Title
Evaluation Set Size and Purchase: Evidence from a Product Search Engine☆

Journal
Journal of Interactive Marketing, 37(1)

ISSN
1094-9968

Authors
Choudhary, Vidyanand
Currim, Imran
Dewan, Sanjeev
et al.

Publication Date
2017-02-01

DOI
10.1016/j.intmar.2016.07.003

Peer reviewed

Evaluation Set Size and Purchase: Evidence from a Product Search Engine[1]

Vidyanand Choudhary[2] Imran Currim[2] Sanjeev Dewan[2]
Ivan Jeliazkov[3] Ofer Mintz[4] John Turner[2]

Abstract

The last decade has seen a dramatic increase in the popularity of product search engines, yet the analysis of consumer behavior at such sites remains a challenging research problem despite its timeliness and importance. In this article, we develop and estimate a copula model of evaluation set size and purchase behavior employing data from 3,182 hotel searches by customers at a large travel search engine. The model allows us to jointly study purchase behavior, evaluation sets, and their antecedents. Our results reveal that evaluation set size and purchase are negatively correlated and that factors typically presumed to be associated with purchase – i.e., when users sort search results by price or quality, request many rooms, disclose that there are many guests in their party, or arrive from other search engines and/or partner sites – actually relate to larger evaluation sets but lower purchase probability. In contrast, when users filter the search results, we observe smaller evaluation sets and higher purchase probability. The theoretical background and practical implications of our findings suggest that efforts to increase purchases need not necessarily be predicated on cultivating larger evaluation sets.

Keywords: Evaluation Set Size; Purchase; Product Search Engines; Copula Model; E-Commerce; Bayesian estimation.

Introduction

Product search engines such as Amazon, Angie's List, CNET, EBay, or Travelocity, offer access to information on a large number of products and services and cater to customers with various characteristics and experience. To facilitate product evaluation and purchase, search results can typically be manipulated by sorting and filtering on characteristics such as brand, price, or quality. The number of options actually evaluated by a customer—the Evaluation Set Size (ESS)—and whether a purchase is made are key components in the interaction between consumers and the product search engine. However, in spite of the widespread use of product

[1] The authors thank S. Siddarth (USC) for providing valuable feedback and the Wharton Customer Analytics Initiative (WCAI) for providing the data.
[2] Paul Merage School of Business, University of California, Irvine CA 92697-3125.
[3] Department of Economics, University of California, Irvine CA 92697-5100.
[4] E. J. Ourso College of Business, Louisiana State University, Baton Rouge LA 70803.

search engines, there is little empirical evidence on customer behavior at such sites, including how past experience, filtering and sorting tools, referring site, volume of search results, or user characteristics such as usage magnitude (# rooms, # guests, # nights) or time to consumption, influence ESS and purchase probability.

Consequently, our investigation centers on the following main research questions: What are the antecedents of ESS? What are the drivers of purchase decisions in product search engines? And, what is the relationship between ESS and purchase? The evaluation set is analogous to the consideration set in marketing, which in prior studies employing supermarket scanner data has remained unobserved and has subsequently been imputed and considered exogenous to purchase (e.g., Andrews and Srinivasan 1995; Mehta, Rajiv, and Srinivasan 2003; Siddarth, Bucklin, and Morrison 1995). In contrast, as a result of a customer's interaction with the search engine, in our context the evaluation set is observed, enabling the investigation of its antecedents, and the relationship between ESS and purchase. Survey-based studies in which the consideration set is stated (Hauser and Wernerfelt 1990; Roberts and Lattin 1991), have considered the composition of this set but not its size. However, there are important reasons to focus on ESS. On the one hand, larger evaluation sets can enhance customers' involvement, knowledge, and confidence with a particular purchase. On the other hand, larger sets may deter purchase due to information processing costs. Hence, understanding the net effect is an empirical question of high relevance. These issues are particularly important for product search engine managers, who care about how purchases in general can be facilitated, in contrast to the more narrow goals of product managers, whose focus is on how purchases of a specific product can be facilitated. To address these questions, we study data on Budapest hotel searches by 3,182 customers on a large travel website. Here ESS corresponds to the number of hotel links a consumer clicked on, where after each click detailed information about that option was revealed, e.g., price, amenities, and geographical location.

The empirical analysis of such data can proceed by either structural or reduced-form modeling, each having its advantages and drawbacks. Structural models take an explicit stand on the process by which interactions are structured and data are generated, e.g., how (sets of) alternatives are evaluated, how purchase decisions are reached, whether choices are sequential or simultaneous, and so on. For example, Moe (2006) has successfully employed a structural two-stage model to show that product attributes of nutritional products, such as price and size, employed in the evaluation stage, are different from the ingredient attributes employed in the purchase stage. However, a structural model can be particularly challenging to formulate in the absence of sufficient *a priori* theory, in the presence of conflicting theories, or if the imposition of a realistic structure (e.g., flexible correlations) could make identification tenuous or estimation prohibitively difficult, especially when a plethora of available alternatives exist and the majority of customers do not purchase, as is common at product search engines. In contrast, reduced form models are more robust to theoretical conflicts, can be specified more parsimoniously, and can still be informative about the role of key covariates or the interactions among multiple outcomes. Consequently, we design a reduced form model to help investigate the antecedents of ESS, link ESS and purchase probability, and deal with a setting where there are few purchases relative to the number of shoppers.

Our econometric strategy is to start with two simple marginal models—one for ESS and the other for purchase—and then employ a copula to couple them into a joint model for the two outcomes. The first component entering the copula is a negative binomial model for ESS, which is a count variable. The specification involves the antecedents of ESS, i.e., past experience with the site, the referring site (where customers arrive from), the use of search tools (e.g., sorting or filtering), the volume of the corresponding search results, and user characteristics such as usage magnitude (# rooms, # guests) and time to consumption (the number of days the user's search precedes their planned travel). The second component entering the copula is a probit model for purchase, with the complication, however, that the binary purchase decision is unavailable for individuals whose evaluation sets are empty. Unobserved factors and individual heterogeneity can simultaneously affect both of the equations in the model, leading to correlation between the count and binary outcomes. Therefore, we focus on joint modeling and simultaneous estimation that allows for interactions and flexibility, and formally accounts for non-randomly missing outcomes. The econometric methodology also allows us to address model uncertainty and compare relevant competing specifications.

Our study reveals three main results. First, ESS depends positively on past experience with product category searches on the site, arrival from another search engine or partner site, use of sorting tools, volume of search results on browsed pages, and usage magnitude (# rooms, # guests). Second, ESS depends negatively on use of filtering tools, volume of search results, and average price of search results. Third, there is a significant negative correlation between ESS and purchase, whereby smaller evaluation sets are associated with higher purchase probability. Purchase is also found to depend on the average quality of all options included in the search results. The results link increases in ESS and decreases in purchases to price- and quality-based sorting strategies, while brand-name and other filtering strategies are related to reductions in ESS and increases in purchases. Arrival from other search engines and partner sites, as well as usage magnitude (# rooms, # guests), both presumed to promote purchase, are associated with larger ESS and lower purchase probability.

Our work offers theoretical, methodological, and managerial contributions. The paper advances a conceptual framework on the antecedents of ESS and purchase and presents expectations based on economic search and behavioral information-processing theories, which are then confronted with evidence from the data. To deal with the estimation challenges in this setting, we employ methodology that is tailored to the data. A key commonality in the product search engine context is that very few purchases are present relative to the number of evaluated alternatives. This makes the estimation of saturated single-equation reduced-form models for purchase infeasible; instead, identification issues can be handled by carefully combining a richer model for the data on products evaluated with a parsimonious model for purchase, thereby leveraging inference in the joint model. In addition, there are very few evaluated alternatives relative to the number of available alternatives, which together with other data limitations, makes the estimation of multinomial choice models or models with sophisticated parameter heterogeneity across customers impractical. On the managerial side, our results suggest that executives' efforts to increase purchases need not necessarily be coupled with efforts to increase ESS, and that this task can be better accomplished by promoting filtering, presenting search results in ways that focus attention to only a few options at a time, or providing more higher-priced (rather than lower-priced) options, particularly to low-usage (in terms of # rooms, # guests) users. Our conclusions section revisits the results for the antecedent variables and presents a number of managerial implications.

Literature

In marketing, there is significant literature on how consumers form evaluation or consideration sets in a retail shopping context. Early research focused on imputing consideration sets from supermarket scanner data. Wu and Rangaswamy (2003) focus on how reducing "fuzziness" alters consideration set sizes in an online grocer setting; however, the products customers consider or evaluate are imputed (not observed), and the links between the antecedents of consideration or evaluation set size and fuzziness are not examined. van Nierop et al. (2010) propose and estimate a model that captures unobserved consideration from discrete choice data. Kuksov and Villas-Boas (2010) investigate how search or evaluation costs may lead consumers not to search and not to choose if too many or too few alternatives are offered. Yet, neither van Nierop et al. (2010) nor Kuksov and Villas-Boas (2010) investigate behavioral causes of ESS. Shi and Zhang (2014) research how experience with decision aids impacts online purchase behavior over time, but do not look at antecedents or consequences of evaluation sets. Moe (2006) models customers' choice decisions as a sequence of two-stage choice models that predict which products customers view and purchase based on observed clickstream data from a retailer's website; however, the article does not focus on the antecedents of ESS or address *why* some customers evaluate more products than others.

Recently, with the advent of internet-based product search engines, considerable research interest has been directed towards search and shopping behavior in this online context. A primary focus of this research is on the drivers of click-through and purchase behavior, including the special case of the hotel-booking context that we study. Ghose, Ipeirotis, and Li (2014) study the effect of different search engine rankings on search engine revenue, based on weekly data for U.S. hotels listed on Travelocity. A hierarchical Bayes model is employed to jointly examine click-through, conversion, ranking, and rating. In addition, they employ a randomized experiment using a hotel search engine application. The results of their two-pronged empirical analysis show that a customer utility-based ranking mechanism can lead to higher search engine revenues; incorporating signals from social media into the ranking algorithm further increases revenues; and employing an active personalized ranking system, where users can customize the ranking algorithm, outperforms a static one. However, they do not focus on ESS, its antecedents, or their relationship to purchase.

A stream within the product search engine literature has focused on the quantitative estimation of user search cost. Koulayev (2014) develops a structural model to estimate search cost for online hotel bookings at a major website. He implements an identification strategy to separate the effect of search costs from other customer preferences, and estimates the median search cost for processing a page of results to vary from \$4 to \$16. Ghose, Ipeirotis, and Li (2012) also provide numerical estimates of search costs in the hotel booking context of \$39.15 per page of results, and an incremental cost of \$6.24 for evaluating one additional offer on a given page of results. A working paper by Chen and Yao (2015) also develops a structural model of search, estimates it using data from a hotel reservation site, and involves search tools, focusing on the relationship between search tools, searches, and the utility of the purchased product.

This work on product search engines also relies on the established literature on the economics of search. Search theory views consumers as rational maximizers of surplus, which in turn depends on price and quality attributes. Because consumers can improve surplus by identifying better bargains (i.e., products with higher quality, lower price, or both), they are motivated to search at the cost of spent time and cognitive effort. This trade-off between cost and expected gain motivates a stopping rule in sequential search models whereby consumers conclude the search process when the marginal cost of additional search exceeds the expected benefit (Stigler 1961; Weitzman 1979; Bakos 1997). The literature has examined the impact of changes in search cost on the amount of consumer search and the distribution of prices charged by sellers (see Stiglitz 1989 for a review), including the possibility of market failure due to high search costs. Since it is optimal for consumers who exhibit low search costs to search more extensively than consumers with higher search costs, variability in search costs could reasonably induce differences in search and purchase behavior among consumers. Recent research includes work by De los Santos et al. (2012), Honka (2014), and Moraga-González et al. (2013), who focus on the influence of search costs on search evaluations and choice, but do not consider ESS nor its antecedents and consequences.

Consumer search—including both information processing and choice—have also been studied from a behavioral perspective. One stream of work in marketing focuses on the adverse impact of large amounts of information and products on choice behavior (e.g., Malhotra, Jain, and Lagakos 1982; Townsend and Kahn 2014). While there has been disagreement over the impact of overload on choice (e.g., Jacoby 1984; Scheibehenne, Greifeneder, and Todd 2010), studies in this area, predominately employing data from forced-choice laboratory experiments, show a "paradox of information and choice" in that increasing the amount of information and products customers must choose from shifts customers' choice-processing towards simple heuristic-based strategies (Lurie 2004; Payne, Bettman, and Johnson 1993), decreases choice accuracy (Keller and Staelin 1987; Malhotra 1982) and confidence (Iyengar and Lepper 2000; Jacoby, Speller, and Kohn 1974). This literature is extended to describe why customers can experience worse decision quality and satisfaction in their selections when search costs are lowered in ordered environments (Diehl 2005; Diehl and Zauberman 2005; Goodman et al. 2013). However, the aim of this literature is to describe the rationale behind the information and choice overload phenomenon based on forced choice experiments, not to link the information and choice load phenomenon to shoppers' purchases. In addition, because of contextual differences between behavior in lab experiments and purchases on product search engine sites, the behavioral literature is unable to study many of the antecedents of consideration set size that we study.

Our results aim to quantify and validate various aspects of the theories discussed here. To the best of our knowledge, our work is the first empirical investigation of the antecedents of ESS (variables that relate to some customers choosing larger evaluation sets than others) and the connection between ESS and purchase. Notably, in contrast to previous work on consideration sets (that does not investigate behavioral causes of ESS), information and choice overload (that does not link behavioral laboratory findings with shoppers' purchases), and consumer information search (which focuses on diagnosing the effects of search costs on choice but not the antecedents of ESS), our research contributes to an area that has received scant prior attention in the marketing, information technology, and economics literatures.

Theory

Our article builds upon theoretical paradigms from previous work, including economic search theory (Moe 2006; Weitzman 1979), which models the amount of search as determined by the trade-off between the expected improvement in utility from search and the incremental cost of search. We also build upon behavioral search theory (e.g., Alba and Hutchinson 1987) which invokes theories from psychology to understand cognitive processes underlying consumer search. In particular, we focus on the modeling framework of Moorthy, Ratchford, and Talukdar (1997) (henceforth referred to as MRT), who combine economic and behavioral search theories to generate empirically testable implications.

MRT provide a bridge between the behavioral and economic literatures on consumer search and information processing. The model emphasizes the importance of consumers' prior beliefs about product attributes. Key insights from MRT are that the amount of search should be increasing in: (1) relative brand uncertainty (uncertainty about which brand provides higher utility); (2) individual brand uncertainty (uncertainty about the utility of a brand); (3) involvement (relevance of a product class to a consumer, see, e.g., Beatty and Smith 1987); and (4) the unit cost of search. MRT note that all of these factors are related to consumer expertise, i.e., knowledge about product attributes, and ability to evaluate a larger number of product attributes. Consumer expertise is generally increasing in the level of consumer experience. With respect to the purchase funnel (see, e.g., Alba et al. 1997; Howard and Sheth 1969)**,** the further a consumer is from the final purchase decision (i.e., higher up in the funnel), the greater the relative and individual brand uncertainty. It then follows from MRT that such consumers would search more. The framework behind our analysis is discussed next.

Conceptual Framework

Figure 1 presents the conceptual framework for our analysis. The key dependent variables are: (i) ESS and (ii) whether the customer purchased a hotel reservation. These variables correspond to the two "stages" in two-stage search theory (e.g., Moe 2006). In the context of the purchase funnel in marketing (see, e.g., Alba et al. 1997; Howard and Sheth 1969), wherein a user's consideration set tends to get progressively narrower as the user approaches a purchase decision, that decision would correspond to the end of the funnel, whereas the formation of the evaluation set could be at any point above it. Figure 1 also shows the relations between the dependent variables and the explanatory variables, which are grouped into four sets: Experience with Site, Use of Search Tools, Search Results, and Customer Characteristics. The constituent variables of each set are described in Table 1. We next provide a theoretical discussion of how the antecedent variables in our empirical model relate to ESS and purchase.

Experience with Site

Prior to the visit of interest—the focal visit—customers may have had a history of interactions with the product search engine that could affect behavior. Prior visits can be a combination of search related to the current trip and to other unrelated trips. Search theories (Stigler 1961; Alba and Hutchinson 1987; Zhang, Fang, and Sheng 2007) suggest that customers who have more prior experience with the site and have gained familiarity with evaluating the product category, will have lower unit search cost during the focal visit. Alba and Hutchinson (1987) provide a theoretical pathway, resulting from the development of a more refined cognitive structure and an improved ability to identify important information, through which experience and familiarity reduce the cognitive effort required to recognize and evaluate relevant alternatives. Based on MRT, this reduction in the marginal search cost reduces the cost of considering additional products and implies a larger evaluation set for the focal visit.

However, one could argue that prior experience can also be associated with smaller ESS because previous search related to a trip would have moved the customer down the purchase funnel (Alba et al. 1997; Howard and Sheth 1969), narrowing down the set of options and leading to a smaller ESS. Prior unrelated search would also have increased the knowledge of the customer about the search process on the product search engine, resulting in a more efficient and streamlined search, also yielding a smaller ESS. Therefore, the net association between prior experience and ESS would depend on which of these opposing effects dominates, making it an empirical question of considerable interest.

Use of Search Tools

Filtering and sorting are the two tools that are most widely used at product search engines to truncate and organize the initial search results along a variety of price and quality dimensions and make them more suitably aligned with customers' preferences. In particular, filtering enables customers to pare down the results based on specific requirements (e.g., five star hotels or other must-haves). It is more likely employed by customers who are in later stages of decision-making (Hauser and Wernerfelt 1990; Payne 1976) and, therefore, have lower individual and relative brand uncertainty. It is also possible that the use of filtering reduces relative brand uncertainty during the search process itself, due to the presence of fewer brands in the filtered results. Either way, MRT predict that the use of a filtering tool would be associated with a smaller ESS.

In contrast, sorting is likely to be employed by customers who are more flexible, less certain about their requirements, or are in earlier stages of the purchase funnel. In the MRT framework, such customers are likely to have higher individual and relative brand uncertainty. Sorting helps reorder the search results along price and quality dimensions to better understand trade-offs among product attributes and more easily identify suitable options (Häubl and Trifts 2000). For example, sorting hotels by star rating could help customers understand the trade-off between quality and price across different star ratings; repeating the sorting exercise for distance from an ideal location could elicit trade-offs across other product attributes as well. As a result, customers who sort are expected to evaluate more options to understand various trade-offs, which reflects their flexibility, uncertain requirements, or the fact that they are at an earlier stage in their decision-making process. In addition, sorting exposes customers to larger cross-sections of options, whereby they are likely to evaluate more of them.

Customers can arrive at a product search engine in several ways. For instance, they may directly navigate to the product search engine, or they may follow a link from a search engine (e.g., Google or Yahoo) or partner site (e.g., another product search engine). Customers arriving from other referring sites are likely to be conducting a relatively broader search (across multiple websites), because they may have greater individual and relative brand uncertainty (MRT), be at an earlier stage of search in the purchase funnel (Hauser and Wernerfelt 1990; Payne 1976), or have lower search costs (MRT, Zhang et al. 2007). In every case, customers coming from referral sites are likely to search more and have higher ESS.

Search Results

When the number of options that match a user's search query is large, relative brand uncertainty will tend to be high as well. This is because we expect customers presented with more options (in their search results) will tend to consider more options (in their evaluation sets), and consequently be less certain about which brand is the best choice. Following MRT, greater brand uncertainty would be associated with a larger ESS. However, it is an empirical question whether customers presented with more options will indeed consider more options.

Customer Characteristics

Customers vary in the magnitude (e.g., number of nights, number of adults, etc.) and specificity (e.g., family amenities) of potential purchases, both of which relate directly to customer involvement. In the MRT framework, greater involvement is associated with more searching. Another customer characteristic is the time to consumption (i.e., the number of days between the search and the actual trip). The greater the time to consumption the earlier the consumer is likely to be in the purchase funnel, thereby having greater individual and relative brand uncertainty and a larger ESS. Also, more days to check-in would be associated with lower time pressure, which would tend to decrease unit search costs (the fourth factor in MRT), resulting in more searching.

Relationship Between Evaluation Set Size and Purchase Decision

The larger the evaluation set, the more trade-offs customers need to consider when making their purchase decisions. This increases the decision difficulty (Schwartz 2004), complicates choice, and reduces the likelihood of purchase (Iyengar and Lepper 2000). In addition, search cost theory predicts that customers who evaluate more options likely do so because they face lower search costs (Stiglitz 1989). Customers who face lower search costs are likely to evaluate more options at any given website. They are also likely to evaluate options across websites and visit more websites. Thus, the probability of such customers purchasing from any one of these websites is likely to be lower. Also, customers at an earlier stage in the purchase funnel consider a larger number of alternatives and because they are in earlier stages of decision-making, they are less likely to purchase. On the other hand, it is also possible that when evaluating a larger set of options, customers simplify their decision and make a choice; however, in our case we would expect to find a negative association between ESS and purchase.

Statistical Model

In line with the conceptual framework in Figure 1, our empirical model aims to address the relationship between ESS and purchase behavior given their antecedents. Although it is desirable to pursue joint modeling to allow dependence between the variables, mitigate misspecification problems, and exploit gains in efficiency, specifying a suitable joint distribution in our context is difficult because the dependent variables are of different types – ESS is a count variable and purchase is binary. Until recently, this difficulty has required sacrifices in modeling generality or flexibility for the sake of retaining computational tractability. To deal with these drawbacks, in our application we start with two popular univariate models—a negative binomial model for count data and a probit model for binary data—and then employ a copula to couple them into a joint model for the two outcomes.

Copula models are very valuable, general, flexible and conceptually simple, but have yet to be employed widely in applications despite their potential. The premise behind the use of copula models rests on the idea that one can establish an isomorphic relationship between one random variable and another via a transformation through the uniform distribution. In fact, random number generators proceed in precisely that way – they produce random draws from a given cumulative distribution function (cdf) by starting with uniform random variables that are then transformed by the inverse of that cdf. The key insight is that if one is interested in joint modeling of random variables for which there is no suitable joint distribution, the isomorphic mapping of each of them to a distribution that allows flexible joint modeling (e.g., Gaussian) provides the necessary pathway; reversing the mapping recovers the original random variables and marginal distributions, but they can now be correlated (the original marginal models are a special case of the copula when there is no correlation). Formal details are presented next.

The origins of copula models can be traced back to Sklar (1959). More recent reviews, applications, and estimation techniques are presented in Joe (1997), Nelsen (2007), Trivedi and Zimmer (2005), Pitt et al. (2006), Jeliazkov and Lloro (2011), and Danaher and Smith (2011). Formally, a copula C maps the unit hypercube $[0,1]^q$ to the unit interval $[0,1]$ that satisfies:

1. $C\left(1,\ldots,1,a_p,1,\ldots,1\right)=a_p$ for every $p\in\{1,\ldots,q\}$ and all $a_p\in[0,1]$;

2. $C\left(a_1,\ldots,a_q\right)=0$ if $a_p=0$ for any $p\in\{1,\ldots,q\}$;

3. C is q-increasing, i.e., any hyperrectangle in $[0,1]^q$ has non-negative C-volume.

As a consequence of these conditions, the copula can be viewed as a q-dimensional distribution function with uniform marginals, each of which can be related to an arbitrary known cdf $F_j(\cdot)$ for $j=1,\ldots,q$. Specifically, if a random variable u_j is uniform $u_j \sim U(0,1)$ and $y_j=F_j^{-1}\left(u_j\right)$, then basic probability arguments imply that $y_j \sim F_j(\cdot)$. Therefore, if the variables $y_1,\ldots,y_q$ have corresponding univariate cdfs $F_1(y_1),\ldots,F_q(\cdot)$, each taking values in $[0,1]$, then a copula can be used to link or "couple" those univariate marginal distributions to produce the joint distribution function $F(y_1,\ldots,y_q)$ that can be written as

$$F(y_1,\ldots,y_q) = C(F_1(y_1),\ldots,F_q(y_q)). \tag{1}$$

The key feature embodied in (1) is that copulas provide a way to model dependence among multiple random variables when their joint distribution is not easy to specify, e.g., in cases where the marginal distributions $\{F_j(\cdot)\}$ belong to different parametric classes, as in our application.

There are many families of copulas, but the Gaussian copula is a natural modeling choice with desirable theoretical properties (Danaher and Smith 2011; Joe 1997; Nelsen 2007; Pitt, Chan, and Kohn 2006; Trivedi and Zimmer 2005).[5] The Gaussian copula is given by

$$C(u \mid \Omega) = \Phi_q(\Phi^{-1}(u_1),\ldots,\Phi^{-1}(u_q) \mid \Omega), \tag{2}$$

where $u = (u_1,\ldots,u_q)'$, Φ represents the standard normal cdf, and Φ_q is the cdf of a multivariate normal vector $z = (z_1,\ldots,z_q)'$, $z \sim N(0,\Omega)$, where Ω is in correlation form with ones on the main diagonal. The data generating process implied by the Gaussian copula in (2) is given by

$$y_{ij} = F_{ij}^{-1}\left(\Phi(z_{ij})\right), \quad z_i \sim N(0,\Omega), \quad i = 1,\ldots,n, \quad j = 1,\ldots,q, \tag{3}$$

where F_{ij} is a cdf specified in terms of a vector of parameters θ_j and covariates x_{ij}, q is the dimension of each vector $y_i = (y_{i1},\ldots,y_{iq})'$, and n is the sample size. Note that the correlation matrix Ω for the latent z_i induces dependence among the elements of y_i and that the copula density will typically be analytically intractable due to the cdf transform.

An important benefit of the modeling framework in (3) is that it provides a useful way of linking observed choice outcomes to an underlying behavioral random utility model. Researchers

[5] Gaussian copulas, among several others, are not suitable for the study of extreme events because of asymptotic independence in the extreme tails, but this is not a concern in common practical situations or in the current context.

trained in thinking about discrete data models from a random utility perspective may recognize that z_i in (3) may be thought of as a latent utility vector with the mapping from z_i to y_i linking those latent utilities to observed outcomes. The structures in (1), (2), and (3) are general and apply to both discrete and continuous outcomes, with the distinction that $F^{-1}(\cdot)$ is one-to-one when y_{ij} is continuous and many-to-one when it is discrete. In the latter case, it is necessary to integrate over the relevant values of z_i to obtain the joint distribution of the observed y_i.

In our paper, the Gaussian copula model is also used to addresses the problem of sample selection, which arises because our count variable (ESS) determines whether our binary purchase outcome is observed. This is because an explicit purchase decision is only conceivable for customers who click on at least one option and not for those who do not click at all. Since the copula allows for correlation (unlike independent models), our model can mitigate potential misspecification, due to the possible presence of incidental truncation (informative missingness).

In our joint model, the count variable for size of the evaluation set y_{i1} follows a negative binomial distribution

$$y_{i1} \sim NB(\lambda_i, \alpha), \tag{4}$$

with probability mass function (pmf) given by

$$\Pr(y_{i1} \mid \beta_1, \alpha) = \frac{\Gamma(\alpha + y_{i1}) r_i^{\alpha} (1 - r_i)^{y_{i1}}}{\Gamma(1 + y_{i1}) \Gamma(\alpha)}, \quad \lambda_i > 0, \quad \alpha > 0, \tag{5}$$

where $r_i = \alpha / (\alpha + \lambda_i)$ and $\lambda_i = \exp(x_{i1}' \beta_1)$, and x_{i1} is a vector of covariates that determines y_{i1}. The negative binomial distribution in (5) has mean λ_i and variance $\lambda_i (1 + \lambda_i / \alpha)$. It can therefore be seen that the parameter α controls the degree of over-dispersion – the conditional variance can be much larger than the conditional mean for small values of α, but in the limit, as

$\alpha \to \infty$, the negative binomial distribution approaches the Poisson, where the conditional mean and variance are equal (i.e., we have equidispersion). Therefore, small estimated values of α would be evidence that the Poisson specification is not suitable and overdispersed models are more appropriate. An important reason for considering the negative binomial distribution is also that, by construction, the model is a mixed Poisson model that allows for unobserved heterogeneity (cf. mixed logit or random effects models). This can be seen by employing the well-known result that the negative binomial pmf in (5) can be represented as

$$f_{NB}\left(y_{i1} \mid \lambda_i, \alpha\right) = \int f_{Po}\left(y_{i1} \mid \lambda_i u_i\right) f_G\left(u_i \mid \alpha\right) du_i,$$

where the individual heterogeneity u_i has a gamma distribution $f_G\left(u_i \mid \alpha\right) = \alpha^{\alpha} e^{-\alpha u_i} u_i^{\alpha-1} / \Gamma(\alpha)$, and $\left[y_{i1} \mid \lambda_i u_i\right]$ is Poisson. This model feature is important because individual heterogeneity may be a key source of correlation between the outcomes. These and other aspects of negative binomial models are carefully reviewed in Cameron and Trivedi (1998) and Winkelmann (2008).

The second outcome of interest in our application is the binary purchase decision, which is modeled (marginally) through a probit model

$$\Pr\left(y_{i2} = 1 \mid \beta_2\right) = \Phi\left(x_{i2}' \beta_2\right). \tag{6}$$

The size of the evaluation set in (5) and the decision to buy in (6) are linked through the copula in (2) – a key feature of the joint model is that it can accommodate non-randomly missing purchase decisions which are unavailable when evaluation sets are empty (independent equation models necessarily treat this as ignorable truncation). Therefore, the likelihood consists of two parts corresponding to the subsamples of customers with empty and non-empty evaluation sets

$$f\left(y \mid \beta, \alpha, \rho\right) = \left\{ \prod_{i: y_{i1}=0} f\left(y_{i1} \mid \beta_1, \alpha\right) \right\} \left\{ \prod_{i: y_{i1}>0} f\left(y_i \mid \beta, \alpha, \rho\right) \right\}, \tag{7}$$

where the terms in the first product are given by the pmf in (5), the terms in the second product are obtained from the joint cdf (2), where $\beta = (\beta_1', \beta_2')'$, and ρ is the off-diagonal correlation element in $\Omega = \begin{pmatrix} 1 & \rho \\ \rho & 1 \end{pmatrix}$. To evaluate the terms in the second product in (7), we use the following steps. First, we compute the cdf $F_{NB}(y_{i1} \mid \lambda_i, \alpha)$ for the negative binomial distribution by summing the pmf in (5) for values less than or equal to y_{i1}. This determines unique cutpoints

$$\begin{aligned} \gamma_{i1,U} &= \Phi^{-1}\left(F_{NB}(y_{i1} \mid \beta_1, \alpha)\right) \\ \gamma_{i1,L} &= \Phi^{-1}\left(F_{NB}(y_{i1} \mid \beta_1, \alpha) - \Pr(y_{i1} \mid \beta_1, \alpha)\right) \end{aligned} \tag{8}$$

so that $\Pr(z_{i1} \le \gamma_{i1,U}) = F_{NB}(y_{i1} \mid \beta_1, \alpha)$ and $\Pr(\gamma_{i1,L} < z_{i1} \le \gamma_{i1,L}) = \Pr(y_{i1} \mid \beta_1, \alpha)$ for $z_{i1} \sim N(0,1)$. Similarly, the range for $z_{i2} \sim N(0,1)$ that is consistent with the observed outcomes for the probit model from equation (6) is given by $\gamma_{i2,L} = -\infty$ and $\gamma_{i2,U} = -x_{i2}'\beta_2$ if $y_{i2} = 0$, or $\gamma_{i2,L} = -x_{i2}'\beta_2$ and $\gamma_{i2,U} = \infty$ if $y_{i2} = 1$. The joint probability required for the second product in (7) is given by

$$f(y_i \mid \beta, \alpha, \rho) = \int_{\gamma_{i2,L}}^{\gamma_{i2,U}} \int_{\gamma_{i1,L}}^{\gamma_{i1,U}} f_N(z_i \mid 0, \Omega)\, dz_{i1}\, dz_{i2}, \tag{9}$$

where $f_N(\cdot)$ represents the multivariate normal density. The evaluation of (9) can be done by numerical integration in bivariate models, but requires simulation in higher-dimensional settings (Jeliazkov and Lee 2010; Jeliazkov and Lloro 2011). The copula maintains the marginal models $f(y_{i1} \mid \beta_1, \alpha)$ and $f(y_{i2} \mid \beta_2)$ by construction, but the joint probability in (9) and the conditionals $f(y_{i1} \mid y_{i2}, \beta, \alpha, \rho)$ and $f(y_{i2} \mid y_{i1}, \beta, \alpha, \rho)$ depend on all parameters and the correlation ρ, whose sign determines the direction of the dependence. Heterogeneity can be one of the key drivers of that correlation and can be modeled by writing the model in (6) in latent variable form

$$y_{i2} = 1\{x'_{i2}\beta_2 + v_i > 0\},$$

where $v_i \equiv (w_i + \varepsilon_i) \sim N(0,1)$ is a composite error term consisting of a heterogeneity component w_i and idiosyncratic errors ε_i. Whereas in cross-sectional settings, heterogeneity cannot be decoupled from the error term, longitudinal data would allow one to separate the impact of the errors from that of the heterogeneity and explore possible correlation between heterogeneity and observables (Chib and Jeliazkov 2006).[6] Thus, the expected benefit of pursuing longitudinal studies in the future is quite high, as it could dramatically improve our understanding of individual heterogeneity.

In our application, estimation of the model proceeds by the Accept-Reject Metropolis-Hastings (ARMH) algorithm (Chib and Greenberg 1995; Chib and Jeliazkov 2005; Tierney 1994), which simulates a random sample from the joint distribution of the model parameters $\theta = (\beta', \alpha, \rho)'$ given the data. One benefit of the ARMH simulation algorithm is its ability to produce better-behaved draws from the posterior distribution than regular Metropolis-Hastings samplers; another is that the building blocks of the algorithm can be easily employed to estimate the marginal likelihood for competing models for the purposes of formal model comparison. Details on the ARMH algorithm and marginal likelihoods are provided in the Appendix.

Before moving to a discussion of our data set, we compare our approach to the work of Ghose, Ipeirotis, and Li (2014), who build a hierarchical Bayesian model for a setting similar to ours. Specifically, in their approach the individual equations contain unobserved latent factors (modeled in a higher level of the modeling hierarchy) that induce co-variation among the

[6] In cross-sectional settings the normalization $\text{var}(v_i) = \text{var}(w_i + \varepsilon_i) = 1$ fixes the scale in the probit model, while in longitudinal settings it is common to assume $\text{var}(\varepsilon_{it}) = 1$ with $\sigma_w^2 = \text{var}(w_i)$ being free; this renormalization can rescale the point estimate of β_2 by $1/\sqrt{1+\sigma_w^2}$ without affecting covariate effects.

observables. While the two approaches provide alternative ways of constructing a joint model for the data, it is also important to recognize the existence of important complementarities. For instance, a copula can be easily inserted as a component in a larger hierarchical model and, conversely, hierarchical models can also be combined through copulas, offering important synergies that can be exploited in suitable settings. Nevertheless, in our particular context the copula modeling approach was chosen for the following reasons. First, it provides an appealing parsimonious strategy for directly combining the marginal distributions into a joint model that simplifies both modeling and estimation (based on a single-block MCMC algorithm). Second, despite its relative parsimony, the correlation structures captured by the copula model are quite flexible and attain the Frèchet bounds (see, e.g., Trivedi and Zimmer 2005). Third, in working with cross-sectional data (our unit of analysis is a customer visit as opposed to a longitudinal set of weekly hotel observations as in Ghose, Ipeirotis, and Li 2014), latent effects at the individual level would not be identified as they enter the modeling hierarchy through a single observation, making a copula approach preferable; nonetheless, our results can be used to provide real-data evidence on several aspects of the randomized experiments performed by Ghose, Ipeirotis, and Li (2014) to describe consumer behavior and product search engine revenues. Finally, by its very nature, a copula permits the user to retain the marginal distributions of the individual variables y_{ij}, which is rarely the case in non-Gaussian hierarchical models where integration over the latent variables in the hierarchy may make the marginal distributions intractable and difficult to summarize.

Data

The data set for our study is from a leading travel search engine and was obtained as part of a data grant from the Wharton Customer Analytics Initiative (WCAI). The data set includes observations on customers from all over the world who visited the travel search engine website during a two-week window, October 1-14, 2009, and sought hotels in the city of Budapest, Hungary, with check-in dates ranging from October 1, 2009, to November 1, 2010. The records for each customer included referral site (if they arrived from another search engine or partner site), any filtering or sorting employed, the set of hotels the search engine displayed, the subset of hotels the customer clicked on to evaluate in greater detail, and the hotel booked, if any. We extracted the records and other metadata from each user's last visit within the two-week study window, which we call the focal visit, and summarized previous activity in additional data fields. Moreover, we employed clickstream data parsed from server logs from January–October 2009 to get a broader picture of each customer's visit and purchase history prior to the two-week study window, and removed observations with missing or spurious data.[7] The resulting data set comprises 3,182 customer focal visits; for each focal visit, we keep track of the full list of hotels shown, evaluated, and purchased. We also include the frequency of prior visits for hotel and non-hotel searches and purchases, which serve as proxies for involvement and search cost.

Each page of results shown by the travel website listed up to 25 hotels, with fewer hotels displayed when the customer provided filtering criteria, or when there were simply not enough available hotels for the requested date or location. Both clicks and purchases are infrequent. A total of 69,349 options were listed, of which 2,372 were clicked on – 58.4% of customers (1,858 out of 3,182) did not click any hotel, 26.2% (835) clicked on only one, 7.8% (249) clicked on two, 5.3% (168) clicked on three or four, and 2.3% (72) clicked on 5 or more options. Most (63,345 of 69,349 = 91.3%) of the hotel listings were on the first page (listings ranked 1 through 25), indicating that few customers bothered to look past the first page of results, even though a

[7] We eliminated a small set of observations attributable to non-human activity that exhibited excessive and abnormal numbers of searches, likely due to automated scripts ("robots").

small number of bargain-seekers continued to look through as many as 3 pages. Of the 3,182 customers, 121 made a booking, generating customer, click, and listing conversion rates of 3.8%, 5.1%, and 0.17%, respectively.

Several optional filtering criteria could be employed to reduce the number of hotels presented – 555 (17.4%) of customers used such filtering criteria, while 384 customers (12.1%) sorted the search results (e.g., by price, star rating, user rating, hotel name, city name, or distance to a landmark) to help organize and view their results. In addition, of the 3,182 customers, 302 (9.5%) requested more than one room, 182 (5.7%) were in a group of 3+ adults, and 157 (4.9%) were traveling with children. The most commonly sought trip durations were 3 nights (29.5%), followed by 2 nights (24.9%), 4 nights (17.9%), 1 night (12.6%), 5 nights (6.4%), 7 nights (3.4%), and 6 nights (2.2%). Only 100 customers (3.1%) were interested in booking 8+ nights.

We used several metrics to evaluate the quality of the sets of hotels that the product search engine displayed and those that the customer clicked on to define the evaluation set. There are five dimensions of quality: (i) the star rating provided by the travel website (henceforth, "site rating"), (ii) the user rating produced by customers who wrote reviews, (iii) the number of reviews written by customers (henceforth, "user rating count"), (iv) the geographic proximity of the hotel to the most desirable areas in the city (henceforth, "distance to best areas"), and (v) whether the hotel belongs to a well-known brand. Site ratings and user ratings are measured from 1 to 5, the brand flag is a binary indicator variable, and we measure geographic proximity in meters to the nearest highly desirable area. We identified such areas in consultation with graduate students from Budapest, used Google Earth to plot the identified regions, and computed the shortest Euclidean distances to the identified regions using the longitude and latitude coordinates of each hotel. Table 1 lists the full set of variables in our data set, along with their detailed descriptions. In the estimation, all variables except categorical variables were stabilized by the transformation $\ln(1+x)$, which is similar to the usual logarithmic transformation but with the benefit that it retains the sign of the original variable and works even when $x=0$.

Results

We perform statistical inference by tackling both model and parameter uncertainty. In particular, we examine model uncertainty due to variable selection (i.e., which variables ought to be in the model) and specification uncertainty (independent versus copula model, endogeneity, etc.). This is followed by a careful assessment of the parameter estimates from the best model.

To address variable selection, we estimate five model variants, labeled M1-M5, that allow us to consider alternative measures of previous search and hotel quality, in addition to allowing us to build parsimonious models and avoid problems with collinearity.[8] Each model variant has two components, corresponding to the two dependent variables. The ESS component has four groups of independent variables pertaining to the antecedent categories discussed earlier (see Figure 1), in addition to price and quality controls for ESS. The purchase component includes controls that measure the average price and quality of the evaluation set. In the copula model, ESS and purchase probability are linked through the copula correlation parameter ρ. The results for specifications M1-M5 are presented in Table 2.

M1 includes the sets of variables noted above. M2 adds three covariates relevant to past experience to investigate whether proxies for involvement and search cost could influence ESS. Table 2 reveals that the addition of these variables lowers the marginal likelihood of M2 relative to M1, suggesting that the data does not support the presence of these variables in the model. M3

[8] Our Web Appendix presents pairwise correlations for the variables. Apart from the alternative hotel quality and previous search variables (which are employed in separate models), 163 of the 169 correlations (96%) are less than 0.4, and variance inflation factor scores are well below 6, raising no red flags for estimation (Hair et al. 2010).

through M5 include different sets of quality-based controls for ESS and purchase. While M1 and M2 include quality controls provided by the search engine (site rating), M3 includes quality-based controls generated by users (user rating and the number of user ratings), as well as the distance to desirable areas. M4 includes the distance to these areas and an indicator for brand-name hotel. M5 considers the user generated ratings only. Before we delve into a detailed discussion of our empirical results, we mention that in all models, the overdispersion parameter $\ln(\alpha)$ is close to 0 indicating that α is close to 1, confirming that the negative binomial model is a more appropriate choice than a Poisson model.

We begin by discussing M1 because it has the best marginal likelihood. First, we describe the drivers of ESS. Regarding ***past experience***, the number of product category (hotel) searches in the past 3 months is found to be positively associated with ESS during the focal search, whereas the number of products (hotels) evaluated (i.e., clicked on) in the two-week study period prior to the focal visit is not associated with ESS. Concerning the use of ***search tools***, we find that when customers filter their search results, evaluation sets are smaller; in contrast, when customers sort their search results, evaluation sets are larger. We also find that customers who either used a search engine or a partner website immediately prior to arriving at the travel search engine have larger evaluation sets. For ***search results***, we find that the volume of search results (# of options that match the customer's query) is negatively associated with ESS, while the volume of search results on pages browsed (# of options on pages actually seen) is positively associated with ESS (# of options clicked to examine more closely). ***Customer characteristics*** related to the usage magnitude (# adults, # children, and # number rooms requested) are found to increase ESS, while the length of stay (number of nights) is not found to significantly affect ESS. Moreover, time to consumption (the number of days prior to the trip's first night) is not associated with ESS. Finally, for ***hotel-based search result controls***, we find the greater the average price of the hotels in this customer's search results, the smaller the ESS, while average site rating quality does not significantly impact ESS.

Now, we discuss the results of the purchase probability equation. For the ***association between ESS and purchase probability***, we find the copula correlation parameter ρ is negative and statistically different from zero, indicating that a larger ESS is associated with lower purchase probability. Regarding ***ESS hotel-based controls***, we find a significant negative association between average star rating of the hotels in a customer's ESS and purchase probability, but do not find a significant relationship for the average price of the hotels in a customer's ESS.

As noted earlier, the formal model comparison between M1 and M2 demonstrates that the inclusion of additional controls for past experience is not warranted in the ESS model. Similarly, comparisons between M1 and M3-M5 imply that additional user-generated and data-based controls of hotel quality do not contribute to explaining ESS and purchases beyond the search engine quality controls employed in M1, while the results of the original price and site rating quality control in both stages are consistent with whichever other controls are included in the different variations of the models. This is of particular interest because the price and quality covariates in the second equation of M1 are highly correlated; however, the fact that models M1-M5 include different sets of covariates in the second equation, several of which do not include the price and site rating simultaneously, helps alleviate concerns about the impact of collinearity, because M1 is still preferred despite the correlation.

It is also important to emphasize, however, that our conclusions are robust not only because of the large log-marginal likelihood differences between M1 and M2-M5, but also because the parameter estimates in Table 2 support similar conclusions about the underlying theoretical expectations as those supported by M1. In other words, although the data support M1 overwhelmingly, the other models we have examined do not violate our main conclusions.

Additional Models

The model comparisons and parameter estimates in Table 2 provide empirical evidence on many of the theoretical considerations that motivate our analysis by helping to gauge the practical relevance of alternative sets of covariates. To supplement our investigation, we now present additional analysis aimed at validating key modeling features of our empirical approach. In Table 3, we report the log marginal likelihood fit statistics and numerical standard errors of the additional models.

Independent Equations Model. One main reason for employing a copula is that it allows for correlation between the outcome variables. While it can be seen in Table 2 that estimates of the correlation parameter ρ differ from zero across models M1-M5, we also conducted formal marginal likelihood comparisons between the copula specification and its independent equation counterpart (the copula model reduces to a set of two independent equations when $\rho = 0$). Table 3 shows that when ρ in M1 is restricted to 0, the log marginal likelihood drops by 11.54, providing strong evidence in favor of joint modeling.[9] Not only does this imply that the outcomes may be jointly determined, hinge on common unobserved factors, or be correlated because of individual heterogeneity, but that incidental truncation is indeed a central feature in this context that must be properly accommodated (recall that y_{i1} determines if y_{i2} is observed). In contrast, the independent equation model necessarily implies non-informative (ignorable) missingness.

[9] Similarly, when ρ was restricted to 0 in M2-M5, the log marginal likelihoods of the restricted (independent equations) models were lower by at least 10 than those of the corresponding copula models.

Conditional Model. Next, we estimated another version of M1 where the count ESS variable y_{i1} enters the purchase equation for y_{i2} as an additional covariate (beyond determining whether y_{i2} is observed). The conditional model had a slightly lower (by 2.6 on the natural log scale) marginal likelihood than our baseline model, and although it can be criticized on theoretical grounds,[10] it nonetheless supported one of our main conclusions that a larger evaluation set relates to lower probability of purchase. In particular, the coefficient on y_{i1} in the y_{i2} equation was -0.235 with a standard deviation of 0.09, indicating a negative association between ESS and purchase likelihood. Interestingly, the conditional model also provided evidence in favor of joint modeling – Table 3 shows that restricting ρ to equal 0 reduces the log marginal likelihood of the conditional model by 6.94 (from -4189.5 when $\rho \neq 0$ to -4196.44 when $\rho = 0$).

Endogenous Clicked Price and Quality (Fitted Regressors) Model. Finally, we sought to address potential concerns about endogeneity of the price and quality hotel-based control variables in the second equation. This was done by modifying our baseline model, M1, so that the covariates in the second equation were replaced, similarly to the second stage of two-stage estimators, by their fitted values from a regression on consumer and search characteristics. The instruments included average user rating, average count of user ratings, average proportion of brand-name hotels, and average proximity to good locations, all of which were computed for clicked hotels. The modeling produced comparable parameter and marginal likelihood estimates with the latter being slightly higher than, yet statistically indistinguishable from, that of M1. However, inclusion of the fitted regressors reduced covariate variability and led to considerable

[10] This model assumes not only that y_{i1} and y_{i2} are jointly determined, but implies a direction of causality from y_{i1} to y_{i2}. However, in theory, the opposite is also possible, e.g., customers who know they must make a purchase may decide to adjust their search effort and evaluate a different number of alternatives (ESS).

deterioration of the likelihood surface. In turn, that adversely affected parameter identification, mixing and convergence of the estimation algorithm, and the stability of the marginal likelihood estimate, as can be seen by its abnormally high NSE. Therefore, on balance, the copula model with fitted regressors was unsuitable for our application and we are hesitant to advocate it over the simpler and better-behaved model M1, even though the modeling could be promising in other similar contexts.

Interpretation of Key Results

A striking and robust result is that the evaluation set size (ESS) is negatively associated with purchase. This finding is consistent with the notion that consumers further along in the purchase funnel, and who are closer to making a purchase decision, tend to evaluate fewer options. Our four categories of ESS antecedents yield some nuanced observations. First, consider the role of past experience. We have observed that greater past experience in the product category at the site is connected to customers evaluating a larger set of options, which is consistent with such customers having lower search costs due to their familiarity with the site and their ability to make quicker and finer comparisons of product attributes.

Second, consider the role of search tools. We find that visitors arriving at the product search engine from outside referral sites (search engines and partner sites) tend to evaluate more options. A potential explanation is that customers who arrive from such sites are, at minimum, in search mode, if not having already conducted search and developed domain knowledge leading to lower search costs, making it easier to evaluate more options. We also find that customers who sort tend to have larger ESS, while customers who filter tend to have smaller ESS. Therefore, the act of sorting or filtering may indicate how far along customers are in their search process. Also, customers that use the sorting tool may have high relative and/or individual brand uncertainty, and may thus be more likely to evaluate many options and be less likely to purchase. The reverse is true for customers who filter.

Third, consider the role of the search results. Interestingly, we find the volume of search results (# of options that match the customer's search query) is negatively associated with ESS (and thus positively associated with purchase). On the other hand, a related metric exhibits the opposite association, with the volume of search results on pages browsed (# of options on pages actually seen) being positively associated with ESS (and thus negatively associated with purchase). This latter relationship can be explained by MRT, since browsing more options on more pages is an indication of high search activity, which is consistent with low search costs, high involvement, and high relative and individual brand uncertainty -- all reasons to expect high ESS. However, the former relationship may have a behavioral explanation: if the product search engine returns more results, the customer may believe that the search engine has done a better job at finding good options. The customer is then content to scrutinize fewer options, hence lower ESS, when the volume of search results is high. We conclude that it may be prudent to organize search results into many pages that have a few options on each page. This would create the feeling that the product search engine provides value by offering many choices, yet at the same time would limit the number of results per page to avoid overwhelming customers, with the intent to keep information processing costs low, ESS small, and purchase probability high.

Fourth, consider the role of customer characteristics. Consistent with MRT, we have shown that the variables room count, adult count, and child count, which capture the magnitude of the potential purchase and are thus akin to measuring involvement, are associated with larger evaluation sets. However, we did not observe a significant relationship between time to consumption, as measured by days until check-in, and ESS. This may, in part, be due to an inability to distinguish time-constrained patrons whose demand may be inelastic from last-minute bargain hunters who may be quite flexible and need not necessarily commit to one option or another. Developing marketing tools for segmenting this part of the market may well prove very profitable. We also note that some customer characteristics can proxy for others not available in our data (e.g., child count could be used to classify business vs. leisure travelers), which can help control for unobserved heterogeneity.

Finally, consistently across the variations of our models, we find the lower the average hotel quality ratings in the ESS, the greater the purchase likelihood of consumers even while controlling for the average price of the hotels in ESS. One reason for why this may occur is that consumers on this product intermediary's website are more interested in lower quality alternatives, and for lower quality alternatives, price may be unimimportant or have less variance, meaning it is less important to consumers' purchase decisions. In addition for our empirical sample, perhaps consumers who are more willing to purchase higher quality alternatives are less likely to purchase on the product intermediary's website but instead purchase the higher quality alternative hotel directly on its own website or reservation system, which causes conversion issues for product intermediaries. Interestingly, we also find that the average price of hotels in the ESS does not appear to impact purchase decisions for our sample, but do find the average price of all hotels listed in the search results does influence ESS. Consequently, it appears price plays an inital screening type role in the first-stage of decision making which influences ESS, but then has less an impact in the second-stage of decision making stage once hotels are within the ESS and the consumer must decide whether to purchase a hotel.

Conclusions

This paper has developed a conceptual framework linking four categories of antecedents to ESS and purchase by integrating elements of search cost, information, two-stage, context effects, and consumer behavior theories. We use MRT's concepts of relative brand uncertainty, individual brand uncertainty, involvement, and search cost to integrate the relevant economic and behavioral search theories and study the joint determination of ESS and purchase. Our methodology rests on formal model comparisons and the examination of parameter estimates to determine the empirical relevance of alternative sets of covariates and modeling assumptions. Our baseline model addresses key features of the data by combining a profligate model for the number of evaluated products with a parsimonious model of purchase in a joint copula model.

An interesting result we find is that smaller ESS corresponds to higher purchase likelihood. This result is consistent with observing customers that are at different stages of the purchase cycle, with those that are earlier in their search (or novices) constructing a larger ESS and being less likely to buy as compared to those that are later in their search (or experts) who construct a smaller ESS and are more likely to buy. Further, we find that some elements of search behavior (use of sorting tools and larger size of purchase) are associated with larger ESS, while other types of search behavior (the use of filtering tools) is associated with a smaller ESS --- and therefore higher probability of purchase.

Managerially, we conclude that consumers' search behavior on the website is predictive of ESS and purchase propensity, which can be fruitfully leveraged by the website. A well-designed search engine should provide both sorting and filtering capabilities, to cater to both early-stage and late-stage customers. However, since filtering is consistent with small ESS and correspondingly high purchase likelihood, it is prudent for a product search engine to design their webpage to encourage filtering and thus help customers finalize their purchase decisions. Moreover, rich filtering capabilities may also attract more late-stage customers who are looking for a quick way to complete a purchase.

The key objective of the website from a marketing perspective would be to accurately predict where the user is in her purchase funnel, and deploy tactics to move her down the funnel --- towards purchase. For example, the website could provide a dashboard that prominently displays the current evaluation set (i.e., the hotels clicked on by the consumer), including comparative tools that make it easy for the consumer to compare and contrast the evaluation set options along the dimensions that are important to her. The dashboard could also incorporate an "editor" that allows the consumer to easily add, substract or reorder options in the evaluation set. The editor could also determine and display display "dominated alternatives" --- i.e., options that are dominated by others in terms of price, quality and proximity to desired neighborhood --- so the consumer can readily winnow the evaluation set, and move closer to identifying the best option to purchase. Along the same lines, the website could dynamically allocate customer support resources (such as instant messaging or video chat) or dynamically offer marketing incentives (such as discounts) to assist customers who are close to finalizing their purchase.

As with any study, there are limitations to what we can infer. First, it is possible some searches result in purchases on the site subsequent to the time window we studied. Nevertheless, insofar as the results demonstrate support for our hypotheses in spite of this limitation, they can be viewed as conservative lower-bound estimates that would be strengthened had data on ultimate purchase been available. Second, it is possible some searches result in purchases at other travel sites during or after the time window for which we have data. While a similar argument can be made in this case as well, this limitation suggests the potential for future research based on consumer panel data collected over many travel sites and larger time windows, within the context of a joint model such as ours. Third, we note that our study employs observational data, rather than an experimental or quasi-experimental design. Therefore, even though our results and conclusions may be relevant for theory and practice, they should not be interpreted causally until causality has been verified. Fortunately, the setting of online product search engines allows for quick and inexpensive experimentation by management through which causality or lack thereof can be ascertained before our recommendations are put to practice. We suspect that each market and market segment may present different challenges, yet our basic framework should be useful in forming, evaluating, and estimating treatment models and effects. We view the model and methodology presented in this paper as an important step in the analysis of data from product search engines, providing heretofore unavailable insights about the antecedents of ESS and purchase, enhancing our understanding of consumer behavior at product search engines, and laying the groundwork for future academic and commercial research.

References

Alba, Joseph W., John Lynch, Barton Weitz, Chris Janiszewski, Richard Lutz, Alan Sawyer, and Stacy Wood (1997), "Interactive Home Shopping: Consumer, Retailer, and Manufacturer Incentives to Participate in Electronic Marketplaces," *Journal of Marketing*, 61, 3, 38–53.

Alba, Joseph W. and J. Wesley Hutchinson (1987), "Dimensions of Consumer Expertise," *Journal of Consumer Research*, 13, 4, 411–54.

Andrews, Rick L. and T.C. Srinivasan (1995), "Studying Consideration Effects in Empirical Choice Models Using Scanner Panel Data," *Journal of Marketing Research*, 32, 1, 30–41.

Bakos, J. Yannis (1997), "Reducing Buyer Search Costs: Implications for Electronic Marketplaces," *Management Science*, 43, 12, 1676–92.

Beatty, Sharon E. and Scott M. Smith (1987), "External Search Effort: An Investigation Across Several Product Categories," *Journal of Consumer Research*, 14, 1, 83–95.

Cameron, A. Colin and Pravin K. Trivedi (1998), *Regression Analysis of Count Data*, Cambridge, UK: Cambridge University Press.

Chen, Yuxin and Song Yao (2015), "Sequential Search with Refinement: Model and Application with Click-Stream Data," working paper, Northwestern University, http://papers.ssrn.com /sol3/papers.cfm?abstract_id=2130646.

Chib, Siddhartha and Edward Greenberg (1995), "Hierarchical Analysis of SUR Models with Extensions to Correlated Serial Errors and Time-Varying Parameter Models," *Journal of Econometrics*, 68, 2, 339–60.

--- and Ivan Jeliazkov (2005), "Accept–Reject Metropolis–Hastings Sampling and Marginal Likelihood Estimation," *Statistica Neerlandica*, 59, 1, 30–44.

--- and --- (2006), "Inference in Semiparametric Dynamic Models for Binary Longitudinal Data," *Journal of the American Statistical Association*, 101, 474, 685–700.

Danaher, Peter J. and Michael S. Smith (2011), "Modeling Multivariate Distributions Using Copulas: Applications in Marketing," *Marketing Science*, 30, 1, 4–21.

De los Santos, Babur, Ali Hortaçsu, and Matthijs R. Wildenbeest (2012), "Testing Models of Consumer Search Using Data on Web Browsing and Purchasing Behavior," *American Economic Review*, 102, 6, 2955–80.

Diehl, Kristin (2005), "When Two Rights Make a Wrong: Searching Too Much in Ordered Environments," *Journal of Marketing Research*, 42, 3, 313–22.

--- and Gal Zauberman (2005), "Searching Ordered Sets: Evaluations from Sequences under Search," *Journal of Consumer Research*, 31, 4, 824–32.

Ghose, Anindya, Panagiotis G. Ipeirotis, and Beibei Li (2012), "Designing Ranking Systems for Hotels on Travel Search Engines by Mining User-Generated and Crowdsourced Content," *Marketing Science*, 31, 3, 493–520.

---, ---, and --- (2014), "Examining the Impact of Ranking on Consumer Behavior and Search Engine Revenue," *Management Science*, 60, 7, 1632–54.

Goodman, Joseph K., Susan M. Broniarczyk, Jill G. Griffin, and Leigh McAlister (2013), "Help or Hinder? When Recommendation Signage Expands Consideration Sets and Heightens Decision Difficulty," *Journal of Consumer Psychology*, 23, 2, 165–74.

Hair, Joseph F., William C. Black, Barry J. Babin, and Rolph E. Anderson (2010), *Multivariate Data Analysis*, 7th edition, Upper Saddle River, NJ: Prentice Hall.

Häubl, Gerald and Valerie Trifts (2000), "Consumer Decision Making in Online Shopping Environments: The Effects of Interactive Decision Aids," *Marketing Science*, 19, 1, 4–21.

Hauser, John H. and Birger Wernerfelt (1990), "An Evaluation Cost Model of Consideration Sets," *Journal of Consumer Research*, 16, 4, 393–408.

Honka, Elisabeth (2014), "Quantifying Search and Switching Costs in the US Auto Insurance Industry," *The RAND Journal of Economics*, 45, 4, 847–84.

Howard, John A. and Jagdish N. Sheth (1969), *Theory of Buyer Behavior*, New York: John Wiley & Sons.

Iyengar, Sheena S. and Mark R. Lepper (2000), "When Choice Is Demotivating: Can One Desire Too Much of a Good Thing?," *Journal of Personality & Social Psychology*, 79, 6, 995–1006.

Jacoby, Jacob (1984), "Perspectives on Information Overload," *Journal of Consumer Research*, 10, 4, 432–35.

---, Donald E. Speller, and Carol A. Kohn (1974), "Brand Choice Behavior as a Function of Information Load," *Journal of Marketing Research*, 11, 1, 63–69.

Jeliazkov, Ivan and Esther H. Lee (2010), "MCMC Perspectives on Simulated Likelihood Estimation," *Advances in Econometrics: Maximum Simulated Likelihood*, 26, 3–39.

--- and Alicia Lloro (2011), "Maximum Simulated Likelihood Estimation: Techniques and Applications in Economics," in *Computational Optimization, Methods and Algorithms*, Slawomir Koziel and Xin-She Yang, eds., Springer Berlin Heidelberg, 85–100.

Joe, Harry (1997), *Multivariate Models and Multivariate Dependence Concepts*, London: Chapman & Hall.

Keller, Kevin Lane and Richard Staelin (1987), "Effects of Quality and Quantity of Information on Decision Effectiveness," *Journal of Consumer Research*, 14, 2, 200–213.

Koulayev, Sergei (2014), “Search for Differentiated Products: Identification and Estimation,” *The RAND Journal of Economics*, 45, 3, 553–75.

Kuksov, Dmitri and J. Miguel Villas-Boas (2010), “When More Alternatives Lead to Less Choice,” *Marketing Science*, 29, 3, 507–24.

Lurie, Nicholas H. (2004), “Decision Making in Information-Rich Environments: The Role of Information Structure,” *Journal of Consumer Research*, 30, 4, 473–86.

Malhotra, Naresh K. (1982), “Information Load and Consumer Decision Making,” *Journal of Consumer Research*, 8, 4, 419–30.

---, Arun K. Jain, and Stephen W. Lagakos (1982), “The Information Overload Controversy: An Alternative Viewpoint,” *Journal of Marketing*, 46, 2, 27–37.

Mehta, Nitin, Surendra Rajiv, and Kannan Srinivasan (2003), “Price Uncertainty and Consumer Search: A Structural Model of Consideration Set Formation,” *Marketing Science*, 22, 1, 58–84.

Moe, Wendy W. (2006), “An Empirical Two-Stage Choice Model with Varying Decision Rules Applied to Internet Clickstream Data,” *Journal of Marketing Research*, 43, 4, 680–92.

Moorthy, Sridhar, Brian T. Ratchford, and Debabrata Talukdar (1997), “Consumer Information Search Revisited: Theory and Empirical Analysis,” *Journal of Consumer Research*, 23, 4, 263–77.

Moraga-González, José Luis, Zsolt Sándor, and Matthijs R. Wildenbeest (2013), “Semi-Nonparametric Estimation of Consumer Search Costs,” *Journal of Applied Econometrics*, 28, 7, 1205–23.

Nelsen, Roger B. (2007), *An Introduction to Copulas*, . 2nd edition, New York: Springer.

Payne, John W. (1976), “Task Complexity and Contingent Processing in Decision Making: An Information Search and Protocol Analysis,” *Organizational Behavior & Human Performance*, 16, 2, 366–87.

---, James R. Bettman, and Eric J. Johnson (1993), *The Adaptive Decision Maker*, Cambridge, UK: Cambridge University Press.

Pitt, Michael, David Chan, and Robert Kohn (2006), “Efficient Bayesian Inference for Gaussian Copula Regression Models,” *Biometrika*, 93, 3, 537–54.

Roberts, John H. and James M. Lattin (1991), “Development and Testing of a Model of Consideration Set Composition,” *Journal of Marketing Research*, 28, 4, 429–40.

Scheibehenne, Benjamin, Rainer Greifeneder, and Peter M. Todd (2010), “Can There Ever Be Too Many Options? A Meta-Analytic Review of Choice Overload,” *Journal of Consumer Research*, 37, 3, 409–25.

Schwartz, Barry (2004), *The Paradox of Choice*, New York: HarperCollins.

Shi, Savannah Wei and Jie Zhang (2014), "Usage Experience with Decision Aids and Evolution of Online Purchase Behavior," *Marketing Science*, 33, 6, 871–82.

Siddarth, S., Randolph E. Bucklin, and Donald G. Morrison (1995), "Making the Cut: Modeling and Analyzing Choice Set Restriction in Scanner Panel Data," *Journal of Marketing Research*, 32, 3, 255–66.

Sklar, A (1959), "Fonctions de Répartition À N Dimensions et Leurs Marges," *Publications de l'Institut de Statistique de L'Université de Paris*, 8, 229–31.

Stigler, George J. (1961), "The Economics of Information," *Journal of Political Economy*, 69, 3, 213–25.

Stiglitz, Joseph E. (1989), "Imperfect Information in the Product Market," in *Handbook of Industrial Organization*, Richard Schmalensee and Robert Willig, ed., Elsevier, Volume 1:769–847.

Tierney, Luke (1994), "Markov Chains for Exploring Posterior Distributions," *The Annals of Statistics*, 22, 4, 1701–28.

Townsend, Claudia and Barbara E. Kahn (2014), "The 'Visual Preference Heuristic': The Influence of Visual versus Verbal Depiction on Assortment Processing, Perceived Variety, and Choice Overload," *Journal of Consumer Research*, 40, 5, 993–1015.

Trivedi, Pravin K. and David M. Zimmer (2005), "Copula Modeling: An Introduction to Practitioners," *Foundations and Trends in Econometrics*, 1, 1–111.

van Nierop, Erjen, Bart Bronnenberg, Richard Paap, Michel Wedel, and Philip Hans Franses (2010), "Retrieving Unobserved Consideration Sets from Household Panel Data," *Journal of Marketing Research*, 47, 1, 63–74.

Weitzman, Martin L. (1979), "Optimal Search for the Best Alternative," *Econometrica*, 47, 3, 641–54.

Winkelmann, Rainer (2008), *Econometric Analysis of Count Data*, 5th edition, Berlin: Springer.

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

Zhang, Jie, Xiao Fang, and Olivia R. Liu Sheng (2007), "Online Consumer Search Depth: Theories and New Findings," *Journal of Management Information Systems*, 23, 3, 71–95.

Appendix

The Accept-Reject Metropolis-Hastings (ARMH) Algorithm. The ARMH algorithm is a useful modification of the traditional accept-reject simulation method. In particular, let the posterior density $\pi(\theta \mid y) \propto f(y \mid \theta)\pi(\theta)$, which need only be known up to proportionality, be the target density of interest and let $h(\theta \mid y)$ denote a source density from which one can easily draw random variates. Also, let c denote a constant that defines the region of integration

$$D = \{\theta : f(y \mid \theta)\pi(\theta) \leq ch(\theta \mid y)\},$$

and let D^c denote the complement of this set. Then, the ARMH algorithm is defined by the following two steps:

1. Accept-reject step: Generate a proposed draw $\theta' \sim h(\theta \mid y)$ and accept it with probability

$$\alpha_{AR}(\theta' \mid y) = \min\left\{1, \frac{f(y \mid \theta')\pi(\theta')}{ch(\theta' \mid y)}\right\},$$

continuing the process until a draw has been accepted.

2. Metropolis-Hastings step: Given the current value θ and the proposed θ':

a) if $\theta \in D$, let $\alpha_{MH}(\theta, \theta' \mid y) = 1$;

b) if $\theta \in D^c$ and $\theta' \in D$, let $\alpha_{MH}(\theta, \theta' \mid y) = \dfrac{ch(\theta \mid y)}{f(y \mid \theta)\pi(\theta)}$;

c) if $\theta \in D^c$ and $\theta' \in D^c$, let $\alpha_{MH}(\theta, \theta' \mid y) = \min\left\{1, \dfrac{f(y \mid \theta')\pi(\theta')h(\theta \mid y)}{f(y \mid \theta)\pi(\theta)h(\theta' \mid y)}\right\}$.

With probability $\alpha_{MH}(\theta, \theta' \mid y)$, return θ' and otherwise return θ.

This Markov chain Monte Carlo (MCMC) algorithm is ergodic and converges to the density $\pi(\theta \mid y)$ as the sampling is iterated. Moreover, as shown in Chib and Jeliazkov (2005) its building blocks can easily be employed to produce an estimate of the marginal likelihood through the ratio

$$m(y)=\frac{c\int \alpha_{AR}(\theta \mid y)h(\theta \mid y)d\theta}{\int \alpha_{MH}(\theta,\theta' \mid y)\pi(\theta \mid y)d\theta},$$

where we have suppressed any model indicators in the conditioning set. These model indicators will be introduced next, as we discuss the marginal likelihoods and model comparison issues.

Model comparison. Given two competing models M_i and M_j with corresponding parameter vectors θ_i and θ_j, the posterior odds ratio, i.e., the ratio of posterior model probabilities given the data y, is given by

$$\frac{\Pr(M_i \mid y)}{\Pr(M_j \mid y)}=\frac{m(y \mid M_i)}{m(y \mid M_j)}\frac{\Pr(M_i)}{\Pr(M_j)}.$$

This expression follows from Bayes' formula and shows that to obtain the posterior odds, the prior odds (ratio of model probabilities before seeing the data) are updated by the ratio of marginal likelihoods (called the Bayes factor). The marginal likelihood is given by

$$m(y \mid \theta_k, M_k)=\int f(y \mid \theta_k, M_k)\pi(\theta_k \mid M_k)d\theta_k$$

and represents an important ingredient in comparing competing models.

Well-known properties of marginal likelihoods and Bayes factors are that they lead to finite sample model probabilities, do not require competing models to be nested, and have appealing asymptotic properties that give rise to the information criterion of Schwarz (1978) (see O'Hagan 1994, Ch. 3; Greenberg 2008, Ch. 3). A less-known, yet very important, point is that marginal likelihoods provide a measure of *sequential out-of-sample predictive fit*. This can be seen by writing:

$$
\begin{aligned}
f\left(y \mid M_l\right) &= \prod_{i=1}^{n} f\left(y_i \mid \left\{y_j\right\}_{j<i}, M_l\right) = \\
&= \prod_{i=1}^{n} \int f\left(y_i \mid \left\{y_j\right\}_{j<i}, \theta_l, M_l\right) \pi\left(\theta_l \mid \left\{y_j\right\}_{j<i}, M_l\right) d\theta_l,
\end{aligned}
$$

where the first line uses the law of total probability to represent the marginal likelihood as the product of n one-step-ahead sequential predictive densities and the second line makes it explicit that the adequacy of a model, as captured by its marginal likelihood, corresponds to its cumulative out-of-sample predictive record where the fit of observation i is measured with respect to the posterior density based only on data $\{y_j\}_{j<i}$ up to the ith data point, without conditioning on $\{y_j\}_{j\geq i}$. In contrast, in-sample measures of fit condition on the entire data set. There are also important advantages of the model comparison framework presented here relative to customary out-of-sample comparisons in which a researcher would estimate the model using part of the data and then examine how successfully that model can predict the remainder of the data. The sequential out-of-sample fit measure provided by marginal likelihoods overcomes key difficulties of traditional out-of-sample comparisons. In particular, note that the marginal likelihood is invariant to permutation of the indices of the data: the same value will be obtained if the data were rearranged. This invariance is desirable because, in contrast, typical out-of-sample comparisons depend on what part of the data is used in estimation and what is retained for the purpose of comparison.

We use the approach of Chib and Jeliazkov (2005) to estimate the marginal likelihood using the building blocks of the ARMH algorithm, but MCMC sampling can also be employed for evaluating the likelihood function of our model when it involves more outcomes and higher-dimensional integrals using the methods discussed in Jeliazkov and Lee (2010).

Figure 1: Conceptual Framework

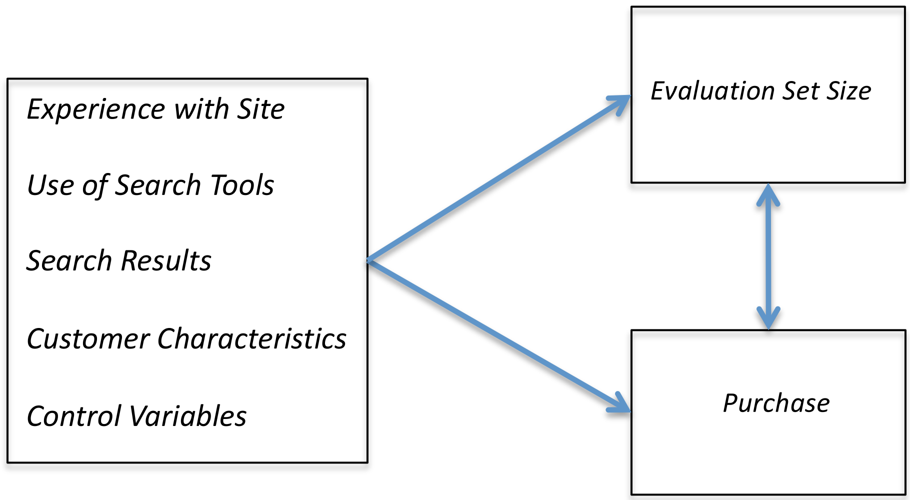

Table 1: Descriptive Statistics and Definitions of Variables

Component 1 - Evaluation Set Size

Variable	Definition	Mean	St. Dev.
Dependent Variable			
Evaluation Set Size	The number of hotels in the search results that this customer clicked on to reveal the details about, and thus evaluate.	0.75	1.30
Independent Variables			
Experience with Site			
Number of Hotel Searches	Frequency of previous hotel searches made by this customer on this travel website during the 3 months prior to the customer's current search.*	27.31	88.17
Number of Non-Hotel Searches	Frequency of previous non-hotel searches made by this customer on this travel website during the 3 months prior to the customer's current search.*	51.43	166.02
Number of Hotel Purchases	Frequency of previous hotel purchases made by this customer on this travel website during the 3 months prior to the customer's current search.*	0.12	0.89
Number of Non-Hotel Purchases	Frequency of previous non-hotel purchases made by this customer on this travel website during the 3 months prior to the customer's current search.*	0.26	1.45
Previous Clicks	Frequency of previous clicking activity by this customer on this travel website that may be related to the customer's current search, as measured by the number of hotels this customer clicked on which were (a) part of our study (related to searches made Oct 1 – Oct 14, 2009 for hotels in Budapest), and (b) clicked on prior to the current search	0.51	1.57
Use of Search Tools			
Filter	Binary indicator set to 1 if the customer filtered their results using a textual keyword or phrase (e.g., hotel name, brand, point of interest), or 0 otherwise	0.17	0.38
Sort	Binary indicator set to 1 if customer sorted their results (e.g., by hotel name, distance to point of interest, site rating, user rating, or price), or 0 otherwise	0.12	0.33
Search Engine	Binary indicator set to 1 if this customer used a search engine and followed a link generated by that search engine to get to our travel site, or 0 otherwise	0.26	0.44
Partner Site	Binary indicator set to 1 if the customer used a partner site (i.e., other travel website that shares data with the travel website in our study) and followed a link from that partner site to get to our travel website, or 0 otherwise	0.61	0.49
Search Results			
Total Volume	Number of hotels matching a customer's search criteria	181.49	112.22
Volume On Pages Browsed	Number of hotels matching a customer's search criteria that were presented on-screen. (Note: Unless the customer browses through all pages of search results, the number of hotels seen will be less than the number of matched results).	21.79	9.20
Customer Characteristics			
Nights Stay	Number of nights the customer wants to book, as measured by the planned hotel departure date minus the planned arrival date plus one	3.23	1.89
Room Count	Number of rooms the customer wants to book	1.14	0.55
Adult Count	Number of adults in the customer's party	1.83	0.60
Child Count	Number of children in the customer's party	0.06	0.30
Days Until Check-In	Immediacy of this customer's travel, as measured by the number of days between their search date and their planned hotel arrival date	52.21	63.97
Controls			
Price	Average price of the hotels in this customer's search results	105.68	33.30
Site Rating	Average star rating of the hotels in this customer's search results	40.27**	5.42
User Rating	Average user review rating of the hotels in this customer's search results	4.26	0.25
User Rating Count	Average user review count of the hotels in this customer's search results	21.13	20.80
Distance to Best Areas	Proximity of the hotels in this customer's search results to popular areas, as measured by the average distance (in meters) to the nearest desirable location.	1396.64	819.09
Branded Hotel	Proportion of brand-name hotels in this customer's search results	0.44	0.24

** Past experience applies for any worldwide destination. Non-hotel searches and purchases could be for flights, rental cars, or other travel products that the travel website offers. ** The search engine multiplies this star rating by 10.*

Component 2 - Purchase

Variable	Definition	Mean	St. Dev.
Dependent Variable			
Purchase	Binary indicator set to 1 if this customer purchased a hotel, or 0 otherwise	0.09	0.29
Independent Variables (Controls)			
Price	Average price of the hotels in this customer's evaluation set	107.48	51.32
Site Rating	Average star rating of the hotels in this customer's evaluation set	42.18	6.99
User Rating	Average user review rating of the hotels in this customer's evaluation set	4.31	0.44
User Rating Count	Average user review count of the hotels in this customer's evaluation set	25.10	30.90
Distance to Best Areas	Proximity of the hotels in this customer's evaluation set to popular areas, as measured by the average distance (in meters) to the nearest desirable location.	713.29	1261.3
Branded Hotel	Proportion of brand-name hotels in this customer's evaluation set	0.53	0.46

Table 2: Estimation Results

	M1		M2		M3		M4		M5	
	Mean	SD	Mean	SD	Mean	SD	Mean	SD	Mean	SD
Component 1 - Evaluation Set Size										
Intercept	-0.94	1.02	-0.91	1.01	-0.19	1.10	0.47	0.65	-0.24	1.04
Experience with Site										
Number of Hotel Searches*	**0.03**	**0.02**	0.03	0.02	**0.03**	**0.02**	**0.03**	**0.02**	**0.03**	**0.02**
Number of Non-Hotel Searches*	---	---	0.00	0.02	---	---	---	---	---	---
Number of Hotel Purchases*	---	---	-0.10	0.13	---	---	---	---	---	---
Number of Non-Hotel Purchases*	---	---	0.12	0.10	---	---	---	---	---	---
Previous Clicks**	0.11	0.06	0.11	0.06	0.11	0.06	**0.11**	**0.06**	**0.11**	**0.06**
Use of Search Tools										
Filter (binary)	**-0.62**	**0.16**	**-0.63**	**0.16**	**-0.68**	**0.16**	**-0.70**	**0.16**	**-0.66**	**0.15**
Sort (binary)	**0.34**	**0.09**	**0.34**	**0.09**	**0.33**	**0.09**	**0.35**	**0.09**	**0.33**	**0.09**
Search Engine (binary)	**0.32**	**0.08**	**0.33**	**0.08**	**0.32**	**0.08**	**0.32**	**0.09**	**0.32**	**0.08**
Partner Site (binary)	**0.28**	**0.08**	**0.28**	**0.08**	**0.27**	**0.08**	**0.28**	**0.08**	**0.28**	**0.08**
Search Results										
Total Volume	**-0.24**	**0.04**	**-0.24**	**0.05**	**-0.26**	**0.04**	**-0.25**	**0.04**	**-0.26**	**0.04**
Volume on Pages Browsed	**0.22**	**0.11**	**0.23**	**0.11**	**0.24**	**0.11**	**0.22**	**0.11**	**0.23**	**0.10**
Customer Characteristics										
Nights Stay	0.04	0.08	0.04	0.08	0.05	0.08	0.05	0.08	0.05	0.08
Room Count	**0.43**	**0.16**	**0.42**	**0.16**	**0.43**	**0.15**	**0.41**	**0.16**	**0.43**	**0.15**
Adult Count	**0.30**	**0.15**	**0.30**	**0.15**	0.27	0.14	**0.30**	**0.14**	0.27	0.15
Child Count	**0.53**	**0.14**	**0.54**	**0.14**	**0.49**	**0.15**	**0.52**	**0.14**	**0.48**	**0.14**
Days Away from Check-In	-0.03	0.02	-0.03	0.02	-0.02	0.02	-0.02	0.02	-0.03	0.02
Controls										
Price***	**-0.35**	**0.14**	**-0.34**	**0.14**	-0.23	0.12	**-0.28**	**0.11**	**-0.23**	**0.11**
Site Rating***	0.47	0.33	0.46	0.33	---	---	---	---	---	---
User Rating***	---	---	---	---	0.40	0.72	---	---	0.42	0.69
User Rating Count***	---	---	---	---	-0.04	0.06	---	---	-0.04	0.06
Distance to Best Areas***	---	---	0.00	0.03	0.00	0.03	0.00	0.03	---	---
Branded Hotel (binary)***	---	---	---	---	---	---	0.22	0.15	---	---
Component 2 - Purchase										
Intercept	0.93	1.01	1.20	1.04	-1.06	0.99	-0.74	0.51	-1.22	0.97
Controls										
Price****	0.26	0.15	0.21	0.15	-0.10	0.12	-0.01	0.11	-0.05	0.12
Site Rating****	**-0.83**	**0.37**	**-0.80**	**0.37**	---	---	---	---	---	---
User Rating****	---	---	---	---	0.26	0.68	---	---	0.12	0.67
User Rating Count****	---	---	---	---	0.08	0.05	---	---	0.09	0.05
Distance to Best Areas****	---	---	**-0.04**	**0.02**	-0.04	0.02	**-0.04**	**0.02**	---	---
Branded Hotel (binary)****	---	---	---	---	---	---	-0.09	0.11	---	---
Model Parameters										
Copula Correlation ρ: Relationship b/n ESS & Purchase	**-0.35**	**0.09**	**-0.31**	**0.10**	**-0.27**	**0.10**	**-0.27**	**0.10**	**-0.32**	**0.09**
ln (α): Negative Binomial Dispersion Parameter	-0.14	0.07	-0.14	0.07	-0.14	0.07	-0.14	0.07	-0.14	0.07
Model Fit Statistics										
Log Marginal Likelihood	-4186.90		-4208.19		-4205.20		-4199.92		-4196.90	
Numerical Standard Error	0.017		0.034		0.021		0.020		0.020	

** From the three months immediately preceding the focal search*
*** Prior to the focal search but during the two-week study period of Oct 1-Oct 14, 2009*
**** These quality metrics are averaged over all hotels in the search result set*
***** These quality metrics are averaged over all hotels in the evaluation set*
Bold parameter estimates indicate that the 95% highest posterior density region did not include 0.

Table 3: Additional Model Comparisons for Versions of M1

Model	**Log Marginal Likelihood (NSE)**
Copula (M1, $\rho \neq 0$)	-4186.90 (0.017)
Independent Equations (M1, $\rho = 0$)	-4198.44 (0.05)
Conditional Model (Copula, $\rho \neq 0$)	-4189.50 (0.097)
Conditional Model (Indep. Eqs, $\rho = 0$)	-4196.44 (0.05)
Copula Model with Fitted Regressors ($\rho \neq 0$)	-4186.14 (3.137)

Web Appendix Table 1. Correlation Matrices

Panel A: Equation 1

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)	(11)	(12)	(13)	(14)	(15)	(16)	(17)	(18)	(19)	(20)	(21)	(22)	(23)
(1) Hotel Search Frequency	1																						
(2) Non-Hotel Search Frequency	.51	1																					
(3) Hotel Purchases Frequency	.37	.41	1																				
(4) Non-Hotel Purchases Frequency	.33	.49	.65	1																			
(5) Previous Clicks	.20	.06	.00	.01	1																		
(6) Filtered Search	.02	.01	.04	.02	.05	1																	
(7) Sorted Search	.03	.00	.00	.00	.05	-.10	1																
(8) Search Engine Visit	-.04	.00	-.03	-.04	.03	-.01	.00	1															
(9) Partner Site Visit	.08	.07	.05	.07	.03	.06	.03	-.59	1														
(10) Vol. of Search Results, Total	-.03	-.02	-.02	-.03	-.07	-.51	.09	.04	-.10	1													
(11) Vol. of Search Results, Pages Browsed	-.02	-.01	-.03	-.03	-.05	-.72	.14	.02	-.08	.69	1												
(12) Nights Stay	.01	.00	-.02	-.03	.00	-.04	.00	.04	-.04	.07	.04	1											
(13) Room Count	-.03	-.03	-.02	-.02	.00	-.01	.04	.04	-.02	.01	.02	-.03	1										
(14) Adult Count	.00	-.04	-.03	-.05	.03	-.06	.02	.09	-.08	.06	.05	-.04	.00	1									
(15) Child Count	-.01	-.01	.01	.02	.01	-.05	.02	.03	-.04	.04	.05	.05	.01	.00	1								
(16) Days Away	.03	.02	-.04	-.04	-.04	-.03	.04	.06	-.01	.05	.03	.14	.12	.09	.04	1							
(17) Ave. Price of Results	.02	.00	-.01	.02	-.01	.03	-.32	-.02	.02	-.25	-.21	-.10	-.02	.11	.08	.09	1						
(18) Ave. Site Rating of Results	.02	.00	.00	.02	.03	.06	-.33	-.01	.02	-.45	-.30	-.02	.01	-.06	-.08	.00	.66	1					
(19) Ave. User Rating of Results	-.01	-.01	-.04	-.02	.02	.02	-.33	.04	-.04	-.14	-.12	.03	-.05	.01	-.01	.04	.46	.54	1				
(20) Ave. User Rating Count of Results	-.01	-.03	-.03	-.02	-.01	.02	-.11	.00	-.01	-.01	.01	-.01	.00	-.05	-.03	.04	.11	.13	.25	1			
(21) Prop. of Brand Hotels in Results	.03	.01	.05	.05	.00	.12	-.30	.00	-.01	-.26	-.24	-.01	.04	-.08	-.08	-.01	.44	.63	.30	.04	1		
(22) Results Ave. Proximity to Desirable Location	.00	.00	-.01	-.01	.00	-.30	.29	-.01	-.02	.42	.45	-.02	-.03	.03	.04	-.03	-.42	-.48	-.34	-.07	-.45	1	
(23) *Evaluation Set Size*	.04	-.01	.00	.00	.05	-.07	.10	.04	.03	-.06	.03	.00	.05	.03	.06	.02	-.01	.02	-.02	-.04	-.01	.02	1

Panel B: Purchase Equation

	(1)	(2)	(3)	(4)	(5)	(6)	(7)
(1) Average Price of Alternatives Evaluated	1						
(2) Average Site Rating of Alternatives Evaluated	.70	1					
(3) Average User Rating of Alternatives Evaluated	.44	.49	1				
(4) Average User Rating Count of Alternatives Evaluated	.22	.21	.31	1			
(5) Proportion of Branded Hotels of Alternatives Evaluated	.18	.38	-.03	-.01	1		
(6) Results Average Proximity to Desirable Location	-.24	-.28	-.16	-.12	-.28	1	
(7) *Did Buy Hotel*	.01	-.02	.03	.04	-.01	-.06	1