects Estimation Results in Ctrip Data (DV= STD_{ij})

VARIABLES	(1) Service Rating	(2) Location Rating	(3) Overall Rating
<i>SEQ</i>	0.000123*** (2.66e-05)	-0.000119*** (1.64e-05)	2.20e-05 (2.09e-05)
Constant	0.882*** (0.00244)	0.749*** (0.00151)	0.803*** (0.00192)
Observations	299,259	299,275	299,275
R-squared	0.009	0.013	0.000
Number of productid	5,439	5,439	5,439

Note: Cluster robust standard errors in parentheses, standard errors adjusted for all clusters (Groups)
Coefficients significant at *** $p < 0.001$, ** $p < 0.01$, * $p < 0.05$

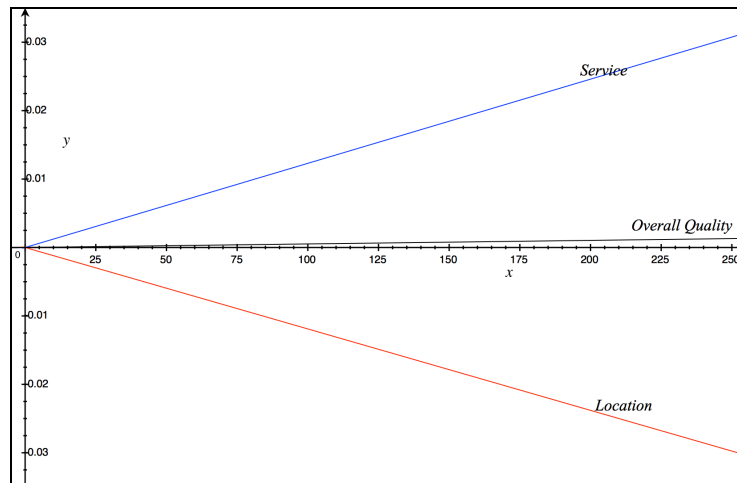


Figure 5. Empirical Observations of the Regression Line for Different Hotel Attributes

5. Discussions and Implications

5.1 Key Findings and Contributions

This study tries to classify products based on the dynamics of review variances and offers insights for research and practice. By leveraging theories from economics and statistics, we build an empirical framework, which offers an analytics tool to leverage the large amount of free micro-level user-generated data to categorize product based on the paradigm of experience versus search attributes. This research provides several unique contributions: first, we proposed an analytics tool to categorize products based on the statistical properties of the online product reviews. Our theoretical model indicates that, for a search product, as number of reviews increases as more consumers rate the product, variance of the mean rating will decrease. And for a product with dominant experience attributes, when number of reviews increases, the variance of the mean rating will increase, and the speed of increase depends on the weight of experience attributes in a product. Second, using the user-generated rating data from several sources (Amazon, Yelp and Ctrip), we identified and re-categorized the product type on a continuum of search-to-experience for 15 popular product categories on Amazon, restaurant subgroups on Yelp, and product attributes for hotels on Ctrip. Most importantly, this paper provides an empirical framework grounded with theory support that can be directly employed by researchers and practitioners to understand their product type, not only at the category level, but also at the product level and attribute level. That is, anyone can employ our analytical tool to know where a particular product/service or attribute stands, relatively to other products/services and attributes.

5.2 Implications for Research on Product Type

First, the theoretical results of this study offer new mechanism to classify products, which can be utilized by academic researchers in information systems, marketing and economics alike. The mechanism derived in this paper can be applied for any other online product review data that do not severely violate the assumptions. Second, the empirical method proposed in this paper translates the analytical propositions into several intuitive sets of results for both commodity goods and services. Third, our study is among a

few efforts to measure product type, which is an update to Nelson (1970)'s measure with data on non-merchandise receipts in 1960. The field has expressed that although the theoretical model on consumer information search still applies, due to the impacts of the reduced search cost (Bakos 1997) on different products are not the same, using the measure about 50 years ago (Nelson 1970) may limit the confidence of drawing accurate theoretical implications and actionable findings (Lucas 1976). For example, to avoid ambiguity, many studies sample products based on extreme examples of search or experience goods (Senecal and Nantel 2004), while some others acknowledged the measure for product type could potentially be problematic (Caves and Williamson 1985, Huang et al. 2009), especially for products outside of Nelson (1970)'s original list (Mudambi and Schuff 2010). What is more, some scholars used Nelson's original classification even though they acknowledged, "Internet is likely to change the traditional relationship between search and experience goods" (Huang et al. 2009, p. 56). In this respect, our study tries to answer the call by leveraging free and large amount of data generated by users to inform scholarly research. Future research can use either the analytics tool or the results from our study and design experiments and archival analyses involving the construct of "product type" with more confidence. Also, since we are proposing a mechanism for measuring product type, as time go by, consumers' information search behavior and information search costs may change and product type might also change. However, our analytical tool can still be applied for future data for updated evidence.

5.3 Theoretical Implications for Information Content of Online Product Reviews

In order to make reviews helpful for consumer decision-making, we argue that for different types of products, different online product review systems should be created. The information role of the product reviews' distributional characteristics also indicate that, for pure search goods, consumer text comments might be redundant information that may make the product review less informative due to information overload. Since people generally have the capacity to interpret statistical properties of everyday events (such as political polls) based on heuristics (Darke et al. 1998, Nisbett et al. 1983), especially when they

are in a scenario (online shopping task) that motivates them to require high accuracy²⁰, consumers no longer need to read many text comments because the distributional characteristics already conveys what they want to know (average rating for expectation of product quality and variance of rating for product quality uncertainty).

This study also points out the need for a multi-dimensional design of online product review systems. The current review systems (e.g., amazon.com and yelp.com) allow consumers to post a rating (An integer between 1 and 5) and a text comment to complement the rating. We feel it is useful to expand the uni-dimensional review system into multiple dimensions to separate main search attributes and experience attributes to help match consumer preference with the products attributes. For example, for a laptop computer, there can be dimensions such as build quality, processing efficiency, aesthetics, etc. Also, since consumers are bounded in rationality (Simon 1997), too much unstructured information may be detrimental to decision making (Jones et al. 2004), thereby undermining the information role of online product reviews. By correctly identifying a product's type, online product review systems could be designed towards higher consumer search efficacy accordingly. For example, a pure search product would benefit from a uni-dimensional numeric rating, while products with dominant experience attributes would benefit from more structured text comments, discussions and even product forums, which allow later consumers to identify "consumer-product fit" from heuristics.

5.4 Managerial Implications

As Villas-Boas observes, for experience goods consumers can gain further information regarding how well a product fits their preferences only by experiencing it after purchase. Therefore, for manufacturers and retailers, educating consumers about product attributes could lower consumer's pre-purchase uncertainty and enhance performance. For example, online retailers can adopt the state-of-the-art technologies such as "3D virtual experience" (Jiang and Benbasat 2007) to simulate "touch and feel" experiences for potential consumers. Presumably adopting such technologies is costly, therefore correctly

²⁰ We argue that online shopping indeed is a task that motivates consumers to achieve high decision accuracy because they are clearly driven to maximize utility in this task.

understanding product type is crucial. We also argue that for experience attributes that cannot be easily transferred, it will be beneficial to advertise to consumers the brand instead of the quality. It will also in the best interest of firms to offer a product forum in which consumers can interact with each other to find whether the product actually fits them. Another related issue is product returns for online retailers (WSJ 2008). The value of product returns (products returned for any reason within 90 days of sale) exceeded \$100 billion annually in the US (Guide et al. 2006). And returns are costly for both retailers and consumers (Hong and Pavlou 2010). Therefore, a better designed review systems and proper education could reduce consumer product fit uncertainty for experience goods and reduce product returns.

Another implication for managers of online or offline shops is that by looking at the variance trend of their products with product review data, they can get to know whether the consumers view their products with the same standard (searchable attributes), or different standards (preferences). For example, in the restaurant sample, even though 68% of all restaurants are found to have experience attributes (and therefore any restaurant is more likely to have experience attributes), there are still 32% out of all restaurants that are dominated by search attributes. Let's consider two restaurants, one only serves fast food such as KFC (search) and the other is an exotic French restaurant (experience). In the former case, to boost sales, taking a vertical differentiation strategy is optimal (e.g., the restaurant can increase volume per meal, or lower price). In the latter case, for the French restaurant that features a lot of experience attributes, only people who like the taste matter since later adopters can figure out the taste. Thus there might be no need for the French restaurant to cater to the general taste of the public; the effort of the restaurant should instead be focused on facilitating "matching".

5.5 Limitations and Future Research

This study has several limitations. First, with limited resources, we only analyze and present results based on a limited number of product categories. Therefore, the category level analysis might be noisy, for example, different digital cameras (a digital point-and-shoot versus a digital single-lens-reflex may be comprised of different levels of search or experience attributes and may exhibit different patterns in rating

dynamics). Future research could generalize our results as reviews for more fine-tuned product categories become available.

Second, fraudulent reviews may bias the estimates. Dellarocas (2006) shows that firms may strategically manipulate online reviews by hiring professional reviewers. Although fraudulent reviews have been assumed not to have an effect when review sample is large, identifying fraudulent review can improve the precision of the estimates. Some websites have designed algorithms to filter out “probable” fake reviews (such as Dianping.com). For Amazon data, one way of alleviating the problem will be using data only from consumers who have purchased from Amazon (“Amazon verified purchase”), future research could test whether there is any difference between reviews from Amazon verified purchases and other reviews. Since other marketplaces, such as eBay and buy.com, also host product reviews, future research could look at the results for the same products from different review sources.

6. Concluding Remark

There is a great hype today about the new phenomenon of social media, business analytics, and micro level data. The digital revolution and digital innovation (Brynjolfsson and McAfee 2011) have brought about the data that researchers have dreamed about a decade ago. One challenge that hangs over the air is how to make sense of and leverage the available data to inform research and practice. Resolving this challenge is bound to put forth both theoretical and empirical challenges of offering proper analytical tools. And this challenge has already caught attention of many scholars in the management science field²¹. As a type of consumer generated content, online product reviews are prime examples of the micro level data, which are free and readily available in the online space. Making sense of such data for our own research, as we continue to revise and advance our understandings of new IT phenomenon, can be beneficial. Our paper provides an initiative to make sense of online product review data for the purpose of informing both scholarly research and business practice.

²¹As a matter of fact, the *Workshop on Statistical Challenges in E-commerce Research* (SCECR) was organized as an effort to encourage IS and marketing researchers to take advantage of the newly available data to inform research and practice.

References

- Archak, N., Ghose, A., and Ipeirotis, P.G. 2011. "Deriving the Pricing Power of Product Features by Mining Consumer Reviews," *Management Science* (57:8), pp 1485-1509.
- Bain, J.S. 1956. *Barriers to New Competition*. Cambridge: Harvard University Press.
- Bakos J. Y. 1997. "Reducing Buyer Search Costs: Implications for Electronic Marketplaces," *Management Science* (43:12), pp. 1676-1692
- Brynjolfsson, E. and 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.
- Caves, R.E., and Williamson, P.J. 1985. "What is Product Differentiation, Really?" *The Journal of Industrial Economics* (34:2), pp 113-132.
- Chamberlin, E.H. 1950. *The Theory of Monopolistic Competition: A Re-Orientation of the Theory of Value*. Harvard University Press, Cambridge, Massachusetts.
- Chen, Y. and Xie, J. 2008. "Online Consumer Review: Word-of-Mouth as a New Element of Marketing Communication Mix," *Management Science* (54:3), pp 477-491.
- Chevalier, J., and Mayzlin, D. 2006. "The Effect of Word of Mouth Online: Online Book Reviews," *Journal of Marketing Research* (43:3), pp 345-354.
- Clarkson, K.W., and Miller, R.L.R. 1982. *Industrial Organization: Theory, Evidence, and Public Policy*. McGraw-Hill Companies.
- Clemons, E.K., Gao, G.G., and Hitt, L.M. 2006. "When Online Reviews Meet Hyper-differentiation: A Study of the Craft Beer Industry," *Journal of Management Information Systems* (23:2), pp 149-171.
- Dimoka, A., Hong, Y. and Pavlou, P.A. 2012. "On Product Uncertainty in Online Markets: Theory and Evidence", *MIS Quarterly* (36:2), pp 395-426.
- Darke, P.R., Chaiken, S., Bohner, G., Einwiller, S., Erb, H.P., and Hazlewood, J.D. 1998. "Accuracy Motivation, Consensus Information, and the Law of Large Numbers: Effects on Attitude Judgment in the Absence of Argumentation," *Personality and Social Psychology Bulletin* (24:11), pp 1205-1215.
- Dellarocas, C. 2003. "The Digitization of Word of Mouth: Promise and Challenges of Online Feedback Mechanisms," *Management Science* (49:10), pp 1407-1424.
- Dellarocas, C. 2006. "Strategic Manipulation of Internet Opinion Forums: Implications for Consumers and Firms," *Management Science* (52:10), pp 1577-1593.
- Dellarocas, C., Zhang, X.M., and Awad, N.F. 2007. "Exploring the Value of Online Product Reviews in Forecasting Sales: The Case of Motion Pictures," *Journal of Interactive Marketing* (21:4), pp 23-45.
- Duan, W., Gu, B., and Whinston, A.B. 2008a. "Do Online Reviews Matter? An Empirical Investigation of Panel Data," *Decision Support Systems* (45:4), pp 1007-1016.
- Duan, W., Gu, B., and Whinston, A.B. 2008b. "The Dynamics of Online Word-of-Mouth and Product Sales - An Empirical Investigation of the Movie Industry," *Journal of Retailing* (84:2), pp 233-242.
- Forman, C., Ghose, A., and Wiesenfeld, B. 2008. "Examining the Relationship between Reviews and Sales: The Role of Reviewer Identity Disclosure in Electronic Markets," *Information Systems Research* (19:3), pp 291-313.
- Garvin, D.A. 1984. "What Does Product Quality Really Mean?" *Sloan Management Review* (26:1), pp 25-43.
- Ghose, A., and Ipeirotis, P. 2010. "Estimating the Helpfulness and Economic Impact of Product Reviews: Mining Text and Reviewer Characteristics," *IEEE Transactions on Knowledge and Data Engineering* (23:10), pp 1498-1512.
- Godes, D. and Silva, J. C. 2012. "Sequential and Temporal Dynamics of Online Opinion," *Marketing Science* (31:3), pp 448-473.

- Greer, D.F. 1984. *Industrial Organization and Public Policy*. MacMillan Publishing Company.
- Gu, B and Ye, Q. 2012. "Firm's First-Step in Online Social Media – Measuring the Influence of Online Management Response on Customer Satisfaction," *Production and Operation Management*, *Forthcoming*
- Guide, V., G. Souza, L. Wassenhove, J. Blackburn. 2006. "Time Value of Commercial Product Returns," *Management Science* 52(8), pp 1-15.
- Hamilton, J.D. 1994. *Time Series Analyses*. Princeton University Press.
- Hofstede, G. 1984. "The Cultural Relativity of the Quality of Life Concept," *The Academy of Management Review* (9:3), pp 389-398.
- Holbrook, M.B. 1978. "Beyond Attitude Structure: Toward the Informational Determinants of Attitude," *Journal of Marketing Research* (15:4), pp 545-556.
- Holbrook, M.B. 1999. "Popular Appeal Versus Expert Judgments of Motion Pictures," *Journal of Consumer Research* (26:2), pp 144-155.
- Hong, Y., and Pavlou, P.A. 2010. "Fit Does Matter! An Empirical Study on Product Fit Uncertainty in Online Marketplaces," *Proceedings of the 31st International Conference on Information Systems*, *St. Louis*.
- Hotelling, H. 1929. "Stability in Competition," *The Economic Journal* (39:153), pp 41-57.
- Hu, N., Pavlou, P.A., and Zhang, J. 2006. "Can Online Reviews Reveal a Product's True Quality? Empirical Findings and Analytical Modeling of Online Word-of-Mouth Communication," *Proceedings of the 7th ACM Conference on Electronic Commerce*, pp. 324-330.
- Huang, P., Lurie, N.H., and Mitra, S. 2009. "Searching for Experience on the Web: An Empirical Examination of Consumer Behavior for Search and Experience Goods," *Journal of Marketing* (73:2), pp 55-69.
- Jiang, Z., and Benbasat, I. 2007. "The Effects of Presentation Formats and Task Complexity on Online Consumers' Product Understanding," *MIS Quarterly* (31:3), pp 475-500.
- Jones, Q., G. Ravid, S. Rafaeli. 2004. "Information Overload and The Message Dynamics of Online Interaction Spaces: A Theoretical Model and Empirical Exploration," *Information Systems Research* 15(2), pp 194-210.
- Juran, J.M., Godfrey, A.B., and Williams, D.H. 1999. *Juran's Quality Handbook*. McGraw-Hill London.
- Klein, L.R. 1998. "Evaluating the Potential of Interactive Media through a New Lens: Search Versus Experience Goods," *Journal of Business Research* (41:3), pp 195-203.
- Kuehn, A.A., and Day, R.L. 1962. "Strategy of Product Quality," *Harvard Business Review* (40:6), pp 100-110
- Laband, D.N. 1991. "An Objective Measure of Search Versus Experience Goods," *Economic Inquiry* (29:3), pp 497-509.
- Lawton, C. 2008. "The War on Returns. Wall Street Journal. Available Online at: [Http://Online.Wsj.Com/Article/Sb121020824820975641.Html](http://Online.Wsj.Com/Article/Sb121020824820975641.Html), Accessed on April 19, 2010,").
- Li, X., and Hitt, L. 2010. "Price Effects in Online Product Reviews: An Analytical Model and Empirical Analysis," *MIS Quarterly* (34:4), pp 809-831.
- Li, X., and Hitt, L.M. 2008. "Self Selection and Information Role of Online Product Reviews," *Information Systems Research* (19:4), pp 456-474.
- Liu, Y. 2006. "Word of Mouth for Movies: Its Dynamics and Impact on Box Office Revenue," *Journal of Marketing* (70:3), pp 74-89.
- Lucas, R. E. 1976. "Econometric Policy Evaluation: A Critique," *Rochester Conference Series on Public Policy*, available online at: <http://people.bu.edu/rking/REmodels/lucascr.pdf>
- Malone, T. W., Yates, J. and Benjamin R. I. 1987. "Electronic Markets and Electronic Hierarchies," *Communications of the ACM* (30:6), pp 484-497.
- Martin, S. 1988. *Industrial Economics: Economic Analysis and Public Policy*. Macmillan College.
- Moe, W. W and Schweidel, D. A. 2012. "Online Product Opinions: Incidence, Evaluation, and Evolution," *Marketing Science* (31:3), pp 372-386.

- Mudambi, S. M., and Schuff, D. 2010. "What Makes a Helpful Online Review? A Study of Customer Reviews on Amazon. Com," *MIS Quarterly* (34:1), pp 185-200.
- Nelson, P. 1970. "Information and Consumer Behavior," *Journal of Political Economy* (78:2), pp 311-329.
- Nelson, P. 1974. "Advertising as Information," *Journal of Political Economy* (82:4), pp 729-754.
- Nelson, P. 1981. "Consumer Information and Advertising," in *Economics of Information*. Springer. pp 42-82.
- Nisbett, R.E., Krantz, D.H., Jepson, C., and Kunda, Z. 1983. "The Use of Statistical Heuristics in Everyday Inductive Reasoning," *Psychological Review* (90:4), pp 339-363.
- Olney, T.J., Holbrook, M.B., and Batra, R. 1991. "Consumer Responses to Advertising: The Effects of Ad Content, Emotions, and Attitude toward the Ad on Viewing Time," *Journal of Consumer Research* (17:4), pp 440-453.
- Rosen, S. 1974. "Hedonic Prices and Implicit Markets: Product Differentiation in Pure Competition," *The Journal of Political Economy* (82:1), pp 34-55.
- Senecal, S., and Nantel, J. 2004. "The Influence of Online Product Recommendations on Consumers' Online Choices," *Journal of Retailing* (80:2), pp 159-169.
- Sheffett, M.J. 1983. "An Experimental Investigation of the Documentation of Advertising Claims," *Journal of Advertising* (12:1), pp 19-29.
- Simon, H.A. 1997. *Models of Bounded Rationality: Empirically Grounded Economic Reason*, Vol. 3. MIT Press.
- Smith, D.B. 1990. "The Economics of Information: An Empirical Approach to Nelson's Search-Experience Framework," *Journal of Public Policy & Marketing* (9), pp 111-128.
- Stigler, G. J. 1961. "The Economics of Information," *The Journal of Political Economy* (69:3), pp 213-225.
- Stock, J.H., and Watson, M.W. 2003. *Introduction to Econometrics*. Addison Wesley Boston.
- Sun, M. 2012. "How Does Variance of Product Ratings Matter?" *Management Science* (Forthcoming).
- Theil, H. 1971. *Principles of Econometrics*. Wiley New York.
- Tirole, J. 1988. *The Theory of Industrial Organization*. MIT Press.
- Tuchman, B.W. 1980. "The Decline of Quality," *New York Times Magazine* (2), pp 38-41.
- Villas-Boas, J. M. 2004. "Consumer Learning, Brand Loyalty and Competition," *Marketing Science* (23:1), pp 134-145.
- Ullmann, S. 1957. *The Principles of Semantics*. Oxford: Basil Blackwell.
- Weathers, D., Sharma, S., and Wood, S.L. 2007. "Effects of Online Communication Practices on Consumer Perceptions of Performance Uncertainty for Search and Experience Goods," *Journal of Retailing* (83:4), pp 393-401.
- Zhu, F., and Zhang, X. 2010. "Impact of Online Consumer Reviews on Sales: The Moderating Role of Product and Consumer Characteristics," *Journal of Marketing* (74:2), pp 133-148.