Examining the Impact of Ranking on Consumer Behavior and Search Engine Revenue[1]

Anindya Ghose [a], Panagiotis G. Ipeirotis [a], Beibei Li [b]

[a] Stern School of Business, New York University
(aghose, panos)@stern.nyu.edu
[b] Heinz College, Carnegie Mellon University
beibeili@andrew.cmu.edu

(Last Updated: March 2013)

Abstract

In this paper, we study the effects of three different kinds of search engine rankings on consumer behavior and search engine revenues: direct ranking effect, interaction effect between ranking and product ratings, and personalized ranking effect. We combine a hierarchical Bayesian model estimated on approximately one million online sessions from Travelocity, together with randomized experiments using a real-world hotel search engine application. Our archival data analysis and randomized experiments are consistent in demonstrating the following: (1) a consumer utility-based ranking mechanism can lead to a significant increase in overall search engine revenue. (2) Significant interplay occurs between search engine ranking and product ratings. An inferior position on the search engine affects "higher-class" hotels more adversely. On the other hand, hotels with a lower customer rating are more likely to benefit from being placed on the top of the screen. These findings illustrate that product search engines could benefit from directly incorporating signals from social media into their ranking algorithms. (3) Our randomized experiments also reveal that an "active" (wherein users can interact with and customize the ranking algorithm) personalized ranking system leads to higher clicks but lower purchase propensities and lower search engine revenue compared to a "passive" (wherein users cannot interact with the ranking algorithm) personalized ranking system. This result suggests that providing more information during the decision-making process may lead to fewer consumer purchases because of information overload. Therefore, product search engines should not adopt personalized ranking systems by default. Overall, our study unravels the economic impact of ranking and its interaction with social media on product search engines.

---

[1]Author names are in alphabetical order. We thank Jason Chan, Vasant Dhar, Sam Hui, Prasanna Tambe, Duncan Simester, Yong Tan, Russ Winer, and seminar participants at Columbia University, MIT, HBS, UT Austin, University of Connecticut, UT Dallas Marketing Conference, and NYU for very helpful comments.

1. Introduction

Over the last couple of decades, search engines have emerged as a significant channel for promoting and selling products. In information search engines (e.g., Google) the ranking of the search results is an immediate signal of the relevance of the result to the query. However, in *product* search engines, the ranking of the displayed products is often based on criteria such as price, product rating, etc. In such a setting, we may often have multiple, potentially conflicting signals given to the customer about what the ranking of the products. For example, if we rank by price, then the cheapest products sometimes have low product ratings; or products appearing on top of the list may be too expensive for the customer. Effectively consumers have to observe multiple, competing ranking signals and come up with their own ranking in their minds; in some settings, the product search engine will also generate personalized results, trying to rank the products according to the preferences of the consumer. In such an environment, we want to understand which factors influence the decision-making process of the customers and what the magnitude of the influence is. Are consumers influenced by the display ranking order, by the product rating, by price, and in what degree? How does this interplay affect the revenue that a search engine can generate?

1. 1 Related Work

In the last 10 years, the literature in ecommerce has shown the existence of a strong primacy effect in environments wherein consumers make choices among offers displayed in *information search engines* such as Google, Yahoo, or Bing. Specifically, we have learned that online position effect exists and that rank order has a significant impact on the click-through rates and conversion rates (e.g., Rutz and Bucklin 2007, Ghose and Yang 2009, Yang and Ghose 2010, Agarwal et al. 2011, Jerath et al. 2011, Rutz and Trusov 2011, Narayanan and Kalyanam 2011, Abhishek et al. 2011). These papers focused primarily on evaluating the effect of screen position on user behavior, controlling for the quality of the advertisement. However, in *product search engines*, the observed demand patterns can be influenced by the joint variation in product ratings (either professional rating or user rating) and online screen position. The first goal of our study is to examine the position effect in product search engines, conditional on its interaction with product ratings.

Search engines are beginning to adopt signals from social media sites directly into their ranking mechanism design (e.g., Bing Social Search, TripAdvisor). Recent work has found that a utility-based ranking mechanism on product search engines that incorporates multidimensional consumer preferences and social media signals can lead to significant surplus gain for consumers (Ghose et al. 2012). However, given that price was not the top priority considered in the ranking recommendation, whether such a mechanism can actually benefit product search engines is not clear, because their revenues are normally commission based. Therefore, the second goal of our study is to examine the effect of different ranking mechanisms on product search engine revenue.

Outside of search, one of the most important ways for shoppers to discover products has been through recommendation engines (Chittor 2010). However, although some online retailers use recommendation systems, many product-specific search engines (e.g., travel search engines) still do not provide personalized ranking results in response to consumer queries, presumably because these product search engine companies are unsure whether providing extra information to consumers will lead to an increase in profit. Existing research holds two different opinions on the effects of personalization. One stream of work is supportive of personalization (e.g., Malthouse and Elsner 2006, Rossi et al. 1996, Ansari and Mela 2003, Arora and Henderson 2007, Yao and Mela 2011), whereas another stream of work is a bit more skeptical (e.g., Zhang and Wedel 2009, Aral and Walker 2011, Goldfarb and Tucker 2011, Lambrecht and Tucker 2011), suggesting that although personalization can lead to higher customer satisfaction and profits, it will not work as well universally.[2] However, none of these papers have examined the effect of information availability and personalization in a search engine context. Koulayev (2010) examines consumer search behavior on travel search engines through the formation of consideration sets. Chen and Yao (2012) use secondary data to examine how the sorting and filtering tools on travel search engines influence consumer hotel search. They find these tools result in a significant increase in total search activities, but also lead to lower overall welfare due to the disproportional engagement induced by the refinement tools. With these findings in mind, our third goal is to examine how different kinds of personalized ranking mechanisms in product search engines affect consumer behavior and search engine revenues. Specifically, does allowing users to interact with the ranking algorithm to proactively personalize their search results lead to more or fewer purchases?

1.2 Contributions and Results

We situate our study in a travel search engine context, looking specifically at consumer selection of a hotel. We first apply archival data analysis to gain insights into the product rating effects and ranking effects on consumers' click and purchase behaviors. Using a panel data set from November 2008 to January 2009, containing approximately one million online user search sessions—including detailed information on consumer searches, clicks, and transactions obtained from Travelocity—we propose a hierarchical Bayesian framework in which we build a simultaneous equation model to jointly examine the interrelationship between consumers' click and purchase behavior, search engine ranking decisions, and customers' ratings.

Toward the first goal, we examine the variation in the ratings of different hotels (both hotel "class" rating and customer rating) at the same rank on the travel search engine over time. In addition, our data setting has variation in rank of the same hotel over time because the same hotel appears at different positions at different points in time. Controlling for room prices, such variation allows us to model the interaction effect of hotel class and customer ratings with rank, and to measure its effect on demand.

[2] For a good review of the stream of work on personalization, refer to Arora et al. (2007).

Toward the second goal, we examine how different ranking mechanisms affect the search engine revenue. We achieve this goal by conducting a set of policy experiments. We consider six different ranking designs: utility-based, conversion rate (CR)-based, click-through rate (CTR)-based, price-based, customer rating-based and Travelocity default algorithms. Then we estimate our model and predict future search engine revenues under each ranking mechanism.

Toward our third goal, we examine how different levels of personalized ranking mechanisms affect consumer behavior and search engine revenue. Particularly, we compare two types of personalization mechanisms used to drive the ranking of results in response to a query: *active personalized ranking* and *passive personalized ranking*. In our context, a ranking system that allows consumers to proactively interact with the recommendation algorithm prior to the display of results from a search query is classified as "active." By contrast, a ranking system that does not allow customers to interact with the recommendation algorithm is classified as "passive."

As of today, no hotel search engine has explicitly adopted a personalization-based approach to hotel ranking because they are still grappling with the issue of whether such an approach is useful.[3] Hence, to our knowledge, no archival data in any product search engine have information on the effect of personalized ranking on user behavior. Therefore, we designed randomized experiments using a hotel search engine application that we built. Our randomized experimental results are based on a total of 900 unique user responses over a two-week period via the Amazon Mechanical Turk (AMT) crowd-sourcing platform. We use a customized behavior-tracking system to observe the detailed information of consumers' search, evaluation, and purchase decision-making process. By manipulating the default ranking method and by enabling or disabling a variety of personalization features on the hotel search engine website, we are able to study the effect of personalized ranking on consumer behavior.

Our archival data analysis and randomized experiments are consistent in demonstrating the following: (1) A utility-based ranking mechanism can lead to a significant increase in the overall search engine revenue. (2) Significant interplay occurs between search engine ranking and product ratings. An inferior rank affects "higher-class" hotels amore adversely. On the other hand, hotels with a lower customer rating are more likely to benefit from being placed on the top of the screen. These findings illustrate that product search engines could benefit from directly incorporating signals from online social media into the ranking algorithms. (3) Our randomized experiments also reveal that an active personalized ranking mechanism that enables consumers to specify both search context and individual preferences leads to more clicks but lower purchase propensities and lower search engine revenue, compared to passive personalized ranking mechanisms. A plausible explanation is related to theories of consumer cognitive cost. Prior theoretical work has shown that information overload and non-negligible search costs can discourage decision makers from evaluating

[3] This finding is based on our personal communication with Travelocity.

choices, leading to a scenario where they make no choices at all (Kuksov and Villas-Boas 2010). Our empirical finding dovetails with the theoretical conclusion by Kuksov and Villas-Boas that providing more information can actually lead to fewer purchases. It is also consistent with Dyzabura (2012), who shows that consumers who do not have well-formed preferences at the start of their search may be better off with uncertainty about product attribute levels rather than perfect knowledge of the attributes of all available products. Therefore, although an active personalized ranking recommendation may help consumers discover what they want to buy, product search engines should not ubiquitously adopt it.

2. Data

Our dataset consists of detailed information on a total of 969,033 online sessions from Travelocity.com, including consumer searches, clicks, and conversions that occurred within these sessions between November 2008 to January 2009. In addition, we have the hotel-related information, such as hotel class, brand, online reviewer rating, and number of reviews. We collected customer reviews from Travelocity.com. We collected the online reviews and reviewers' information on a daily basis up to January 31, 2009 (the last date of transactions in our database). This process provides us with a final dataset containing 29,222 weekly observations for 2,117 hotels in the United States.[4]

We define an "online session" to capture a set of activities by an online user, identified by a unique cookie. In our data, a starting indicator and an ending indicator with a corresponding time stamp (provided by the company) can characterize each unique online session. More specifically, a typical online session involves the initialization of the session, the search query, the results (in a particular rank order) returned from that search query, the sorting method, the click(s) on hotel(s) if any exist, the login and actual transaction(s) if any conversion occurs, and the termination of the session. The ending indicator marks the termination of a session.

We count a "display" for a hotel if that hotel appears visible to a consumer on the web page in an online search session. Meanwhile, we count a "click" if a consumer selects the hotel, and a "conversion" if a consumer has completed the payment in that online session. We only consider sessions with at least one display.[5] A display can lead to a click, but it may not lead to a purchase. Each hotel that counts for a display is associated with a page number and a screen position, which capture the corresponding page order and

---

[4] We aggregate our data to a weekly level mainly to make them computationally tractable. For a robustness check, we have also tried using data from a daily level directly. Due to the size of the data (approximately 1 million user sessions with more than 14 million individual events [impressions, clicks, or conversions]), we randomly select 10% of the observations from our original data set. We then conduct the estimation on the random selected sample at a daily level. We find the estimated coefficients are qualitatively consistent with the ones from a weekly level. We have also selected 15%, 20%, and 25% of the observations to form different random samples. We find the results are similar. We provide the estimation results from the 10% sample at a daily level in Appendix C.

[5] In some cases, users may initiate a session and look for general travel information, such as the area of the city, rather than search for any hotels; thus no hotels will be displayed on any web page. We exclude such sessions from our analysis.

(within-page) rank order of that hotel in the search results. Notice that when Travelocity displays the hotel search results on a web page, it only shows 25 hotels per page.[6] This design restricts the rank order for each hotel within the range from 1 to 25. Meanwhile, to facilitate consumer search, Travelocity provides a sorting criterion called "Travelocity Pick" by default. It also provides multiple alternative sorting criteria: Price, Hotel Class, Hotel Name, and Customer Review Rating. To capture consumers' particular sorting preferences that may potentially influence the position effect, we include a set of control variables in our study to indicate how frequently a hotel appears in a result list under different sorting criteria. In particular, we use a vector (*SpecialSort*) that contains six control variables to capture the frequency of six sorting criteria that consumers use during their searches: default (*DFT*), price ascending (*PRA*), class descending (*CLD*), class ascending (*CLA*), city name ascending (*CNA*), and hotel name ascending (*HNA*).

In summary, each observation in our dataset contains the hotel id, week id, number of competing hotels, number of displays, number of clicks, number of conversions, average screen position (i.e., rank on the result page), average page number, and the corresponding hotel characteristics in that week. For a better understanding of the variables in our setting, we present the definitions and the summary statistics of our data variables in Table 1.

3. Empirical Model

In this section, we discuss how we develop our simultaneous model in a hierarchical Bayesian framework. Then we describe how we apply the Markov Chain Monte Carlo (MCMC) methods (Rossi and Allenby 2003) to empirically identify the effects of product quality and ranking position on consumer search and purchase behavior. More specifically, our model is motivated by the work of Ghose and Yang (2009). The general idea is as follows: We propose to build a simultaneous equations model of click-through, conversion, and rank. We model the click-through and conversion behavior as a function of hotel brand, price, rank, page, sorting criteria, and hotel characteristics (available from either the hotel search summary page or the hotel landing page, depending on the stage in a search process). The rank of a hotel is modeled as a function of hotel brand, price, sorting criteria, hotel characteristics that are available from the hotel landing page, and performance metrics such as previous conversion rate. Each function contains an unobserved error that is normally distributed with mean zero. To capture the unobserved co-variation among clickthroughs, conversions, and rank, we assume the three error terms are correlated and follow the multivariate normal distribution with mean zero. We describe our model next.[7]

---

[6] Recently, Travelocity upgraded its webpage design by showing 10 hotels per page. However, during our examination time period, the number was still 25.

[7] For a robustness check, we also tried a count data model, the Poisson Model. The qualitative nature of our results stays consistent. Due to brevity, we do not describe it in this paper. The results are available upon request.

3.1. Model Setup

First we define our unit of observation to be "hotel-week." Thus, for hotel *j* in week *t*, we use n_{jt} to denote the clickthroughs among N_{jt} displays ($n_{jt} \le N_{jt}$ and $N_{jt} > 0$). We also denote with m_{jt} the conversions among the n_{jt} clickthroughs ($m_{jt} \le n_{jt}$). We further denote with p_{jt} the probability of having a click-through and with be q_{jt} the probability of having a conversion, conditional on a click-through. The consumer decision process involves two steps: In the first step, the consumer sees a hotel displayed on the search result web page and decides whether or not to click on it; in the second step, if the consumer clicks on the hotel, decides whether or not to purchase it. Accordingly, we would expect to observe three types of events:

(1) A consumer sees a hotel, but does not click or purchase. The probability of such an event is $1 - p_{jt}$.

(2) A consumer sees a hotel, clicks through, but does not purchase. The probability of such an event is $p_{jt}(1 - q_{jt})$.

(3) A consumer sees a hotel, clicks through, and makes a purchase. The probability of such an event is $p_{jt}q_{jt}$.

Therefore, we can derive the probability of observing the joint occurrence of n_{jt} click-throughs and m_{jt} conversions, (n_{jt}, m_{jt}), to be the following:

$$\begin{aligned}\Pr(n_{jt}, m_{jt} \mid p_{jt}, q_{jt}) &= C_{N_{jt}}^{n_{jt}} \cdot (p_{jt})^{n_{jt}} \cdot (1-p_{jt})^{N_{jt}-n_{jt}} \cdot C_{n_{jt}}^{m_{jt}} \cdot (q_{jt})^{m_{jt}} \cdot (1-q_{jt})^{n_{jt}-m_{jt}} \\ &= \frac{N_{jt}!}{m_{jt}!(n_{jt}-m_{jt})!(N_{jt}-n_{jt})!} \cdot (p_{jt}q_{jt})^{m_{jt}} \cdot [p_{jt}(1-q_{jt})]^{n_{jt}-m_{jt}} \cdot (1-p_{jt})^{N_{jt}-n_{jt}}.\end{aligned} \tag{1}$$

3.2. A Simultaneous Equation Model of Clickthrough, Conversion, Rank, and Rating

We model the click-through, conversion, rank, and customer rating simultaneously in a hierarchical Bayesian framework. In particular, we divide our model into four interactive components.

(1) Clickthrough Rate Model

First, a consumer's decision to click on a hotel is based on the information available on the Travelocity search results page. Figure 1 provides a screen shot of a sample web page of hotel search results on Travelocity. As denoted in Figure 1, information that enters the consumer decision-making process includes *hotel price, hotel class, reviewer rating, review count, rank order* and *page number*. Prior literature has shown that *rank order* and *page number* are significant determinant of clicks on the results of a search engine query (e.g., Rutz and Bucklin 2007, Ghose and Yang 2009, Jerath et al. 2011, Rutz and Trusov 2011). In addition, previous studies have found that *rank* has a significant and non-linear effect in the context of keyword advertising (e.g., Ghose and Yang 2009; Agarwal et al. 2011). To account for the potential non-linear ranking effect in hotel search, we consider an additional quadratic term of rank in the model. Recent theoretical work has argued that product price affect consumer actions such as click and conversions and search engine decisions (Dellarocas 2012). De Los Santo and Koulayev (2012) and Yao and Mela (2011) have shown that user ratings affect click-through rates on search engines. Hence, we incorporate the volume and valence of reviews. Recent studies have shown that online search refinement tools such as the sorting selection menu can affect consumers' searches and intentions to purchase (Chen and Yao 2012). Therefore, to capture the effect associated with the search refinement tools and to control for consumers' particular sorting preferences, we include a vector $SpecialSort_{jt}$ that contains six control variables to capture the frequency of six sorting criteria that consumers use during the search process for hotel *j* in week *t*. Moreover, previous research has shown that product brand can influence consumers' perceptions of quality and willingness to buy (e.g., Dodds et al. 1991, Nevo 2001). Thus we include hotel *brand* dummies to control for the unobserved hotel characteristics. Finally, prior literature has demonstrated that the number of competitors in the local market can affect consumers' clicks for a product online (e.g., Baye et al. 2009). Therefore, to control for the competition in the local market, we include the *total number of hotels* in *j*'s city, H_j, as a control variable. This setting gives us the following equation:

$$p_{jt} = \frac{\exp(U_{jt}^{p})}{1+\exp(U_{jt}^{p})}$$

$$\text{where,}\quad \begin{aligned} U_{jt}^{p} = {} & \beta_{j0} + \beta_{j1}\text{Rank}_{jt} + \beta_{j2}\text{Rank}_{jt}^{2} + \beta_{j3}\text{Page}_{jt} + \beta_{j4}\text{Price}_{jt} + \beta_{j5}\text{Rating}_{jt} \\ & + \beta_{j6}\text{ReviewCount}_{jt} + \alpha_1\text{Class}_{j} + \alpha_2\text{H}_{j} + \alpha_3\text{Brand}_{j} + \alpha_4\text{SpecialSort}_{jt} + \varepsilon_{jt}. \end{aligned} \tag{2}$$

To capture the unobserved heterogeneity, we model β, the intercept and the coefficients for the time-varying variables, to be random coefficients [8]:

$$\beta = \begin{bmatrix} \overline{\beta}_{j0} \\ \cdots \\ \overline{\beta}_{j6} \end{bmatrix} + \Pi^{\beta} D_j + \begin{bmatrix} \sigma_{j0}^{\beta} \\ \cdots \\ \sigma_{j6}^{\beta} \end{bmatrix}, \tag{3}$$

where we assume each random coefficient to vary along its population mean and the hotel-specific characteristics. More specifically, D_j is a $d \times 1$ vector of observed hotel-specific characteristics. In our model, we consider three time-invariant variables that capture the hotel quality: *hotel class*, *average hotel price*, and *average reviewer rating* (i.e., *d=3*). Π^{β} is a $Z \times d$ matrix of coefficients that measures how hotel utility varies with observed hotel characteristics (i.e., *Z=7* is the dimension of vector β). Moreover, we model the unobserved error terms to be correlated in the following way:

$$\left[\sigma_{j0}^{\beta}, \ldots, \sigma_{j6}^{\beta}\right]' \sim MVN(0, \Sigma^{\beta}), \text{ where } \Sigma^{\beta} \text{ is a } 7 \times 7 \text{ covariance matrix.} \tag{4}$$

[8] As a robustness check, we have tried an alternative model setting with partial heterogeneity by allowing only the intercept and the Rank variable to be associated with random coefficients. We have considered a similar setting for the click-through model, conversion model, ranking model, and rating model. We find the estimation results are qualitatively consistent with our main model estimation results. We provide the results from the alternative model with partial heterogeneity in Appendix A.

(2) Conversion Rate Model

Second, we notice the set of features denoted in Figure 1 is the key determinant for a consumer's purchase decision making as well. Moreover, prior work has shown that price and quality, as well as the volume and valence of online reviews will affect product sales (e.g., Chevalier and Mayzlin 2006, Ghose et al. 2012). Meanwhile, several studies have shown how screen position and page number are important factors that influence consumer demand on search engines (e.g., Rutz and Bucklin 2007; Ghose and Yang 2009; Jerath et al. 2011; Rutz and Trusov 2011; Agarwal, Hosanagar, and Smith 2011). Thus we model the probability of a consumer's conversion as a function of the set of hotel price-, quality-, review- and screen position-related factors: *hotel price, hotel class, reviewer rating, review count, rank order* and *page number*. To account for the non-linear effect of ranking effect, we include the quadratic term of rank order. Based on the previous findings that market competition (e.g., Baye et al. 2009), product brand (e.g., Dodds et al. 1991, Nevo 2001) and online consumer search refinement tools (Chen and Yao 2012) are key determinants of the elasticities of demand, we include the *total number of hotels, brand, and special sort* as additional control variables. The conversion equation is written as follows:

$$q_{jt} = \frac{\exp(U_{jt}^{q})}{1+\exp(U_{jt}^{q})},$$

where
$$\begin{aligned} U_{jt}^{q} = {} & \gamma_{j0} + \gamma_{j1}\text{Rank}_{jt} + \gamma_{j2}\text{Rank}_{jt}^{2} + \gamma_{j3}\text{Page}_{jt} + \gamma_{j4}\text{Price}_{jt} + \gamma_{j5}\text{Rating}_{jt} \\ & + \gamma_{j6}\text{ReviewCount}_{jt} + \theta_{1}\text{Class}_{j} + \theta_{2}\text{H}_{j} + \theta_{3}\text{Brand}_{j} + \theta_{4}\text{SpecialSort}_{jt} + \eta_{jt}. \end{aligned} \tag{5}$$

Similar to (3), we model γ as random coefficients with the following properties:

$$\gamma = \begin{bmatrix} \bar{\gamma}_{j0} \\ \ldots \\ \bar{\gamma}_{j6} \end{bmatrix} + \Pi^{\gamma} D_{j} + \begin{bmatrix} \sigma_{j0}^{\gamma} \\ \ldots \\ \sigma_{j6}^{\gamma} \end{bmatrix}. \tag{6}$$

D_j also contains *hotel class*, *average hotel price*, and *average reviewer rating*. Moreover, we model the unobserved error terms in (6) to be correlated in the following way:

$$\left[\sigma_{j0}^{\gamma}, \ldots, \sigma_{j6}^{\gamma}\right]' \sim MVN(0, \Sigma^{\gamma}), \text{ where } \Sigma^{\gamma} \text{ is a } 7 \times 7 \text{ covariance matrix.} \tag{7}$$

(3) Ranking Model

Equations (2) through (7) model consumers' behavior of click-through and conversion. Meanwhile, we can model search engines' ranking decision. Prior research in keyword search advertising has found that both the bid price and the quality of the keyword affect ranking (e.g., Ghose and Yang 2009). Building on the previous findings along with our further interaction with Travelocity, we model the rank order of hotel *j* in week *t* as being dependent on the set of hotel price and quality characteristics. In particular, we use *the* *previous conversion rate*, $CR_{j,t-1}$, as a quality performance metric.[9] We consider the same set of control variables used in the previous consumer behavior models. The model is written as[10]

$$\begin{aligned}\ln(Rank_{jt}) &= \omega_{j0} + \omega_{j1}\mathrm{CR}_{j,t-1} + \omega_{j2}\mathrm{Price}_{jt} + \omega_{j3}\mathrm{Rating}_{jt} + \omega_{j4}\mathrm{ReviewCount}_{jt} \\ &+ \kappa_1\mathrm{Class}_j + \kappa_2\mathrm{H}_j + \kappa_3\mathrm{Brand}_j + \kappa_4\mathrm{SpecialSort}_{jt} + \upsilon_{jt}.\end{aligned} \tag{8}$$

Similarly, we model ω as random coefficients to vary along the population mean and the hotel-specific characteristics D_j, which contain *hotel class*, *average hotel price*, and *average reviewer rating*:

$$\omega = \begin{bmatrix}\overline{\omega}_{j0} \\ \dots \\ \overline{\omega}_{j4}\end{bmatrix} + \Pi^{\omega} D_j + \begin{bmatrix}\sigma^{\omega}_{j0} \\ \dots \\ \sigma^{\omega}_{j4}\end{bmatrix}. \tag{9}$$

Meanwhile, we model the unobserved error terms in (9) to be correlated in the following way:

$$\left[\sigma^{\omega}_{j0},...,\sigma^{\omega}_{j4}\right]' \sim MVN(0,\Sigma^{\omega}), \text{ where } \Sigma^{\omega} \text{ is a } 5\times 5 \text{ covariance matrix.} \tag{10}$$

(4) Rating Model

Note that customer ratings on product search engines can be endogenous and often determined by many hotel-specific characteristics, such as price, class, brand, and so on. To account for the endogeneity of rating, we model it as the fourth dependent variable in the simultaneous framework. Prior work has shown that product price and product quality affect customer ratings (Li and Hitt 2010). Therefore, we model the customer rating of hotel *j* in week *t* as being dependent on the set of hotel price and quality-related characteristics. Meanwhile, we include the screen position and sorting method of the hotel in the last period to control for the visibility of the hotel. We also control for hotel brand and the total number of hotels in the local market:

$$\begin{aligned}Rating_{jt} &= \rho_{j0} + \rho_{j1}\mathrm{Rank}_{j,t-1} + \rho_{j2}\mathrm{Rank}^2_{j,t-1} + \rho_{j3}\mathrm{Page}_{j,t-1} + \rho_{j4}\mathrm{Price}_{jt} \\ &+ \rho_{j5}\mathrm{ReviewCount}_{jt} + \chi_1\mathrm{Class}_j + \chi_2\mathrm{H}_j + \chi_3\mathrm{Brand}_j + \chi_4\mathrm{SpecialSort}_{j,t-1} + \psi_{jt}.\end{aligned} \tag{11}$$

We model ρ as random coefficients to vary along the population mean and the hotel-specific characteristics D_j. In the rating model, we consider D_j to contain *hotel class* and *average hotel price*:

---

[9] Using the prior conversion rate as a proxy for quality is similar to using the prior click-through rate (e.g., Ghose and Yang 2009). In addition, based on our communication with Travelocity, their default ranking is a function of commission based on previous revenue. Therefore, we tried alternative performance metrics such as revenue in the previous week, monthly averaged conversion rate, and monthly averaged revenue. The results are consistent across all these specifications.

[10] As a robustness check, we considered an alternative model using an ordered probit for the ranking model. We found the estimation results remain qualitatively consistent with the main model. We also conducted model fit comparisons between the different alternative models. We found the main model provides a better performance in both in- and out-of-sample predictions. The model fit comparison results are provided in Table 3.

$$\rho = \begin{bmatrix} \overline{\rho}_{j0} \\ \dots \\ \overline{\rho}_{j5} \end{bmatrix} + \Pi^{\rho} D_j + \begin{bmatrix} \sigma^{\rho}_{j0} \\ \dots \\ \sigma^{\rho}_{j5} \end{bmatrix}. \tag{12}$$

We model the unobserved error terms in (12) to be correlated in a similar fashion:

$$\left[\sigma^{\rho}_{j0}, \dots, \sigma^{\rho}_{j5}\right]' \sim MVN(0, \Sigma^{\rho}), \text{ where } \Sigma^{\rho} \text{ is a } 6 \times 6 \text{ covariance matrix.} \tag{13}$$

Finally, to capture the unobserved co-variation and the potential endogenous relationship among click-through, conversion, rank, and rating, we assume the four error terms in equations (2), (5), (8), and (11) to be correlated as follows:

$$[\varepsilon_{jt}, \eta_{jt}, \upsilon_{jt}, \psi_{jt}]' \sim MVN(0, \Omega_{jt}), \text{ where } \Omega_{jt} \text{ is a } 4 \times 4 \text{ covariance matrix.} \tag{14}$$

4. Empirical Analyses and Results

To estimate our model, we applied the MCMC methods using a Metropolis-Hastings algorithm with a random walk chain (Chib and Greenberg 1995). In particular, we ran the MCMC chain for 80,000 iterations and used the last 40,000 iterations to compute the mean and standard deviation of the posterior distribution of the model parameters.

4.1. Clickthrough Rate Model

First we present the results of the click-through model in Table 2a. All coefficients are statistically significant at the 5% level. The coefficients of both *Rank* and *Page* are negative and statistically significant, confirming a position effect does exist. A hotel that appears on an earlier page in the search results or on a higher position on the screen will receive more clicks than a hotel that appears on a latter page or on a lower position. A one-position increase in rank leads to a 10.07% increase in click-throughs on average. Moreover, we found a positive coefficient on the quadratic term of rank, suggesting the negative effect of rank on CTR increases at a decreasing rate. Consistent with theory and existing empirical findings (e.g., Baye et al. 2009), *Price* has a negative sign, showing the higher the price of a hotel, the lower the willingness of consumers to click on that hotel. *Class* has a positive sign, showing the higher the hotel class, the lower the CTR.

Interestingly, we found the interaction effect between *Rank* and *Class* is negative and statistically significant (i.e., -0.026). The interaction effect between *Rank* and *Price* is also statistically significantly and negative (i.e., -0.019). However, the interaction effect between *Rank* and *Rating* is statistically significant and positive (i.e., 0.020). These findings indicate that higher-class or more expensive hotels are more sensitive to the online ranking effect. They tend to be more adversely affected by an inferior screen position (e.g., at the lower part of the screen). On the other hand, hotels with lower online user ratings are more likely to benefit from being placed on the top of the search results, an effect that also benefits the underlying search engine that is typically paid by click-through or conversion.[11] This finding illustrates the need for product search engines to directly incorporate signals from online social media into the ranking algorithms.

4.2. Conversion Rate Model

The coefficient estimates from the conversion model are presented in Table 2b. Most of the coefficients are statistically significant at the 5% level. *Rank* and *Page* have a negative and statistically significant effect, indicating that screen position not only affects click-throughs, but also significantly affects conversions. Consumers are more likely to book a hotel that is positioned on an earlier page in the search results and at the top of a web page. In particular, a one-position increase in rank corresponds to a 5.63% increase in conversions on an average. Similarly, we found a positive coefficient on the quadratic term of rank, suggesting the negative effect of rank order on conversion rate also increases at a decreasing rate.

As expected, *Price* has a negative effect on hotel demand, whereas *Class* has a positive effect on hotel demand. The online word-of-mouth-related variables, *Rating* and *Review Count*, have a statistically significant and positive effect on hotel demand. We also found similar trends in the interaction effects between *Ranking* and *Price/Class/Rating*, suggesting hotels with a higher class and more expensive hotels are more sensitive to the online ranking effect. And hotels that receive lower ratings from users benefit more when placed on the top of the screen. The total number of hotels in a certain market, *H*, has a negative effect on hotel-level conversion rate. Intuitively, higher the number of choices available to consumers, lower the probability of buying from any given hotel. Thus, on average, the conversion rate for each hotel decreases.

4.3. Ranking Model

The coefficient estimates from the ranking model are presented in Table 2c. This third model sheds light on how search engines' ranking decisions are related to different product inherent characteristics, social media influences, and certain performance metrics such as previous conversions. Not surprisingly, we found that *Price* has a positive sign and *Class* has a negative sign. All else equal, a hotel with a higher price is more likely to appear in a better screen position. A higher-class hotel is more likely to appear in a higher screen position, after controlling for the sorting criteria. Both *Rating* and *Review Count* have a significant and negative effect, showing that hotels with a higher user rating and with more reviews are more likely to appear at the top of a page, controlling for everything else.

4.4. Rating Model

Finally, the coefficient estimates from the rating model are shown in Table 2d. The rating model allows us to account for the potential endogenous nature of the customer ratings. We found that both *Rank* and *Page*

---

[11] We found similar trends in the interaction effects in the conversion-rate model as well, which we briefly discuss in the next subsection.

have a negative and statistically significant effect, suggesting screen position is also correlated with a hotel's rating. Hotels with higher ratings are more likely to be positioned on an earlier page in the search results and at the top of a web page. We also found a similar positive effect from the quadratic term of rank, which suggests the marginal effect of ranking on rating is decreasing.

Notice that in the main model, we assume consumer evaluation (e.g., rating of a hotel, utility of clicking, or booking a hotel) is a quadratic function of the rank order. As a robustness check, we also tried using a simple linear form. We excluded the quadratic term of the rank order from the click-through, conversion, and rating models. The qualitative nature of the estimation results stays consistent. The corresponding estimation results are shown in Appendix B.

4.5. Policy Experiment: Effect of Ranking on Revenue

Previous work has shown that a consumer utility-based search engine ranking system can lead to an increase in consumer surplus (Ghose et al. 2012). However, how such a ranking system affects the search engine's revenues is unclear. Therefore, one question in which we are interested is how different ranking mechanisms would affect search engine revenues.

Toward this goal, we conduct a set of policy experiments. In particular, we consider and compare six different ranking designs based on: consumer-utility, conversion-rate (CR), click-through-rate (CTR), price, customer rating and Travelocity default algorithm [12]. We define the ranking equation in the simultaneous equation model as being based on each of these six ranking criteria to reflect different search engine ranking systems. For the consumer-utility-based ranking, we define the ranking equation based on equation (8)). For the other five ranking designs, we define the ranking equation to contain only the corresponding variable on the right-hand side. For example, in the case of the price-based ranking mechanism, we define the ranking equation to contain the price variable as the independent variable. All other control variables remain the same in each of the six scenarios.

We estimate the simultaneous equation model under each different ranking equation using data from the previous *t*-1 periods. Based on the estimates, we predict the CTR and CR correspondingly for the *t*-th period under each case. This process allows us to predict the future revenue for the search engine under various ranking mechanisms. The overall revenue for the search engine is as follows:

$$\text{Revenue} = \sum_{j=1}^{J} (CR_j * CTR_j * \text{Price}_j) \cdot \tag{15}$$

From our prediction results, we find that although the Travelocity default ranking and price-based ranking mechanisms lead to higher search engine revenue received from the top-ranked hotel, the consumer-utility-

[12] The default ranking algorithm used by Travelocity at the time of our data collection was based on a fixed commission rate (10%) of the last period revenue. Therefore, in the policy experiment we use the last period revenue as the ranking criterion to approximate the Travelocity default ranking.

based ranking mechanism leads to the highest overall revenue received from all hotels. This finding suggests that a utility-based ranking mechanism not only maximizes the surplus for consumers (Ghose et al. 2012), but also maximizes the revenue for search engines.

The main reason for this finding is likely due to the diversity provided in the utility-based ranking. Consistent with the previous results by Ghose et al. (2012), consumers prefer the diversity in the ranking results. More importantly, we find that under the utility-based ranking mechanism, consumers are more likely to click and purchase products that are ranked lower in the list, compared to all the other competing ranking mechanisms. This finding seems to explain why the utility-based ranking outperforms the others (especially the price-based or short-term revenue-based mechanisms) in the overall search engine revenue—the additional conversions received from the lower-ranked products are able to dominate the overall compromise in price. We provide the detailed prediction results in Table 4.

5. Randomized Experimental Design

Our Bayesian analysis provides important insights into the relationship between search engine ranking mechanism and consumer behavior. However, to fully understand how consumers make decisions in the product search engine context, we designed and conducted randomized experiments. Specifically, we tested the effectiveness of four ranking mechanisms and two personalization designs—active (customizable) personalized ranking and passive(non-customizable) personalized ranking—on influencing consumer behavior and search engine revenues.

In a randomized experiment, a study sample is divided into two groups: one receiving the intervention being studied (the treatment group) and the other not receiving it (the control group).[13] Randomized experiments have major advantages over observational studies in making causal inferences. Randomization of subjects to different treatment conditions ensures the treatment groups are, on average, identical with respect to all possible characteristics of the subjects, regardless of whether those characteristics can be measured. In our first experiment, we designed four treatment groups. Each group is exposed to the same search-ranking mechanism except for a different default ranking method. In the second experiment, we have two treatment groups and one control group. The control group is granted full access to the search mechanism with active personalization that allows them to interact with and customize the search engine recommendation algorithm. By contrast, the two key personalization features are disabled for the two treatment groups (which we refer to as passive personalization). Our experimental participants come from Amazon Mechanical Turk (AMT, https://www.mturk.com), which is an online marketplace used for crowd-

[13] In some cases, rather than be compared with the control group, multiple treatment groups can be compared with each other (Ranjith, 2005). We use this method in our first experimental study.

sourcing micro-tasks that require human intervention (i.e., cannot be fully automated using machine learning tools).[14] We discuss the experimental procedure in subsections 5.1-5.5.

5.1 Hotel Search Engine Design

First we designed and built a real-world hotel search engine. This application served as the main instrument for our experimental studies. The main interface of this search engine consists of three components: (1) Search Criteria: including travel destination and search context (e.g., demographics such as income, trip type, and age); (2) Sorting Methods; and (3) Resulting Hotel List: on the right-hand side as the response to (1) and (2). A screenshot of the main search interface is provided in Figure 2.

When consumers start to search for hotels, they are able to define the travel destination, income level, trip type, and age group. We classify consumer trip type into four major categories: *business trip*, *family trip*, *romantic trip*, and *trip with friends*. We classify consumer age into five groups: *17 and below*, *18-24*, *25-34*, *35-64*, and *65 and older*. Meanwhile, we provide consumers with four different sorting methods: *BVR*, *price*, *TripAdvisor.com customer rating*, and *Travelocity.com customer rating*. "BVR" denotes the "Best-Value Ranking" adapted from the utility-based ranking in Ghose et al. (2012). The value-for-money score represents how much additional value consumers can obtain from a hotel after paying the nightly reservation rate. We use the acronym BVR on the search engine to minimize the potential experimenter-expectancy bias that can accrue from displaying the full, expanded label. For each hotel listed on the right-hand side, we provide the summarized hotel information, including the hotel class (i.e., in pink stars), address, price, customer ratings from both Travelocity.com and TripAdvisor.com, and the value for the money (i.e., both in text and indicated by a vertical pink bar).

Users view the summary information in the hotel list and decide whether they want to click on a hotel's URL to acquire more detailed information. If a user chooses to click on a hotel's URL, he/she is directed to that hotel's landing page. A sample hotel landing page is provided in Figure 3. Generally speaking, the landing page consists of three components: (1) Search Criteria: similar to those on the main search page, where consumers can refine the travel destination and search context; (2) Value-for-the-Money Scores: including the hotel's overall value for the money and the breakdown value score for each hotel feature (e.g., price, location, and service and customer reviews); (3) Consumer Decision: a "buy now with 1-click" button that allows consumers to make a simulated purchase, or a "back" button that takes consumers back to the main search-result page to continue searching.

Note that the value-for-the-money score on the landing page exists in two forms: the population's *average* value score and the *personalized* value score. The former represents how much value a hotel feature provides

---

[14] Based on a pilot study, we found the AMT population is generally representative of the overall U.S. Internet population. We provide more details of the pilot study in Appendix E.

to the overall population, whereas the latter represents the personalized value to a specific consumer based on the search context and demographics. Moreover, each hotel feature is associated with a "weight" that ranges from -1 to +1, representing consumer preference from "strongly dislike" to "strongly favor." A consumer can adjust the weight of his/her preference for each hotel feature to obtain a personalized value that most closely represents his/her preference. Overall, by choosing different search criteria or/and weights of preferences, a consumer is able to personalize the ranking results provided by the search engine.

5.2. Consumer-Behavior Tracking System

To better understand the complete decision-making process, we keep track of the exact searching and purchasing behavior of users. This tracking system records the detailed information of every online activity by every consumer. For example, such activity information includes click behavior (e.g., a hotel URL being clicked, corresponding rank position, time spent on the landing page), usage of the search functions (e.g., search criteria changed, sorting methods chosen), hotel landing page browsing behavior (e.g., preference weights adjusted, search criteria changed, etc.), and purchase behavior (e.g., corresponding hotel being booked, corresponding ranking position, and sorting method). Furthermore, each activity is recorded with a time stamp capturing when the activity occurs.

5.3. Experiment I: Evaluating the Impact of the Ranking Mechanism

We now discuss the design of our first randomized experiment, which aims to examine consumer behavior and search engine revenue under different ranking mechanisms. The basic procedure is as follows. We ask the subjects to visit our hotel search engine website, conduct a hotel search using a set of randomly assigned search criteria, and make a simulated purchase at the end. The independent variable is the default ranking method. We are interested in how the ranking mechanism affects the breadth, depth, concentration, and final decision of consumer search. Moreover, we are interested in the resulting revenue for the search engine. Therefore, the dependent variables we focus on are (i) number of clicks, (ii) time spent on evaluation, (iii) number of online activities, (iv) number of conversions (0 or 1), and (v) search engine revenues.

We use a mixed experimental design. First, for the between-subjects design, we use a completely randomized setting with four treatment conditions. We manipulate the independent variable by changing the default ranking method for each of the four treatment groups. Each treatment group is exposed to a different default ranking method. We then randomly assign each subject to only one of the four groups. Meanwhile, to control for the error variance associated with individual subject-level differences, we propose a within-subjects design considering hotel search in two major U.S. cities: *New York City* and *Los Angeles*. We allow each subject to participate in two experiments corresponding to the two cities, but only in the same treatment group. We summarize the design of this study in Table 5a.

5.4. Experiment II: Evaluating the Impact of Personalization

In our second study, we examine consumers' responses to different personalized ranking mechanisms. In particular, we focus on two independent variables that capture two different levels of personalized ranking: (1) whether it allows consumers to change their personalized search context and (2) whether it allows consumers to adjust their weights or preferences for different hotel features. The dependent variables we look into are the CTR and CR at both the subject and group levels. Moreover, we are also interested in the resulting search engine revenue. As before, we propose a mixed experimental design. For the between-subjects design, we apply a completely randomized setting with two treatment groups and one control group. We define the control group as subjects who have full access to our search engine website. For the two treatment groups, everything else is the same as in the control group, except that we remove the two personalization features—the user's ability to change the search context and to adjust weights of preferences—one at a time. Meanwhile, we control for the subject-level fixed effect by using a within-subjects design, similar to that in the first study. We summarize the design of the second study in Table 5b.

5.5. Implementation

We have 900 unique user responses in the experiments, with 100 for each experimental group. We recruit users from the AMT platform. To control for quality, we allow only those AMT workers with a prior approval rate higher than 95 percent to participate in the experiments. AMT provides an approval rate for each worker based on the frequency with which buyers have approved tasks. This approval rate can provide information on the quality of the workers. Moreover, we design an additional survey at the end of the experiment asking the subjects to provide (1) a verification id that is automatically generated once the experiment is properly finished and (2) a short explanation of why they made their final decision, using at least 20 characters. This two-step process helps us avoid negligent participants who have not gone through the entire experiment seriously. With regard to the experimental procedure, we first provide a short introduction about the experiment, as shown in Figure 4. To familiarize subjects with how to use the hotel-search website, we provide a quick two-page demo of the website prior to the experiment. Figure 5 shows the final introduction page leading to the start of the experiment.

6. Results from Randomized Experiments

6.1 Direct Ranking Effect

1) Ranking Effect on Click and Purchase Propensities.

First we look into how the design of ranking mechanisms affects different aspects of user behavior on search engines. We examine the total time spent, number of online activities and number of clicks at the subject level, and the overall purchase propensity[15] from each of the four treatment groups in Study I. Table 6 shows the final purchase propensities under different ranking mechanisms. Subjects who get to see BVR as the default ranking pay more attention and display higher purchase propensities than subjects from other groups. This result is significant at the p=0.05 level based on a post hoc ANOVA test. Price-based ranking provides the second-best performance on these two dimensions, followed by the rankings based on TripAdvisor and Travelocity ratings, respectively. Moreover, this finding is consistent across the two cities, New York City and Los Angeles. This result shows how the design of ranking mechanisms affects the performance of a product search engine.

We also find a significant ranking effect at the individual hotel level. Hotels ranked at the top of the search result list received, on average, 2.39 times more clicks compared to the second-ranked hotels, and 3.42 times more compared to the third-ranked hotels. This trend stays consistent across two cities and regardless of the default ranking method. Table 7 shows the number of clicks received for hotels ranked in the top 10.

We also examine the CTR for the same hotel that appeared in different ranking positions under different default ranking mechanisms. Controlling for everything else, the same hotel in a higher screen position received significantly more clicks. For example, the "Blue Moon Hotel" in New York City received a total of 56 clicks under the BVR ranking, in which it was ranked at position 1. However, the same hotel received zero clicks under the price-based ranking, in which it was ranked 31.

2) Ranking Effect on Search Engine Revenue

Recall we are interested in how different ranking systems affect overall search engine revenues. We compute the overall search engine revenues by multiplying the unit price by the number of conversions for each hotel, and then summing over all hotels in the experiments. We provide the detailed results in Table 8.

Our experimental results are highly consistent with the policy experiment results from the previous archival data analysis (i.e., subsection 4.5). We find that price-based ranking leads to the highest search engine revenue received from the top-ranked hotel. However, BVR (consumer-utility-based) ranking leads to the highest overall revenue from all the hotels. Moreover, we find experimental evidence that under the BVR ranking, a significant part of the overall revenue comes from hotels that are ranked lower on the computer screen, which is different from the other competing ranking mechanisms.

These experimental findings support our previous policy experiment. They indicate consumers prefer the diversity in the utility-based ranking. Diversity presented in the ranking list can lead to a significant increase in conversions, especially from the lower-ranked products. Moreover, these additional conversions can contribute significantly to the overall revenue for search engines.

---

[15] The purchase propensity is defined as the number of subjects who have made a purchase, divided by the total number of subjects in each group.

6.2. Interaction Effect between Ranking and Product Rating

1) Interaction Effect between Ranking and Hotel Class Rating.

We examine the differences in CTR from different ranking positions for two different "classes" of hotels—luxury- and budget-class hotels. In particular, we look into the changes in CTR at different ranking positions for either 4- or 5-star hotels (i.e., luxury hotels) and for 3-star or lower hotels (i.e., budget hotels). We find that as one moves down from the top-ranked position to a lower-ranked position, the decrease in CTR for luxury hotels is much larger than that for budget hotels. For example, moving down from the top to the fifth position leads to a 75% drop in CTR for the luxury hotels compared to a 54% drop for the budget ones. We test different ranking positions using a robustness check and find the results to be very consistent. Table 9a shows the changes in the click-through rate of hotels when moving down from the top position to the third, fifth, and tenth position.

2) Interaction Effect between Ranking and Customer Rating.

Similarly, we also examine the differences in CTR from different ranking positions for hotels with higher customer ratings compared to those with lower customer ratings. In particular, we compare the CTR at different ranking positions for 4- to 5-star hotels, as rated by reviewers, versus 1- to 2-star hotels. We find the increase in CTR resulting from hotels moving from a lower- to a higher-ranked position is greater for hotels with a poor reputation than for hotels with good reputation. For example, moving up from the 10th-ranked position to the top position increases CTR by 245% for hotels with low user ratings compared to an increase of 83% for hotels with high user ratings. Table 9b shows the corresponding changes in the CTR of hotels moving up from the 10th position to the fifth, third, and top position.

Findings in Tables 9a and 9b provide important insights and additional support to the archival data analysis, indicating luxury hotels are more sensitive to the ranking effect and are more adversely affected by an inferior screen position. Meanwhile, hotels that receive a lower reputation from online word-of-mouth are benefiting more when placed at the top of the search results. Our findings strongly illustrate the need for product search engines to directly incorporate signals from online social media into the ranking algorithms.

6.3 Effect of Active versus Passive Personalized Ranking

1) Effect of Personalized Ranking on Click and Purchase Propensities.

Another important goal of our research is to examine how different personalized ranking mechanisms influence the way consumers behave on product search engines. In Study II, we consider three levels of personalization: active personalized ranking with full access (control group, henceforth "FULL_ACCESS"), passive personalized ranking without search context (treatment group 1, henceforth "NO_SEARCH"), and passive personalized ranking without weights of individual preferences (treatment group 2, henceforth

“NO_WEIGHT”). Table 10 summarizes the average user behavior, in terms of total time spent and total number of activities, under the three different personalization mechanisms.

We find the active personalized ranking mechanism results in more user time and more activities than the two passive mechanisms. Each user, on average, spends approximately 351 seconds and conducts 19 activities per session when exposed to active personalized ranking. This finding suggests an active personalized ranking can generate higher online engagement on the search engine. The NO_WEIGHT group with passive personalized ranking demonstrates the lowest level of user engagement. This step provides a sanity check that these different personalization features indeed influence user behavior in our experiments.

Table 11 displays the average number of clicks made by a user and the overall purchase propensity for the two different cities, under the three personalized ranking mechanisms. Interestingly, we find that a travel search engine with an active personalized ranking mechanism can attract significantly more clicks than those with passive mechanisms. However, active personalized ranking leads to a significantly lower purchase propensity. This finding is consistent across the two different cities and is interesting because one would expect the active personalized ranking mechanism to increase, rather than decrease, the purchase propensities. One possible explanation is related to consumer expectations. In most online shopping environments, consumers find active personalization especially useful because it helps them discover what they want to buy before they know it themselves. In other words, the active personalized ranking is more likely to increase sales when consumers have not planned their purchase beforehand. In our setting, we focus on the type of consumers who have planned their purchase before the search starts. Under such a scenario, the major advantage of active personalized ranking is lost on consumers because they already have in mind what they are searching for. What is worse, if the personalization results do not meet consumers’ expectations, they may easily stop the sale. This finding is in line with previous findings by Lambrecht and Tucker (2011), who show the mismatch between the specificity of the ad content and whether a consumer has well-defined preferences can lead to ineffective personalization. Another plausible explanation is related to consumers’ cognitive limitations. The ability to extensively search and change their current consideration sets under the active personalized ranking mechanism can lead to information overload during the decision-making process. As a consequence, consumers may end up being confused or frustrated and therefore skip buying completely.

Comparing the NO_SEARCH group with the FULL_ACCESS group, the additional personalization based on search context and demographics (i.e., “search-based” personalization) results in a larger negative effect on purchase propensity (i.e., 6% larger for LA and 3% larger for NYC) than when we compare the NO_WEIGHT group with the FULL_ACCESS group. This finding provides a plausible explanation: two types of personal information can apparently be used in the personalization process in our context—(i) user-identity-related (i.e., who are you?) and (ii) user-preferences-related (i.e., what do you like?). Search context and demographic information lie closer to the former category, whereas weights of location and service preferences belong to the latter. Our results suggest that when designing a personalized ranking mechanism, using the identity-related information is less beneficial, not only for privacy-preserving purposes, but also for the economic outcomes such as conversions.

The findings above are directly observed at the search engine level. To verify the effects of active and passive personalized ranking mechanisms, we conduct two further analyses at the individual-subject level.

First we consider the user-level number of clicks as the dependent variable in our analysis. The independent variables we are interested in are two dummies: *NOSEARCH* and *NOWEIGHT*, corresponding to the two passive personalized ranking treatment groups, respectively. Because the number of clicks is a nonnegative integer, we use a count data model, the negative binomial model with robust error. For estimation, we apply the maximum likelihood method. To control for the location effect, we include a city dummy variable denoting whether it is New York City or Los Angeles. Moreover, from the previous analysis, we notice the number of consumer activities drops significantly in the case of *NOWEIGHT*. Therefore, to control for the level of online attention, we include the number of total activities at subject level as an additional control variable. The results are qualitatively consistent as displayed in columns 2-4 in Table 12. Both *NOSEARCH* and *NOWEIGHT* show a significant negative effect on the number of clicks, which means the presence of personalization in search context and weights of preferences has significant positive effects on the clicks at the individual level. The ability to define their search criteria on specific contexts and to adjust their preferences toward product features leads to more clicks.

Second, we consider the user-level purchase propensity as the dependent variable in our analysis. As before, we are interested in two independent variables: *NOSEARCH* and *NOWEIGHT*. Note that in our experiment, we ask each subject to make a purchase at the end. However, subjects can still decide not to do so. Thus the purchase outcome is a binary variable: 0 or 1. Therefore, we apply the probit model with maximum likelihood method for estimation. Again, we include two additional control variables: city dummy and number of total activities. We display the results in columns 2-4 in Table 13. Both *NOSEARCH* and *NOWEIGHT* have a statistically significant positive sign. This finding suggests the presence of personalization in search context and individual preferences has significant negative effects on the purchase propensity at the individual level. This result is highly consistent with our previous analysis at the search engine level. It indicates the active personalized ranking mechanism can lead to a significant decrease in consumer purchase propensity.

2) Effect of Personalized Ranking on Search Engine Revenue

Finally, we are interested in how active and passive personalized ranking mechanisms affect the revenue for search engines. Consistent with the previous definition, we sum over all hotels in the experiment to compute the overall search engine revenue. We find the active personalized ranking mechanism can lead to significantly lower overall revenues than the two passive mechanisms in our travel search engine. This finding provides further insight that the decrease in purchases due to the improper use of the active personalized ranking strategy can result in a decrease in the overall revenue for product search engines. Thus implementing the active personalized ranking mechanism may not always be profitable for product search engines. We provide the corresponding results in the last column in Table 11.

6.4 Robustness Tests

To further test the validity of our results, we conduct two robustness tests by considering two additional situations. First we consider a setting with an even higher level of active personalization. Consumers who are randomly assigned to this setting are granted full access to active personalized ranking, as in the previous setting. Moreover, they can adjust their individual weights of preferences not only on the hotel landing page, but also on the main search page. The value score for each hotel and the corresponding BVR ranking will be adjusted instantly based on the weight preferences consumers choose on the search page. The search interface for this robustness test is shown in Figure 6.

We found a similar trend when comparing the case of active personalized ranking with passive personalized ranking. In the new setting, users tend to spend even more time (i.e., an average of 343.02 sec) and conduct even more activities (i.e., an average of 19.27 activities) on the search engine than in the two passive personalized ranking scenarios. These two statistics again serve as good manipulation checks, indicating users are indeed using the personalization features. Furthermore, the high-level active personalization leads to a significantly lower purchase propensity and lower search engine revenue compared to the two passive mechanisms. This result strongly supports our previous findings obtained from both the archival data analysis and the experiment regarding whether excess information discourages consumers from making final decisions. Improper use of the active personalized ranking mechanism can lead to a loss of profit for product search engines. The detailed results are shown in Table 14.

Second, to test consumers' behavior when they have a less structured purchase plan in mind, we consider a more general purchase situation in which, rather than having to make a planned purchase at the end of each search session, consumers can choose to leave the search session without making a purchase. For comparison, consumers who are randomly assigned to this setting receive full access to the active personalized ranking recommendation.

We found that in the case of active personalization with an "unplanned purchase," the average time users spend on the site drops to nearly half of that in the case of active personalization with a "planned purchase" (i.e., 177.01 sec vs. 351.23). However, the average number of activities in which users engage in the two cases remains similar (i.e., 18.18 vs. 19.36). Furthermore, in the case of active personalization with an "unplanned purchase," purchase propensities increase compared to the case of a "planned purchase." The results are consistent across the two cities.

This finding suggests active personalized ranking may be more effective when consumers generally do not have a well-structured purchase plan. In such cases, they are more likely to discover potentially relevant products. However, this scenario is not the case when consumers already have a clear purchase plan. Consumers can be highly discouraged and terminate the search completely if the active personalized ranking results mismatch their original expectations. This test provides additional insights into our main findings, suggesting active personalized ranking should not be adopted blindly, and the level of personalization should be carefully designed based on the search context. The detailed results are provided in Table 15.

7. Conclusions and Implications

In this paper, we focus on investigating three major issues that product search engines are increasingly facing: the direct effect of ranking mechanism on consumer behavior and search engine revenue; the interaction effect of ranking and product ratings; and what kind of personalized ranking mechanism, if any, to adopt. Toward these objectives, we combine archival data analysis with randomized experiments based on a hotel search engine application that we designed. By manipulating the default ranking method and enabling or disabling a variety of active personalization features on the hotel search engine website, we are able to analyze consumer behavior and search engine revenue under different scenarios.

Our experimental results on ranking are consistent with those from the Bayesian model-based archival data analysis, suggesting a significant and causal effect of search engine ranking on consumer click and purchase behavior. In addition to a significant surplus gain found by a previous study (Ghose et al. 2012), a consumer-utility-based ranking mechanism yields the highest purchase propensity and the highest search engine overall revenue compared to existing benchmark systems, such as ranking based on price or star ratings. Moreover, an inferior screen position tends to more adversely affect luxury hotels and more expensive hotels. Hotels with lower reputations are benefiting more from being placed at the top of the search results. This finding illustrates the need for product search engines to directly incorporate signals from online social media into the ranking algorithms. We are beginning to see much of this interplay between search and social media happening in information search engines. Google began to incorporate tweets and other social media status updates into its real-time search function, and then decided to create its own version of the Facebook Like button — the Google +1 – and have it show up in search results. In another example of the interplay between social media and search, Microsoft's search engine Bing is now incorporating Facebook updates in its results.

Our experimental results on personalized ranking show the availability of excess personalization capabilities during the decision-making process may discourage consumers from searching, evaluating, and making final choices. In particular, we find that although active personalized ranking, compared to passive personalized ranking, can attract more online attention from consumers, it leads to a lower purchase propensity and lower search engine revenue. This finding suggests personalized ranking should not be adopted blindly and the level of personalization should be carefully designed based on the search context. Our research sheds light on how consumers search, evaluate choices, and make purchase decisions in response to differences in product search engine designs. We provide empirical and experimental evidence for future studies to build on when designing an efficient ranking system and dynamically modeling consumer behavior on product shopping sites. A good ranking mechanism can reduce consumers' search costs, improve click-through rates and conversion rates of products, and improve revenue for search engines.

Our work has some limitations, some of which we are striving to address in our ongoing work. First, although the AMT platform provides an efficient and cost-friendly framework for randomized experimental design, the inherent heterogeneity in the Internet population makes controlling for subject characteristics across different treatment groups difficult. The randomization process can alleviate this concern to a large extent. However, robustness tests based on offline subjects as well would be helpful. Our current experiments focus on the type of consumers who can make, at most, one purchase in each online shopping session. To better understand the counter-intuitive finding that an active personalized ranking mechanism leads to lower conversion rates, one can extend our experimental design to make a comparison with consumers who are allowed to make *multiple* purchases in a given session. In addition, a study of how the content of consumer search, such as the length and type of search keyword, interacts with the ranking effect would be interesting. Moreover, with regard to examining the ranking mechanism, one can expand the research scope by taking into account consumers' social network neighbors' search and purchase behavior. This expansion would allow one to test the impact of social-signal-based ranking mechanisms on product search engines.

References

- Abhishek, V., K. Hosanagar and P. Fader. 2011. Aggregation Bias in Sponsored Search Data: The Curse and The Cure. *SSRN Working Paper*.
- Agarwal, A., K. Hosanagar, M. Smith. 2011. Location, Location, Location: An Analysis of Profitability of Position in Online Advertising Markets. *Journal of Marketing Research*. 48(6), 1057-1073.
- Ansari, A., and Mela, C. F. 2003. E-customization. *Journal of Marketing Research,* 40(2), 131–145.
- Aral, S. and D. Walker. 2011. Creating Social Contagion Through Viral Product Design: A Randomized Trial of Peer Influence in Networks. *Management Science*. 57(9), 1623-1639.
- Arora, N., and Henderson, T. 2007. Embedded premium promotion: Why it works and how to make it more effective. *Marketing Science*, 26(4), 514–531.
- Baye, M.R., Gatti, J.R.J., Kattuman, P. and Morgan, J. 2009. Clicks, Discontinuities, and Firm Demand Online. *Journal of Economics & Management Strategy*. 18(4), 935-975.
- Chen, Y. and S. Yao. 2012. Search with Refinement. *Working Paper*.
- Chevalier, J. and D. Mayzlin. 2006. The Effect of Word of Mouth on Sales: Online Book Reviews. *Journal of Marketing Research*. Vol. XLIII, 345–354.
- Chib, S., E. Greenberg. 1995. Understanding the Metropolis-Hastings algorithm. *The American Statistician*, 49(4), 327-335.
- Chittor, V. 2010. Online Retail: Getting The Right Product in Front of Your Customers. http://www.inc.com/internet/articles/201002/chittoor.html.

- Dellarocas, C. 2012.Double Marginalization in Performance-based Advertising: Implications and Solutions. *Management Science*, 58 (6), 1178-1195.
- De Los Santos, B. and S. Koulayev. 2012. Optimizing Click-through in Online Rankings for Partially Anonymous Consumers, Working Paper, Indiana University.
- Dodds, W. B., Monroe, K. B., and Grewal, D. 1991. Effects of Price, Brand, and Store Information on Buyers' Product Evaluations. *Journal of Marketing Research*, 28 (3), 307-319.
- Dyzabura, D. 2012. Product Search as Consumers Learn Their Preferences Through Product Evaluation. Working Paper, New York University.
- Ellison, G. and Ellison, S.F. 2009. Search, Obfuscation, and Price Elasticities on the Internet. *Econometrica*, 77, 427-452.
- Ghose A, and Yang, S. 2009. An Empirical Analysis of Search Engine Advertising: Sponsored Search in Electronic Markets. *Management Science*. 55(10), 1605-1622.
- Ghose, A., Ipeirotis, P. G. and Li, B. 2012. Designing Ranking Systems for Hotels on Travel Search Engines by Mining User-Generated and Crowd-Sourced Content, *Marketing Science*, 31(3), 493-520.
- Goldfarb, A., and C. Tucker. 2011. Online Display Advertising: Targeting and Obtrusiveness. *Marketing Science*, 30(3), 389-404.
- Jerath, K., L. Ma, Y. Park and K. Srinivasan. 2011. A "Position Paradox" in Sponsored Search Auctions. *Marketing Science*, 30 (4), 612-627.
- Koulayev, Sergei. 2010. Estimating demand in online search markets, with application to hotel bookings. *Working Paper*.
- Kuksov, D. and Villas-Boas, J.M. 2010. When More Alternatives Lead to Less Choice. *Marketing Science*. 29(3), 507–524.
- Lambrechet, A. and C. Tucker 2012. When Does Retargeting Work? Timing Information Specificity, *Working paper*, SSRN.
- Li, Xinxin and Hitt, Lorin M.. 2010. Price Effects in Online Product Reviews: An Analytical Model and Empirical Analysis. *MIS Quarterly*, 34(4), 809-831.
- Malthouse, E. C., and Elsner, R. 2006. Customization With Cross-basis Sub-segmentation. *Database Marketing & Customer Strategy Management*, 14(1), 40–50.
- Moe, W. and P. S. Fader. 2003. Dynamic Conversion Behavior at e-commerce Sites. *Management Science*. 50(3) 326–335.
- Narayanan, S. and K. Kalyanam. 2011. Measuring Position Effects in Search Advertising: A Regression Discontinuity Approach. *Working Paper*.
- Nevo, A. 2001. Measuring Market Power in the Ready-to-Eat Cereal Industry, *Econometrica*, 69(2), 307-342.
- Ranjith, G. 2005. Interferon-α-Induced Depression: When a Randomized Trial Is Not a Randomized Controlled Trial, *Psychother Psychosom*, 74(6):387.
- Rossi, P. E., McCulloch, R. E., and Allenby, G. M. 1996. The Value of Purchase History Data in Target Marketing. *Marketing Science*, 15(4), 321–340.
- Rossi, P. E. and Greg M. Allenby. 2003. Bayesian Statistics and Marketing. *Marketing Science*, 22(3), 304-328.
- Rutz, O. and R.E. Bucklin. 2007. A Model of Individual Keyword Performance in Paid Search Advertising. *Working paper*, Yale University, New Haven, CT.
- Rutz, O. J. and M. Trusov. 2011. Zooming In on Paid Search Ads—A Consumer-Level Model Calibrated on Aggregated Data. *Marketing Science*, 30(5), 789-800.
- Yang, S. and A. Ghose. 2010. Analyzing the Relationship Between Organic and Sponsored Search Advertising: Positive, Negative, or Zero Interdependence? *Marketing Science*, 29(4), 602-623.
- Yao, S. and C.F. Mela. 2011. A Dynamic Model of Sponsored Search Advertising. *Marketing Science*, 30(3), 447-468.
- Zhang, J. and Wedel, M. 2009. The Effectiveness of Customized Promotions in Online and Offline Stores. *Journal of Marketing Research*. 46(2), 190-206.

Table 1. Definitions and Summary Statistics of Variables

Variable	Definition	Mean	Std. Dev.	Min	Max
PRICE	Transaction price per room per night	120.45	73.25	25.77	978
DISPLAY	Number of displays	213.65	382.28	1	4849
CLICK	Number of clicks	2.99	3.55	0	56
CONVERSION	Number of conversions	1.26	0.66	0	9
PAGE	Page number of the hotel	20.86	13.44	1	192
RANK	Screen position of the hotel within a page	12.09	4.32	1	25
CLASS	Hotel class	3.36	1.37	1	5
REVIEWCNT	Total number of reviews	21.06	29.28	1	202
RATING	Overall reviewer rating	3.84	.85	1	5
SPECIALSORT	Vector of six control variables indicating the frequency of using different sorting methods				
DFT	Default sorting	188.50	369.58	0	4711
PRA	Price Ascending	13.99	23.34	0	338
CLD	Class Descending	1.49	3.42	0	37
CLA	Class Ascending	0.16	0.65	0	11
CNA	City Name Ascending	0.13	0.54	0	9
HNA	Hotel Name Ascending	0.35	0.95	0	15
H	Total number of hotels in a city	24.03	56.48	1	922
BRAND	Dummies for 9 hotel brands: Accor, Best western, Cendant, Choice, Hilton, Hyatt, Intercontinental, Marriott, and Starwood	--	--	0	1
Number of Observations (Weekly-Level): 29,222		**Time Period: 11/1/2008-1/31/2009**			

Table 2 Main Results from Model Estimation

Table 2a. Coefficient Estimates from Clickthrough Model

	Mean	*Class*	*Price*$^{(L)}$	*Rating*
Intercept	1.049(.054)*	.040(.011)*	--	--
Rank	-.062(.007)*	-.026(.004)*	-.019(.004)*	.020(.003)*
Rank^2	.004(.000)*	--	--	--
Page	-.035(.004)*	-.007(.001)*	-.011(.005)*	.016(.002)*
Price$^{(L)}$	-.141(.021)*	.002(.000)*	--	.004(.000)*
Rating	.078(.015)*	.001(.002)	--	--
ReviewCnt$^{(L)}$	.033(.009)*	.029(.032)	-.002(.023)	.017(.003)*
H$^{(L)}$*(Total #of Hotels)*	-.007(.000)*	--	--	--
Brand	Yes			
SpecialSort$^{(L)}$	Yes			

Unobserved Heterogeneity Estimates (Covariance Matrix Σ^{β} *)*

	Intercept	*Rank*	*Page*	*Price*	*Rating*	*ReviewCnt*$^{(L)}$
Intercept	1.012(.041)*	--	--	--	--	--
Rank	-.029 (.003)*	.118(.045)*	--	--	--	--
Page	.016(.001)*	-.025(.002)*	.102(.032)*	--	--	--
Price	-.156(.029)*	-.020(.008)*	.031(.101)	1.443(.058)*	--	--
Rating	.025(.006)*	-.051(.206)	-.042(.067)	-.039(.012)*	.067(.003)*	--
ReviewCnt$^{(L)}$	.003(.000)*	-.109(.099)	.037(.008)*	.060(.297)	-0.116(.004)*	.217(.040)*

$^{(L)}$: The natural logarithm form of the variable. *: Significance level at $p < 5\%$.

Note: *SpecialSort* is a vector of six control variables indicating the frequency of use of different sorting criteria.

Table 2b. Coefficient Estimates from Conversion Model

	Mean	*Class*	*Price*$^{(L)}$	*Rating*
Intercept	1.087(.166)*	.057(.011)*	--	--
Rank	-.021(.003)*	-.009(.002)*	-.010(.001)*	.015(.005)*
Rank^2	.002(.000)*	--	--	--
Page	-.029(.004)*	-.008(.001)*	-.006(.002)*	.003(.002)
Price$^{(L)}$	-.156(.047)*	.014(.011)*	--	.009(.001)*
Rating	.037(.001)*	.002(.003)	-.007(.016)	--
ReviewCnt$^{(L)}$	.019(.001)*	.013(.028)	-.005(.017)	.012(.001)*
H$^{(L)}$*(Total #of Hotels)*	-.008(.001)*	--	--	--
Brand	Yes			
SpecialSort$^{(L)}$	Yes			

Unobserved Heterogeneity Estimates (Covariance Matrix Σ^{γ} *)*

	Intercept	*Rank*	*Page*	*Price*	*Rating*	*ReviewCnt*$^{(L)}$
Intercept	1.225(.032)*	--	--	--	--	--
Rank	-.041 (.012)*	.089(.022)*	--	--	--	--
Page	.038(.007)*	-.070(.031)*	.216(.088)*	--	--	--
Price	-.203(.056)*	.104(.051)*	.044(.093)	2.005(.262)*	--	--
Rating	-.159(.234)	.137(.419)	.028(.036)	.077(.032)*	.108(.024)*	--
ReviewCnt$^{(L)}$	.015(.003)*	-.089(.106)	.020(.001)*	.111(.183)	0.165(.052)*	.304(.086)*

$^{(L)}$: The natural logarithm form of the variable. *: Significance level at $p < 5\%$.

Note: *SpecialSort* is a vector of six control variables indicating the frequency of use of different sorting criteria.

Table 2c. Coefficient Estimates from Ranking Model

	Mean	*Class*	*Price*$^{(L)}$	*Rating*
Intercept	1.487(.059)*	-.017(.002)*	--	--
CR_{t-1}	-.121(.014)*	-.005(.010)	-.004(.001)*	.017(.022)
Price$^{(L)}$	.114(.023)*	.002(.003)	--	-.012(.001)*
Rating	-.019(.000)*	.019(.027)	--	--
ReviewCnt$^{(L)}$	-.017(.000)*	-.003(.000)*	-.006(.002)*	-.002(.000)*
$H^{(L)}$*(Total #of Hotels)*	.010(.001)*	--	--	--
Brand	Yes			
SpecialSort$^{(L)}$	Yes			

Unobserved Heterogeneity Estimates (Covariance Matrix Σ^{ω} *)*

	Intercept	CR_{t-1}	*Price*	*Rating*	*ReviewCnt*$^{(L)}$
Intercept	2.246(.117)*	--	--	--	--
CR_{t-1}	-.107 (.033)*	.282(.057)*	--	--	--
Price	.114(.012)*	-.095(.040)*	.332(.056)*	--	--
Rating	-.201(.023)*	.037(.013)*	-.002(.027)	.838(.126)*	--
ReviewCnt$^{(L)}$	-.032(.002)*	-.043(.155)	.054(.118)	-.069(.033)*	.078(.023)*

$^{(L)}$: The natural logarithm form of the variable. *: Significance level at $p < 5\%$.

Note: *SpecialSort* is a vector of six control variables indicating the frequency of use of different sorting criteria.

Table 2d. Coefficient Estimates from Rating Model

	Mean	*Class*	*Price*$^{(L)}$
Intercept	2.198(.056)*	.035(.008)*	--
Rank	-.028(.007)*	.001(.005)	.003(.002)
Rank^2	.004(.001)*	--	--
Page	-.007(.000)*	-.002(.000)*	-.004(.000)*
Price$^{(L)}$	.005(.001)*	.001(.003)	--
ReviewCnt$^{(L)}$	.003(.000)*	.006(.011)	.017(.015)
$H^{(L)}$*(Total #of Hotels)*	.004(.000)*	--	--
Brand	Yes		
SpecialSort$^{(L)}$	Yes		

Unobserved Heterogeneity Estimates (Covariance Matrix Σ^{ρ} *)*

	Intercept	*Rank*	*Page*	*Price*	*ReviewCnt*$^{(L)}$
Intercept	4.123(.287)*	--	--	--	--
Rank	.195 (.046)*	.086(.030)*	--	--	--
Page	.086(.025)*	.127(.053)*	.326(.068)*	--	--
Price	-.211(.078)*	.061(.080)	-.155(.189)	2.017(.235)*	--
ReviewCnt$^{(L)}$	.001(.003)	-.098(.105)	.072(.034)*	-.209(.276)	.174(.060)*

$^{(L)}$: The natural logarithm form of the variable. *: Significance level at $p < 5\%$.

Note: *SpecialSort* is a vector of six control variables indicating the frequency of use of different sorting criteria.

Table 2e. Covariance Across Clickthrough, Conversion, Rank and Rating Ω_{jt}

	Clickthrough	*Conversion*	*Rank*	*Rating*
Clickthrough	2.721(.087)*	--	--	--
Conversion	2.006(.043)*	.773(.060)*	--	--
Rank	-.214(.022)*	-.626(.051)*	.521(.060)*	--
Rating	.835(.067)*	.304(.038)*	-.409(.079)*	.339(.036)*

*: Significance level at p < 5%.

Table 3: Model Fit Comparison Results

	Main Model (Quadratic Rank Term, Full Heterogeneity)	Model with Quadratic Rank Term, Partial Heterogeneity	Model with Linear Rank Term, Full Heterogeneity	Model with Linear Rank Term, Partial Heterogeneity	Model with Ordered Probit for Rank
	In-sample Model Prediction (Click-Through Rate)				
RMSE	0.0665	0.0759	0.0732	0.0968	0.1020
MSE	0.0044	0.0058	0.0054	0.0094	0.0104
MAD	0.0102	0.0165	0.0152	0.0282	0.0345
	Out-of-sample Model Prediction (Click-Through Rate)				
RMSE	0.0939	0.1068	0.1134	0.1247	0.1601
MSE	0.0088	0.0114	0.0129	0.0156	0.0256
MAD	0.0361	0.0427	0.0464	0.0505	0.0963
	In-sample Model Prediction (Conversion Rate)				
RMSE	0.0816	0.0996	0.0925	0.1127	0.1445
MSE	0.0067	0.0099	0.0086	0.0127	0.0209
MAD	0.0183	0.0237	0.0208	0.0389	0.0490
	Out-of-sample Model Prediction (Conversion Rate)				
RMSE	0.1164	0.1218	0.1292	0.1573	0.1867
MSE	0.0135	0.0149	0.0167	0.0247	0.0349
MAD	0.0386	0.0523	0.0479	0.0688	0.1102

Table 4: Policy Experiment Results for Search Engine Revenue Prediction

Ranking Mechanism	*Predicted Revenues from Top-1 Ranked Hotel ($)*	*Predicted Overall Revenues from All Hotels ($)*
Utility	1846	**423,401**
CR	1866	415,678
Travelocity Default	**2210**	402,349
Rating	1739	367,662
Price	**2003**	361,096
CTR	1476	312,757

Table 5a: Experimental Design – Study I

		(Within-Subject)	
		New York City	**Los Angeles**
(Between-Subject)	**Treatment Group 1**	BVR	BVR
	Treatment Group 2	Price	Price
	Treatment Group 3	TripAdvisor Rating	TripAdvisor Rating
	Treatment Group 4	Travelocity Rating	Travelocity Rating

Table 5b: Experimental Design – Study II

		(Within-Subject)	
		New York City	**Los Angeles**
(Between-Subject)	**Control Group**	Full Access	Full Access
	Treatment Group 1	No Search Context	No Search Context
	Treatment Group 2	No Weight	No Weight

Table 6: Experiment Results - Average User Behavior under Different Ranking Mechanisms

	Purchase Propensity (NYC)	*Purchase Propensity (LA)*
BVR (Utility)	**0.88**	**0.93**
Price	0.65	0.69
TripAdvisor Rating	0.54	0.44
Travelocity Rating	0.47	0.41

Group mean over all users.
Significant ($p<0.05$), Post Hoc ANOVA.

Table 7: Experiment Results - # of Clicks Received at Top-10 Ranking Positions

		Rank1	*Rank2*	*Rank3*	*Rank4*	*Rank5*	*Rank6*	*Rank7*	*Rank8*	*Rank9*	*Rank10*
BVR	NYC	56	24	13	10	9	11	8	2	1	1
	LA	68	20	14	11	10	7	5	4	1	2
Price	NYC	25	10	9	9	7	5	2	1	0	0
	LA	34	15	10	8	6	4	3	2	0	1
TripAdvisor	NYC	31	12	8	8	5	4	4	0	1	1
	LA	23	15	10	9	4	2	0	1	0	1
Travelocity	NYC	23	11	9	6	7	4	4	1	0	2
	LA	17	9	8	8	5	3	2	2	0	0

Table 8: Experiment Results - Search Engine Revenue under Different Ranking Mechanisms

	Revenues from Top-1 Ranked Hotel ($)	*Overall Revenues from All Hotels ($)*
BVR (Utility)	2052	**7162**
Price	**2876**	6898
TripAdvisor Rating	1738	4350
Travelocity Rating	1486	4002

Revenue summed over two cities (NYC and LA).
Significant ($p<0.05$), Post Hoc ANOVA.

Table 9a: Experiment Results – Interaction Effect between Ranking Positions and Hotel Class Ratings

Rank	*Luxury (4-, 5-star)*	*Budget (1-, 2-, 3-star)*
1→ 3	- 69%	- 43%
1→ 5	- 75%	- 54%
1→ 10	- 99%	- 80%

Results are based on average CTR.

Table 9b: Experiment Results – Interaction Effect between Ranking Positions and Hotel Customer Ratings

Rank	*Good (4-, 5-star)*	*Poor (1-, 2-star)*
10→ 5	11%	45%
10→ 3	49%	166%
10→ 1	83%	245%

Results are based on average CTR.

Table 10: Experiment Results - Average User Time and Activities under Different Personalized Ranking Mechanisms

	Time Spent (seconds)	*Total # of Activities*
Active Personalized Ranking with Full Access	351.23	19.36
Passive Personalized Ranking with No Search Context or Demographics [16]	228.52	16.78
Passive Personalized Ranking with No Weights of Individual Preferences [17]	127.01	8.24

Group mean over all users, across two cities (NYC and LA).
Significant ($p<0.05$), Post Hoc ANOVA.

[16] This passive personalized ranking only allows users to personalize their weights of individual preferences.
[17] This passive personalized ranking only allows users to personalize their search contexts and demographics.

Table 11: Experiment Results - User Behavior and Search Engine Revenues under Different Personalized Ranking Mechanisms

	# of Clicks (NYC)	*# of Clicks (LA)*	*Purchase Propensity (NYC)*	*Purchase Propensity (LA)*	*Overall Revenues ($)*
Active Personalized Ranking with Full Access	**2.17**	**2.36**	**0.51**	**0.55**	**5103**
Passive Personalized Ranking with No Search Context or Demographics	1.38	1.40	0.77	0.83	6631
Passive Personalized Ranking with No Weights of Individual Preferences	1.62	1.67	0.72	0.73	6254
Group mean over all users. Significant (p<0.05), Post Hoc ANOVA.					

Table 12: Experiment Results - Negative Binomial Model on # of Clicks

	Coeff.	*Coeff.*	*Coeff.*
NOSEARCH	-.891* (.362)	-.889* (.371)	-.773* (.242)
NOWEIGHT	-.577* (.230)	-.569* (.238)	-.494* (.201)
City	No	Yes	Yes
Activities	No	No	Yes
Log pseudolikelihood	-176.56322	-176.54825	-155.10346
*: Significance level at p < 5%.			

Table 13: Experiment Results - Probit Model on Purchase Propensity

	Coeff.	*Coeff.*	*Coeff.*
NOSEARCH	.587* (.233)	.581* (.228)	.591* (.219)
NOWEIGHT	.076 (.096)	.080 (.089)	.167* (.093)
City	No	Yes	Yes
Activities	No	No	Yes
Log pseudolikelihood	-341.00704	-340.88529	-318.09032
*: Significance level at p < 5%.			

Table 14: Experiment Results - Robustness Test (1)

	Time Spent (seconds)	*Total # of Activitie*	*# of Clicks (NYC)*	*# of Clicks (LA)*	*Purchase Propensity (NYC)*	*Purchase Propensity (LA)*	*Overall Revenues ($)*
High-Level Active Personalized Ranking with Full Access	343.02	19.27	2.28	2.42	0.45	0.44	4622
Passive Personalized Ranking with No Search Context or Demographics	228.52	16.78	1.38	1.40	0.77	0.83	6631
Passive Personalized Ranking with No Weights of Individual Preferences	127.01	8.24	1.62	1.67	0.72	0.73	6254

Table 15: Experiment Results - Robustness Test (2)

	Time Spent (seconds)	*Total # of Activities*	*Purchase Propensity (NYC)*	*Purchase Propensity (LA)*
Active Personalized Ranking with a Planned Purchase	351.23	19.36	0.51	0.55
Active Personalized Ranking with an Unplanned Purchase	177.01	18.18	0.75	0.69

Figure 1: Screenshot of the Search Result Page on Travelocity.com

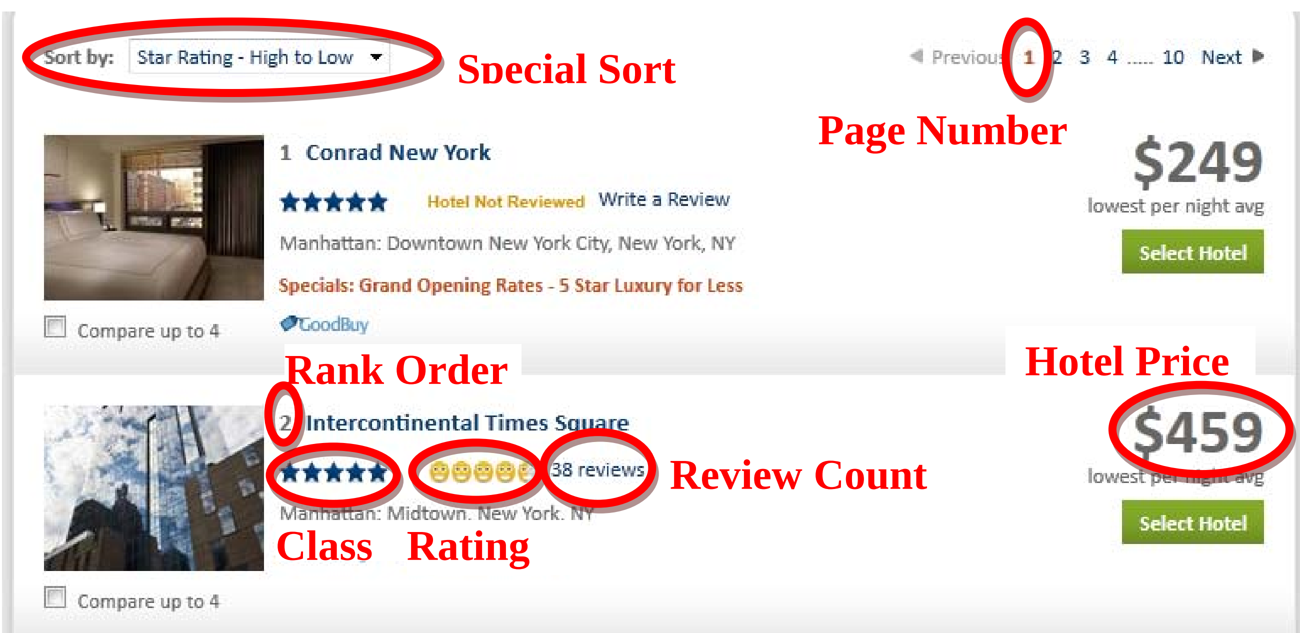


Figure 2: Screenshot of the Main Search Interface of the Hotel Search Engine

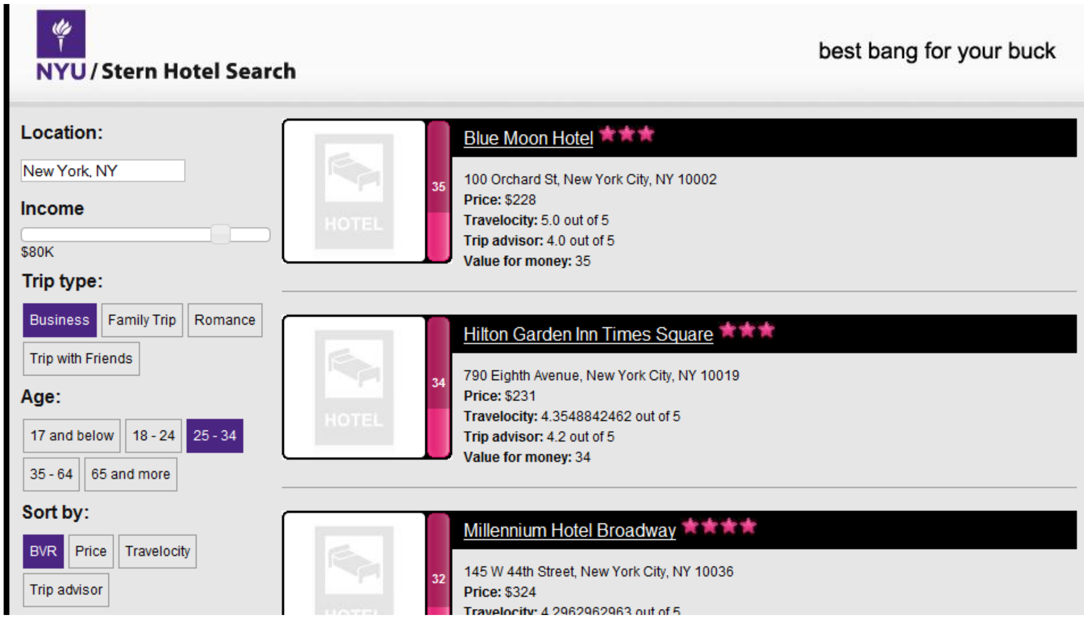
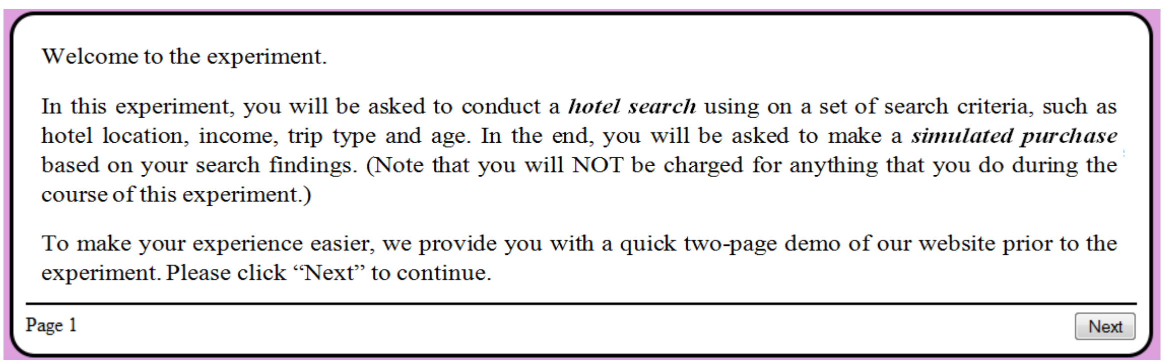


Figure 4: Screenshot of the Introduction Page (1)



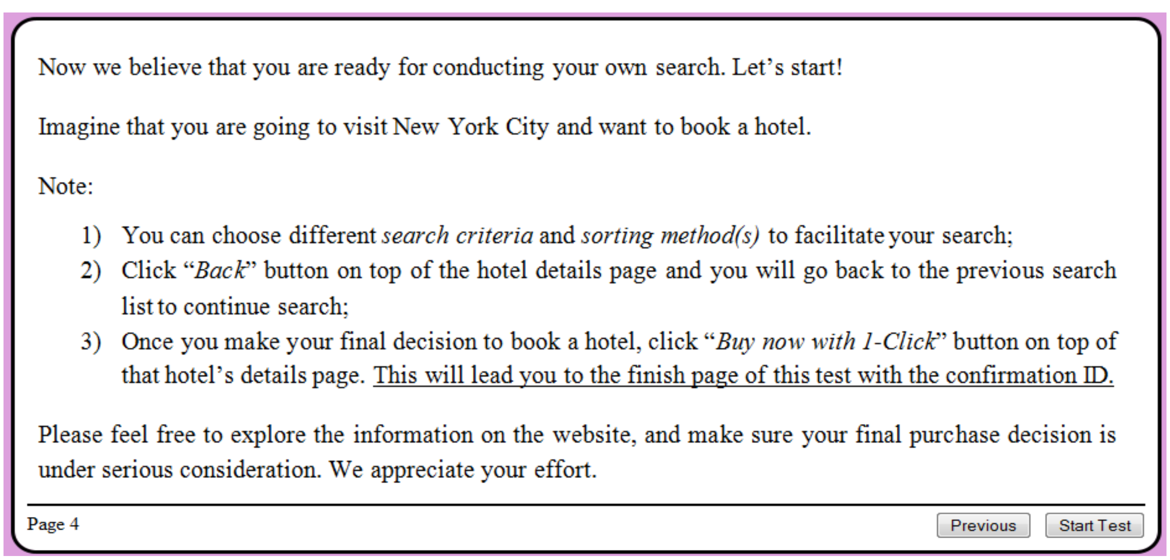
Welcome to the experiment.

In this experiment, you will be asked to conduct a ***hotel search*** using on a set of search criteria, such as hotel location, income, trip type and age. In the end, you will be asked to make a ***simulated purchase*** based on your search findings. (Note that you will NOT be charged for anything that you do during the course of this experiment.)

To make your experience easier, we provide you with a quick two-page demo of our website prior to the experiment. Please click "Next" to continue.

Page 1 Next

Figure 5: Screenshot of the Introduction Page (2)



Now we believe that you are ready for conducting your own search. Let's start!

Imagine that you are going to visit New York City and want to book a hotel.

Note:

1) You can choose different *search criteria* and *sorting method(s)* to facilitate your search;
2) Click "*Back*" button on top of the hotel details page and you will go back to the previous search list to continue search;
3) Once you make your final decision to book a hotel, click "*Buy now with 1-Click*" button on top of that hotel's details page. <u>This will lead you to the finish page of this test with the confirmation ID.</u>

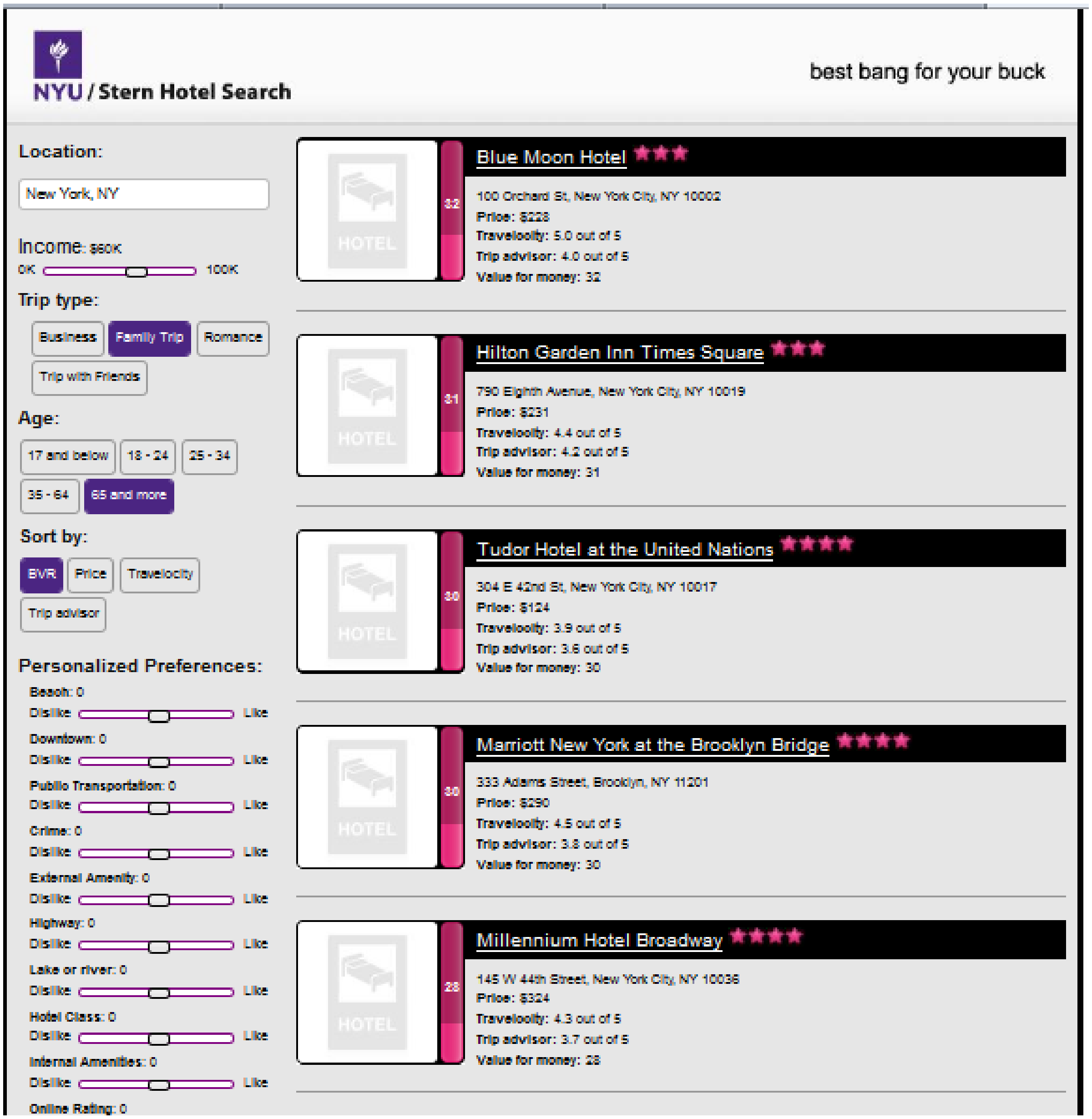
Please feel free to explore the information on the website, and make sure your final purchase decision is under serious consideration. We appreciate your effort.

Page 4 Previous Start Test

Figure 3: Screenshot of a Sample Hotel Landing Page[18]

[18] There are totally 25 hotel features on the landing page. For brevity, we only list 7 features here: price, beach, downtown, hotel class, internal amenities, online rating, and review count.

Figure 6: Screenshot of the Main Search Interface (Robustness Test)



Appendices for "Examining the Impact of Ranking on Consumer Behavior and Search Engine Revenue"

Appendix A: Results from Model with Partial Heterogeneity

Table A1: Coefficient Estimates from Clickthrough Rate Model

	Intercept	*Price*$^{(L)}$	*Class*	*Rating*
Intercept	1.023(.077)*	-.134(.014)*	.056(.009)*	.070(.011)*
Rank	-.053(.004)*	-.014(.003)*	-.022(.003)*	.017(.004)*
Rank^2	.004(.000)*	--	--	--
Page	-.031 (.002)*	--	--	--
ReviewCnt$^{(L)}$	.026(.000)*	--	--	--
$H^{(L)}$*(Total # Of Hotels)*	-.008(.000)*	--	--	--
Brand	Yes			
SpecialSort$^{(L)}$	Yes			

Unobserved Heterogeneity Estimates (Covariance Matrix Σ^{β} *)*

	Intercept	*Rank*
Intercept	1.221(.058)*	--
Rank	-.056 (.005)*	.078(.010)*

$^{(L)}$: The natural logarithm form of the variable. *: Significance level at $p < 5\%$.

Note: *SpecialSort* is a vector of six control variables indicating the frequency of use of different sorting criteria.

Table A2: Coefficient Estimates from Conversion Rate Model

	Intercept	*Price*$^{(L)}$	*Class*	*Rating*
Intercept	1.061(.142)*	-.133(.042)*	.052(.009)*	.036(.001)*
Rank	-.014(.000)*	-.008(.000) *	-.007(.000) *	.010 (.002) *
Rank^2	.001(.000)*	--	--	--
Page	-.025(.002)*	--	--	--
ReviewCnt$^{(L)}$	.016(.001)*	--	--	--
$H^{(L)}$*(Total # Of Hotels)*	-.007(.001)*	--	--	--
Brand	Yes			
SpecialSort$^{(L)}$	Yes			

Unobserved Heterogeneity Estimates (Covariance Matrix Σ^{γ} *)*

	Intercept	*Rank*
Intercept	1.032(.051) *	--
Rank	-.034(.002)*	.049(.004)*

$^{(L)}$: The natural logarithm form of the variable. *: Significance level at $p < 5\%$.

Note: *SpecialSort* is a vector of six control variables indicating the frequency of use of different sorting criteria.

Table A3: Coefficient Estimates from Ranking Model

	Intercept
Intercept	1.212(.046)*
CR_{t-1}	-.110(.009)*
Price$^{(L)}$	.099(.007)*
Class	-.015(.003)*
Rating	-.015(.000)*
ReviewCnt$^{(L)}$	-.014(.001)*
$H^{(L)}$*(Total # Of Hotels)*	.008(.002)*
Brand	Yes
SpecialSort$^{(L)}$	Yes
Unobserved Heterogeneity Estimates (Covariance Matrix Σ^{ω} *)*	
Intercept	1.157(.073)*

$^{(L)}$: The natural logarithm form of the variable.
*: Significance level at $p < 5\%$.
Note: *SpecialSort* is a vector of six control variables indicating the frequency of use of different sorting criteria.

Table A4: Coefficient Estimates from Rating Model

	Intercept	*Price*$^{(L)}$	*Class*
Intercept	3.172(.043)*	.004(.000)*	.030(.009)*
Rank	-.025(.005)*	.002(.002)	.001(.002)
Rank^2	.003(.001)*	--	--
Page	-.002(.000)*	--	--
ReviewCnt$^{(L)}$	.001(.000)*	--	--
$H^{(L)}$*(Total # Of Hotels)*	.003(.001)*	--	--
Brand	Yes		
SpecialSort$^{(L)}$	Yes		

Unobserved Heterogeneity Estimates (Covariance Matrix Σ^{ρ} *)*

	Intercept	*Rank*
Intercept	3.221(.124)*	--
Rank	-.039(.005)*	.182(.069)*

$^{(L)}$: The natural logarithm form of the variable. *: Significance level at $p < 5\%$.
Note: *SpecialSort* is a vector of six control variables indicating the frequency of use of different sorting criteria.

Table A5: Covariance Across Clickthrough, Conversion, Rank and Rating Ω_{jt}

	Click-Through	*Conversion*	*Rank*	*Rating*
Click-Through	1.634(.071)*	--	--	--
Conversion	1.103(.055)*	.802(.052)*	--	--
Rank	-.127(.016)*	-.589(.040)*	.712(.083)*	--
Rating	.729(.085)*	.281(.027)*	-.631(.102)*	.117(.009)*

*: Significance level at $p < 5\%$.

Appendix B: Results from Linear Rank Model

Table B1: Coefficient Estimates from Clickthrough Rate Model

	Mean	*Class*	*Price*$^{(L)}$	*Rating*
Intercept	1.220(.072)*	.044(.002)*	--	--
Rank	-.063(.005)*	-.019(.002)*	-.007(.001)*	.017(.002)*
Page	-.041(.004)*	-.008(.001)*	-.014(.003)*	.022(.008)*
Price$^{(L)}$	-.145(.019)*	.005(.002)*	--	.004(.001)
Rating	.083(.023)*	.003(.007)	--	--
ReviewCnt$^{(L)}$	.028(.010)*	.017(.035)	-.003(.030)	.012(.000)*
$H^{(L)}$*(Total #of Hotels)*	-.008(.000)*	--	--	--
Brand	Yes			
SpecialSort$^{(L)}$	Yes			

$^{(L)}$: The natural logarithm form of the variable. *: Significance level at $p < 5\%$.

Note: *SpecialSort* is a vector of six control variables indicating the frequency of use of different sorting criteria.

Table B2: Coefficient Estimates from Conversion Rate Model

	Mean	*Class*	*Price*$^{(L)}$	*Rating*
Intercept	1.025(.166)*	.042(.009)*	--	--
Rank	-.014(.002)*	-.007(.001)*	-.009(.000)*	.009(.002)*
Page	-.026(.000)*	-.011(.001)*	-.012(.004)*	.002(.015)
Price$^{(L)}$	-.178(.066)*	.026(.009)*	--	.017(.006)*
Rating	.035(.001)*	.001(.006)	-.013(.033)	--
ReviewCnt$^{(L)}$	.014(.000)*	.021(.030)	-.011(.029)	.014(.002)*
$H^{(L)}$*(Total #of Hotels)*	-.007(.000)*	--	--	--
Brand	Yes			
SpecialSort$^{(L)}$	Yes			

$^{(L)}$: The natural logarithm form of the variable. *: Significance level at $p < 5\%$.

Note: *SpecialSort* is a vector of six control variables indicating the frequency of use of different sorting criteria.

Table B3: Coefficient Estimates from Ranking Model

	Mean	*Class*	*Price*$^{(L)}$	*Rating*
Intercept	2.112(.056)*	-.022(.004)*	--	--
CR_{t-1}	-.134 (.013)*	-.008(.017)	-.006(.000)*	.025(.040)
Price$^{(L)}$	.099(.003)*	.004(.011)	--	-.007(.001)*
Rating	-.023(.000)*	.028(.026)	--	--
ReviewCnt$^{(L)}$	-.024(.003)*	-.005(.000)*	-.012(.001)*	-.004(.001)*
$H^{(L)}$*(Total #of Hotels)*	.009(.001)*	--	--	--
Brand	Yes			
SpecialSort$^{(L)}$	Yes			

$^{(L)}$: The natural logarithm form of the variable. *: Significance level at $p < 5\%$.

Note: *SpecialSort* is a vector of six control variables indicating the frequency of use of different sorting criteria.

Table B4: Coefficient Estimates from Rating Model

	Mean	*Class*	*Price*$^{(L)}$
Intercept	2.457(.040)*	.028(.007)*	--
Rank	-.020(.005)*	.001(.004)	.001(.002)
Page	-.009(.001)*	-.003(.000)*	-.004(.001)*
Price$^{(L)}$	.005(.000)*	-.004(.012)	--
ReviewCnt$^{(L)}$	.003(.001)*	.011(.023)	.027(.036)
$H^{(L)}$*(Total #of Hotels)*	.006(.001)*	--	--
Brand	Yes		
SpecialSort$^{(L)}$	Yes		

$^{(L)}$: The natural logarithm form of the variable. *: Significance level at p < 5%.

Note: *SpecialSort* is a vector of six control variables indicating the frequency of use of different sorting criteria.

Table B5: Covariance Across Clickthrough, Conversion, Rank, and Rating Ω_{jt}

	Click-Through	*Conversion*	*Rank*	*Rating*
Click-Through	3.012(.071)*	--	--	--
Conversion	1.006(.041)*	.563(.042)*	--	--
Rank	-.124(.006)*	-.223(.014)*	.786(.067)*	--
Rating	1.081(.117)*	-.149(.007)*	-.371(.122)*	.045(.005)*

*: Significance level at p < 5%.

Appendix C: Model Estimation Results at Daily Level

Table C1: Coefficient Estimates from Clickthrough Rate Model

	Mean	*Class*	*Price*$^{(L)}$	*Rating*
Intercept	2.124(.059)*	.076(.019)*	--	--
Rank	-.077(.012)*	-.041(.006)*	-.031(.007)*	.014(.001)*
Rank^2	.006(.001)*	--	--	--
Page	-.042 (.009)*	-.009(.001)*	-.018(.001)*	.022(.005)*
Price$^{(L)}$	-.201(.033)*	.003(.000)*	--	.005(.000)*
Rating	.087(.023)*	.004(.025)	--	--
ReviewCnt$^{(L)}$	.046(.020)*	.007(.021)	-.001(.019)	.024(.011)*
H$^{(L)}$*(Total Number Of Hotels)*	-.017(.000)*	--	--	--
Brand	Yes			
SpecialSort$^{(L)}$	Yes			

$^{(L)}$: The natural logarithm form of the variable. *: Significance level at p < 5%.

Note: *SpecialSort* is a vector of six control variables indicating the frequency of use of different sorting criteria.

Table C2: Coefficient Estimates from Conversion Rate Model

	Mean	*Class*	*Price*$^{(L)}$	*Rating*
Intercept	.095(.020)*	1.002(.019)*	--	--
Rank	-.044 (.008)*	-.026(.009)*	-.027(.003)*	.017 (.004)*
Rank^2	.003(.000)*	--	--	--
Page	-.038(.005)*	-.014(.003)*	-.011(.002)*	.004(.004)
Price$^{(L)}$	-.191(.032)*	.011(.002)*	--	.012(.003)*
Rating	.057(.014)*	.017(.039)	--	--
ReviewCnt$^{(L)}$	.023(.010)*	.004(.008)	-.002(.031)	.025(.008)*
H$^{(L)}$*(Total Number Of Hotels)*	-.009(.001)*	--	--	--
Brand	Yes			
SpecialSort$^{(L)}$	Yes			

$^{(L)}$: The natural logarithm form of the variable. *: Significance level at p < 5%.

Note: *SpecialSort* is a vector of six control variables indicating the frequency of use of different sorting criteria.

Table C3: Coefficient Estimates from Ranking Model

	Mean	*Class*	*Price*$^{(L)}$	*Rating*
Intercept	.796(.045)*	-.030(.011)*	--	--
CR_{t-1}	-.157(.016)*	-.009(.022)	-.029(.002)*	.029(.068)
Price$^{(L)}$	.187(.036)*	.005(.010)	--	-.019(.001)*
Rating	-.046(.003)*	.039(.103)	--	--
ReviewCnt$^{(L)}$	-.033(.005)*	-.005(.000)*	-.009(.008)	-.007(.001)*
$H^{(L)}$*(Total Number Of Hotels)*	.011(.001)*	--	--	--
Brand	Yes			
SpecialSort$^{(L)}$	Yes			

$^{(L)}$: The natural logarithm form of the variable. *: Significance level at $p < 5\%$.

Note: *SpecialSort* is a vector of six control variables indicating the frequency of use of different sorting criteria.

Table C4: Coefficient Estimates from Rating Model

	Mean	*Price*$^{(L)}$	*Class*
Intercept	1.997(.066)*	--	.043(.016)*
Rank	-.046(.019)*	.007(.021)	.002(.009)
Rank^2	.002(.000)*	--	--
Page	-.009(.000)*	-.005(.001)*	-.004(.001)*
Price$^{(L)}$	.008(.002)*	--	.007(.015)
ReviewCnt$^{(L)}$	.001(.000)*	.034(.029)	.010(.023)
$H^{(L)}$*(Total Number Of Hotels)*	.003(.000)*	--	--
Brand	Yes		
SpecialSort$^{(L)}$	Yes		

$^{(L)}$: The natural logarithm form of the variable. *: Significance level at $p < 5\%$.

Note: *SpecialSort* is a vector of six control variables indicating the frequency of use of different sorting criteria.

Table C5. Covariance Across Clickthrough, Conversion, Rank and Rating Ω_{jt}

	Click-Through	*Conversion*	*Rank*	*Rating*
Click-Through	.608(.101)*	--	--	--
Conversion	1.472(.070)*	.805(.041)*	--	--
Rank	-.288(.035)*	-1.123(.069)*	1.207(.136)*	--
Rating	1.006(.052)*	.774(.085)*	-.966(.063)*	1.290(.067)*

*: Significance level at $p < 5\%$.

Appendix D The MCMC Algorithm

To estimate our model, we applied the MCMC methods using a Metropolis-Hastings algorithm with a random walk chain (Chib and Greenberg 1995). In particular, we ran the MCMC chain for 80,000 iterations and used the last 40,000 iterations to compute the mean and standard deviation of the posterior distribution of the model parameters. In this appendix, we demonstrate our MCMC algorithm for the simultaneous model of click-through rate, conversion rate and rank.

(1) Draw the utilities of click-through and conversion: U_{jt}^{p} and U_{jt}^{q}.

The likelihood function of observing the joint occurrence of n_{jt} clicks and m_{jt} conversions is

$$l\left(U_{jt}^{p}, U_{jt}^{q} \mid n_{jt}, m_{jt}\right) \propto \left(p_{jt} q_{jt}\right)^{m_{jt}} \left[p_{jt}(1-q_{jt})\right]^{n_{jt}-m_{jt}} \left(1-p_{jt}\right)^{N_{jt}-n_{jt}},$$

$$\text{where } p_{jt} = \frac{\exp(U_{jt}^{p})}{1+\exp(U_{jt}^{p})} \quad \text{and} \quad q_{jt} = \frac{\exp(U_{jt}^{q})}{1+\exp(U_{jt}^{q})};$$

$$U_{jt}^{p} = \overline{m}_{jt}^{p} + \varepsilon_{jt} \quad \text{and}$$

$$\overline{m}_{jt}^{p} = \beta_{j0} + \beta_{j1}\text{Rank}_{jt} + \beta_{j2}\text{Rank}_{jt}^{2} + \beta_{j3}\text{Page}_{jt} + \beta_{j4}\text{Price}_{jt} + \beta_{j5}\text{Rating}_{jt} + \beta_{j6}\text{ReviewCount}_{jt} + \alpha_{1}\text{Class}_{j} + \alpha_{2}\text{H}_{j} + \alpha_{3}\text{Brand}_{j} + \alpha_{4}\text{SpecialSort}_{jt};$$

$$U_{jt}^{q} = \overline{m}_{jt}^{q} + \eta_{jt} \quad \text{and}$$

$$\overline{m}_{jt}^{q} = \gamma_{j0} + \gamma_{j1}\text{Rank}_{jt} + \gamma_{j2}\text{Rank}_{jt}^{2} + \gamma_{j3}\text{Page}_{jt} + \gamma_{j4}\text{Price}_{jt} + \gamma_{j5}\text{Rating}_{jt} + \gamma_{j6}\text{ReviewCount}_{jt} + \theta_{1}\text{Class}_{j} + \theta_{2}\text{H}_{j} + \theta_{3}\text{Brand}_{j} + \theta_{4}\text{SpecialSort}_{jt}.$$

Let us denote D and E_{jt} as the conditional covariance matrix and mean vector of $(\varepsilon_{jt}, \eta_{jt})'$, respectively, conditioning on values of υ_{jt} and Ω. We use Metropolis-Hastings algorithm with a random walk chain to generate draws of $U_{jt} = (U_{jt}^{p}, U_{jt}^{q})$ (Chib and Greenberg 1995).

$$U_{jt}^{new} = U_{jt}^{old} + \Delta, \quad where\ \Delta \sim N(0, 0.01I).$$

The draws are accepted with a probability α where

$$\alpha = \min\left\{ \frac{\exp\left[-\frac{1}{2}(U_{jt}^{new} - \overline{m}_{jt} - E_{jt})' D^{-1} (U_{jt}^{new} - \overline{m}_{jt} - E_{jt})\right] l(U_{jt}^{new})}{\exp\left[-\frac{1}{2}(U_{jt}^{old} - \overline{m}_{jt} - E_{jt})' D^{-1} (U_{jt}^{old} - \overline{m}_{jt} - E_{jt})\right] l(U_{jt}^{old})}, 1 \right\}.$$

(2) Draw the random coefficients: $b_{j} = [\beta_{j0} ... \beta_{j6}, \gamma_{j0} ... \gamma_{j6}, \omega_{j0} ... \omega_{j4}, \rho_{j0} ... \rho_{j5}]$.

$$y_{jt1} = U_{jt}^{p} - (\alpha_{1}\text{Class}_{j} + \alpha_{2}\text{H}_{j} + \alpha_{3}\text{Brand}_{j} + \alpha_{4}\text{SpecialSort}_{jt});$$
$$y_{jt2} = U_{jt}^{q} - (\theta_{1}\text{Class}_{j} + \theta_{2}\text{H}_{j} + \theta_{3}\text{Brand}_{j} + \theta_{4}\text{SpecialSort}_{jt});$$
$$y_{jt3} = \ln(Rank_{jt}) - (\kappa_{1}\text{Class}_{j} + \kappa_{2}\text{H}_{j} + \kappa_{3}\text{Brand}_{j} + \kappa_{4}\text{SpecialSort}_{jt});$$

$y_{jt4} = Rating_{jt} - (\chi_1 \text{Class}_j + \chi_2 \text{H}_j + \chi_3 \text{Brand}_j + \chi_4 \text{SpecialSort}_{j,t-1});$

$$x_{jt} = \begin{bmatrix} x'_{jt1} & 0 & 0 & 0 \\ 0 & x'_{jt2} & 0 & 0 \\ 0 & 0 & x'_{jt3} & 0 \\ 0 & 0 & 0 & x'_{jt4} \end{bmatrix} \text{ and } \Sigma = \begin{bmatrix} \Sigma^{\beta} & & & \\ & \Sigma^{\gamma} & & \\ & & \Sigma^{\omega} & \\ & & & \Sigma^{\rho} \end{bmatrix}.$$

$x'_{jt1} = x'_{jt2} = [1 \; Rank_{jt} \; Rank^2_{jt} \; Page_{jt} \; \text{Price}_{jt} \; Rating_{jt} \; \text{ReviewCount}_{jt}]'$,

$x'_{jt3} = [1 \; Page_{jt} \; \text{Price}_{jt} \; Rating_{jt} \; \text{ReviewCount}_{jt}]'$,

$x'_{jt4} = [1 \; Rank_{jt} \; Rank^2_{jt} \; Page_{jt} \; \text{Price}_{jt} \; \text{ReviewCount}_{jt}]'$.

Then, we have the following

$$\bar{b}_j^{\beta_k} = \bar{\beta}_k + \delta_1^{\beta_k} \text{Class}_j + \delta_2^{\beta_k} \text{Price}_j + \delta_3^{\beta_k} \text{Rating}_j, \quad (k = 0...6);$$

$$\bar{b}_j^{\gamma_k} = \bar{\gamma}_k + \delta_1^{\gamma_k} \text{Class}_j + \delta_2^{\gamma_k} \text{Price}_j + \delta_3^{\gamma_k} \text{Rating}_j, \quad (k = 0...6);$$

$$\bar{b}_j^{\omega_k} = \bar{\omega}_k + \delta_1^{\omega_k} \text{Class}_j + \delta_2^{\omega_k} \text{Price}_j + \delta_3^{\omega_k} \text{Rating}_j, \quad (k = 0...4);$$

$$\bar{b}_j^{\rho_k} = \bar{\rho}_k + \delta_1^{\rho_k} \text{Class}_j + \delta_2^{\rho_k} \text{Price}_j + \delta_3^{\gamma\rho} \text{Rating}_j, \quad (k = 0...5);$$

and $b_j \sim MVN(A_j, B_j)$, where $B_j = \left[x'_j \Omega^{-1} x_j + \Sigma^{-1} \right]^{-1}$ and $A_j = B_j \left[x'_j \Omega^{-1} y_j + \Sigma^{-1} \bar{b}_j \right]$.

(3) Draw the homogeneous coefficients: $a_j = [\alpha', \theta', \kappa', \chi']$.

$y_{jt1} = U^p_{jt} - (\beta_{j0} + \beta_{j1}\text{Rank}_{jt} + \beta_{j2}\text{Rank}^2_{jt} + \beta_{j3}\text{Page}_{jt} + \beta_{j4}\text{Price}_{jt} + \beta_{j5}\text{Rating}_{jt} + \beta_{j6}\text{ReviewCount}_{jt});$

$y_{jt2} = U^q_{jt} - (\gamma_{j0} + \gamma_{j1}\text{Rank}_{jt} + \gamma_{j2}\text{Rank}^2_{jt} + \gamma_{j3}\text{Page}_{jt} + \gamma_{j4}\text{Price}_{jt} + \gamma_{j5}\text{Rating}_{jt} + \gamma_{j6}\text{ReviewCount}_{jt});$

$y_{jt3} = \ln(Rank_{jt}) - (\omega_{j0} + \omega_{j1}\text{CR}_{j,t-1} + \omega_{j2}\text{Price}_{jt} + \omega_{j3}\text{Rating}_{jt} + \omega_{j4}\text{ReviewCount}_{jt});$

$y_{jt4} = Rating_{jt} - (\rho_{j0} + \rho_{j1}\text{Rank}_{j,t-1} + \rho_{j2}\text{Rank}^2_{j,t-1} + \rho_{j3}\text{Page}_{j,t-1} + \rho_{j4}\text{Price}_{jt} + \rho_{j5}\text{ReviewCount}_{jt});$

$$x_{jt} = \begin{bmatrix} x'_{jt1} & 0 & 0 & 0 \\ 0 & x'_{jt2} & 0 & 0 \\ 0 & 0 & x'_{jt3} & 0 \\ 0 & 0 & 0 & x'_{jt4} \end{bmatrix};$$

$x'_{jt1} = x'_{jt2} = x'_{jt3} = [\text{Class}_j, \text{H}_j, \text{Brand}_j, \text{SpecialSort}_{jt}]';$

$x'_{jt4} = [\text{Class}_j, \text{H}_j, \text{Brand}_j, \text{SpecialSort}_{jt-1}]';$

$$\bar{a}_0 = 0 \text{ and } \Sigma_0 = 100I.$$

Then $a \sim MVN(A, B)$, where $B = \left[X'\Omega^{-1}X + \Sigma_0^{-1} \right]^{-1}$ and $A = B\left[X'\Omega^{-1}Y + \Sigma_0^{-1}\bar{a}_0 \right]$.

(4) Draw Ω, where Ω represents the covariance matrix of the four error terms in the models:

$$[\varepsilon_{jt}, \eta_{jt}, \upsilon_{jt}, \psi_{jt}]' \sim MVN[0, \; \Omega].$$

$$
\begin{aligned}
y_{jt1} &= U_{jt}^{p} - (\beta_{j0} + \beta_{j1}\text{Rank}_{jt} + \beta_{j2}\text{Rank}_{jt}^{2} + \beta_{j3}\text{Page}_{jt} + \beta_{j4}\text{Price}_{jt} + \beta_{j5}\text{Rating}_{jt} \\
&\quad + \beta_{j6}\text{ReviewCount}_{jt} + \alpha_1\text{Class}_j + \alpha_2\text{H}_j + \alpha_3\text{Brand}_j + \alpha_4\text{SpecialSort}_{jt}); \\
y_{jt2} &= U_{jt}^{q} - (\gamma_{j0} + \gamma_{j1}\text{Rank}_{jt} + \gamma_{j2}\text{Rank}_{jt}^{2} + \gamma_{j3}\text{Page}_{jt} + \gamma_{j4}\text{Price}_{jt} + \gamma_{j5}\text{Rating}_{jt} \\
&\quad + \gamma_{j6}\text{ReviewCount}_{jt} + \theta_1\text{Class}_j + \theta_2\text{H}_j + \theta_3\text{Brand}_j + \theta_4\text{SpecialSort}_{jt}); \\
y_{jt3} &= \ln(Rank_{jt}) - (\omega_{j0} + \omega_{j1}\text{CR}_{j,t-1} + \omega_{j2}\text{Price}_{jt} + \omega_{j3}\text{Rating}_{jt} + \omega_{j4}\text{ReviewCount}_{jt} \\
&\quad + \kappa_1\text{Class}_j + \kappa_2\text{H}_j + \kappa_3\text{Brand}_j + \kappa_4\text{SpecialSort}_{jt}); \\
y_{jt4} &= Rating_{jt} - (\rho_{j0} + \rho_{j1}\text{Rank}_{j,t-1} + \rho_{j2}\text{Rank}_{j,t-1}^{2} + \rho_{j3}\text{Page}_{j,t-1} + \rho_{j4}\text{Price}_{jt} \\
&\quad + \rho_{j5}\text{ReviewCount}_{jt} + \chi_1\text{Class}_j + \chi_2\text{H}_j + \chi_3\text{Brand}_j + \chi_4\text{SpecialSort}_{j,t-1});
\end{aligned}
$$

$$\Omega \sim IW\left(\sum_j \sum_t y_{jt}^{'} y_{jt} + Q_0,\ N + q_0\right),$$

where $Q_0 = 10I,\ q_0 = 10,\ \ N = \text{No. of hotels},$ and IW stands for the inverted Wishart distribution.

(5) Draw $\Sigma^{\beta}, \Sigma^{\gamma}, \Sigma^{\omega}, \Sigma^{\rho}$.

$$\Sigma^{\beta} \sim IW\left(\sum_j (\beta_j - \overline{\beta})'(\beta_j - \overline{\beta}) + Q_0,\ N + q_0\right),$$

$$\Sigma^{\gamma} \sim IW\left(\sum_j (\gamma_j - \overline{\gamma})'(\gamma_j - \overline{\gamma}) + Q_0,\ N + q_0\right),$$

$$\Sigma^{\omega} \sim IW\left(\sum_j (\omega_j - \overline{\omega})'(\omega_j - \overline{\omega}) + Q_0,\ N + q_0\right),$$

$$\Sigma^{\rho} \sim IW\left(\sum_j (\rho_j - \overline{\rho})'(\rho_j - \overline{\rho}) + Q_0,\ N + q_0\right),$$

where $Q_0 = 10I,\ q_0 = 10,\ N = \text{No. of hotels},$ and IW stands for the inverted Wishart distribution.

(6) Draw $f_1 = [\overline{\beta_0}, \delta_1^{\beta_0}, \delta_2^{\beta_0}, \delta_3^{\beta_0}, ..., \overline{\beta_6}, \delta_1^{\beta_6}, \delta_2^{\beta_6}, \delta_3^{\beta_6}]'$.

$$x_j = \begin{bmatrix} x_j^{0}{}' & & & \\ & x_j^{2}{}' & & \\ & & \cdots & \\ & & & x_j^{k}{}' \end{bmatrix}, \text{ where } x_j^{k}{}' = \begin{bmatrix} 1 & \text{Class}_j & \text{Price}_j & \text{Rating}_j \end{bmatrix}', \quad k = 0...6.$$

$f_1 \sim MVN(A, B),$ where $B = [X'\Sigma^{\beta^{-1}} X + \Sigma_0^{-1}]^{-1},\ \ A = B[\Sigma_0^{-1}\overline{a}_0 + X'\Sigma^{\beta^{-1}}\beta],\ \ \overline{a}_0 = 0$ and $\Sigma_0 = 100I.$

(7) Draw $f_2 = [\overline{\gamma_0}, \delta_1^{\gamma_0}, \delta_2^{\gamma_0}, \delta_3^{\gamma_0}, ..., \overline{\gamma_6}, \delta_1^{\gamma_6}, \delta_2^{\gamma_6}, \delta_3^{\gamma_6}]'$, similar to step 6.

(8) Draw $f_3 = [\overline{\omega_0}, \delta_1^{\omega_0}, \delta_2^{\omega_0}, \delta_3^{\omega_0}, ..., \overline{\omega_4}, \delta_1^{\omega_4}, \delta_2^{\omega_4}, \delta_3^{\omega_4}]'$, similar to step 6.

(9) Draw $f_4 = [\overline{\rho_0}, \delta_1^{\rho_0}, \delta_2^{\rho_0}, \delta_3^{\rho_0}, ..., \overline{\rho_5}, \delta_1^{\rho_5}, \delta_2^{\rho_5}, \delta_3^{\rho_5}]'$, similar to step 6.

Appendix E

Comparison of Mechanical Turk Users with overall US Internet Population

	June 2008	October 2008	December 2008
	US Internet Users comscore Data	**Mechanical Turk Users**	**Mechanical Turk Users**
Total Audience	100	100	100
Persons - Age			
Persons: 15+	85.9	100	100
Persons: 18+	80.1	99.6	99.5
Persons: 21+	74.3	92.9	91.1
Persons: 35+	52.4	39.3	37.1
Persons: 50+	24.3	11.2	10.7
Persons: 55+	16.2	5.2	5.4
Persons: 2-11	9.5	0	0
Persons: 2-17	19.9	0.2	0.4
Persons: 6-11	7.4	0	0
Persons: 6-14	12	0	0
Persons: 9-14	8.9	0	0
Persons: 12-17	10.4	0.2	0.4
Persons: 12-24	22.9	19	21.5
Persons: 12-34	38	57.8	60
Persons: 12-49	66.2	87.4	88.2
Persons: 18-24	12.5	18.7	21.1
Persons: 18-34	27.6	57.5	59.7
Persons: 18-49	55.8	87.2	87.8
Persons: 21-34	21.9	53.3	53.9
Persons: 21-49	50	82.9	82
Persons: 25-34	15.1	38.8	38.6
Persons: 25-49	43.2	68.4	66.7
Persons: 25-54	51.3	75.2	72.3
Persons: 35-44	18.7	22.4	21.5
Persons: 35-49	28.2	29.7	28.1
Persons: 35-54	36.2	36.4	33.7
Persons: 35-64	46.8	41.4	38.8
Persons: 45-54	17.6	14	12.2
Persons: 45-64	28.1	19	17.4
Persons: 55-64	10.5	5	5.2
Persons: 65+	5.7	0.7	1.1
Males - Age			
All Males	49.5	28	36.6
Male: 15+	42.1	28	36.6
Male: 18+	39.1	27.8	36.3
Male: 21+	36.1	24.7	32.4
Male: 35+	25.7	9.5	11.3
Male: 50+	12	2.8	2.6
Male: 55+	8.1	1.4	1.1
Male: 2-11	4.9	0	0

Male: 2-17	10.4	0.1	0.2
Male: 6-11	3.9	0	0
Male: 6-14	6.3	0	0
Male: 9-14	4.5	0	0
Male: 12-17	5.5	0.1	0.2
Male: 12-24	11.6	7.5	9.1
Male: 12-34	18.9	17.3	24.2
Male: 12-49	32.5	25	33.9
Male: 18-24	6.1	7.4	8.9
Male: 18-34	13.4	17.2	23.9
Male: 18-49	27.1	24.9	33.7
Male: 21-34	10.4	15.2	21.1
Male: 21-49	24.1	22.9	30.8
Males: 25-34	7.3	9.8	15
Male: 25-49	20.9	17.6	24.8
Male: 25-54	24.8	19	26.3
Males: 35-44	9.1	6	8
Male: 35-49	13.7	7.7	9.7
Male: 35-54	17.5	9.1	11.2
Male: 35-64	22.6	10.6	12.3
Male: 45-54	8.4	3.1	3.2
Male: 45-64	13.5	4.5	4.3
Males: 55-64	5.1	1.4	1.1
Males: 65+	3	0	0.1
Females - Age			
All Females	50.5	72	63.4
Female: 15+	43.8	72	63.4
Female: 18+	41	71.9	63.3
Female: 21+	38.2	68.2	58.7
Female: 35+	26.8	29.8	25.8
Female: 50+	12.3	8.3	8.1
Female: 55+	8.1	3.8	4.3
Female: 2-11	4.6	0	0
Female: 2-17	9.5	0.1	0.1
Female: 6-11	3.6	0	0
Female: 6-14	5.7	0	0
Female: 9-14	4.5	0	0
Female: 12-17	4.9	0.1	0.1
Female: 12-24	11.3	11.5	12.3
Female: 12-34	19.1	40.5	35.9
Female: 12-49	33.6	62.4	54.3
Female: 18-24	6.4	11.5	12.2
Female: 18-34	14.2	40.5	35.8
Female: 18-49	28.7	62.4	54.1
Female: 21-34	11.5	38.1	32.8
Female: 21-49	25.9	60	51.2
Females: 25-34	7.8	28.9	23.6
Female: 25-49	22.3	50.9	41.9
Female: 25-54	26.5	56.2	46
Females: 35-44	9.5	16.4	13.4

Female: 35-49	14.5	21.9	18.4
Female: 35-54	18.7	27.3	22.4
Female: 35-64	24.1	30.8	26.5
Female: 45-54	9.2	10.9	9
Female: 45-64	14.6	14.5	13.1
Females: 55-64	5.4	3.6	4.1
Females: 65+	2.6	0.7	1
HH Income (US)			
HHI USD: Less than 15,000	6	11.4	12.9
HHI US: Under $25K	9.3	22.8	23.1
HHI US: Under $60K	44.5	64.8	60.5
HHI US: $60K+	55.5	34.8	39.1
HHI US: $75K+	43	22.7	27.5
HHI USD: 15,000 - 24,999	3.4	11.4	10.1
HHI USD: 25,000 - 39,999	9.9	21.8	18.9
HHI USD: 40,000 - 59,999	25.3	20.2	18.6
HHI USD: 60,000 - 74,999	12.6	12.1	11.6
HHI USD: 75,000 - 99,999	17.7	10.2	11.5
HHI USD: 100,000 or more	25.3	12.5	16
Region (US)			
Region US:West North Central	7.6	5.8	7.5
Region US:Mountain	6.9	6.4	7.4
Region US:Pacific	15.4	13.3	15.7
Region US:New England	5.5	6.4	4.7
Region US:Mid Atlantic	14.2	13.9	15.8
Region US:South Atlantic	18.7	19.2	19.9
Region US:East South Central	5.1	8.3	5.2
Region US:West South Central	10.5	10.7	9
Region US:East North Central	16.1	15.7	14.8
Children			
Children:No	39.3	52.7	57.6
Children:Yes	60.7	47.3	42.3
HH Size			
HH Size: 1	4.4	17.7	17.3
HH Size: 2	24.2	28.9	30.6
HH Size: 3	21.4	19.7	19.2
HH Size: 4	25.3	20.5	21.9
HH Size: 5+	24.8	12.9	10.7
HH Size: 1-2	28.5	46.6	47.8
HH Size: 3+	71.5	33.5	32.7
Race			
Race:White	87.3	82.7	82
Race:Black	8	6.5	5.3
Race:Asian	1.6	5.7	6.8
Race:Other	3.1	4.9	5.8