A unique aspect of sponsored search advertising is that it allows firms to track what products consumers initially search for and what products they eventually buy after visiting their website – something that is typically hard to acquire in physical world settings. Based on a unique dataset that contains information on what keyword advertisements induced consumers to arrive on a firm’s website and what products they eventually bought in that session, we build a model to map consumers’ search-to-purchase behavior in the online world. We analyze the relationship between consumers’ search for products in a specific category and their propensity to buy products within that category. In addition, we examine spillovers from search resulting in purchases across other product categories in the same session. The model is estimated on purchases from four categories (bath, bedding, kitchen, and home decor) using a unique 6 month panel dataset of several hundred keywords collected from a large nationwide retailer that advertises on Google. Our model allows us to decompose the latent utility of a product category into intrinsic utility from that product category and extrinsic utility from joint purchases of that category with other ones. With regard to the intrinsic utilities, we find that there is a considerable amount of spillovers between the initial search and the final purchase behavior such that consumers who start a search for a product in one category purchase products from a different category, in addition to purchasing from the original category they searched for. However, such search-purchase spillover effects accruing from a given advertisement are not necessarily symmetric between any two given product categories. With regard to the extrinsic utilities, we see evidence of positive cross-category interdependence for retailer-specific keywords. In addition, we find that brand-specific and generic keywords are less likely to induce cross-category purchases. Based on our empirical estimates, we conduct counterfactual experiments to derive insights into changes in overall profitability when firms were to engage in personalized price discounts and targeting via sponsored search advertisements. Our analyses suggest that retailers can benefit from providing personalized targeting and recommendations based on consumer search and cross-category purchase patterns during sponsored search advertising.

Keywords: Online advertising, Search engine marketing, Paid search, Conversion rates, Electronic commerce, Cross-selling, Hierarchical Bayesian estimation.

1. Introduction

The Internet has brought about a fundamental change in the way consumers obtain information, and the way firms advertise products, thereby facilitating a paradigm shift in consumer search and purchase patterns. Though there are many innovative ways firms can advertise online, the bulk of online advertising consists of two main forms: display ad (banner) advertising and paid search advertising (sponsored ads that appear on the search results pages of search engines). The global paid search advertising market is predicted to
have a 37% compound annual growth rate, to more than $33 billion in 2010 and has become a critical component of firm’s marketing campaigns.

Consumers make multi-category decisions in a variety of contexts such as grocery shopping trips, mail-order purchasing, and more recently through online shopping. In the online retailing context, these multi-category decisions result in the formation of consumers' shopping baskets which comprise the collection of categories that consumers purchase in a specific shopping session. Hence, retailers advertising online are interested in understanding how consumer search related factors influence consumer purchase decisions and possible spillovers in terms of the cross-category purchases. This is because if a single click on a given sponsored ad can lead to purchases of different products within a given category or purchases across multiple, then the advertiser is essentially successful in garnering higher revenues at the same cost.

In addition, online advertisers are also interested in implementing micro-marketing programs through the use of individual keyword level data. A combination of cross-category insights and individual keyword-level preferences could allow firms to make advertisement investment decisions across brands in related categories. For example, manufacturers such as “Nate Berkus” or “Nautica”, who in fact market the same brand across different categories, could utilize these insights to optimize marketing expenditures across two or more related categories, e.g., bath and bedding categories. Furthermore, many firms are interested in discovering cross-category dependencies and designing targeted cross-selling programs based on them.

Despite the growth of search advertising, we have little understanding of how sponsored search advertising influences consumer purchase patterns in a given shopping session on the Internet. In this paper, we focus on three main issues that are yet to be explored in prior work. First, given a conversion, how do search-related factors affect the categories of products that are eventually purchased? In particular, we are interested in determining the extent to which the initial category for which consumers search (which is also reflected in the ad they click) influences the final categories from which they purchase. We also aim to examine how keyword-related factors such as the presence of retailer or brand information in the keyword, the position of the advertisement on the search engine results page, and the extent of latency is associated with consumers’ purchase intention across purchases. Second, given a conversion through a sponsored ad, are consumer pur-
chase behaviors interdependent across product categories? If so, what keyword-level factors drive this interdependence and what are the different patterns of interdependence (positive or negative) between any two sets of product categories? Third, how do firm profits change with a change in either price discount or with a change in the cross-category interdependence induced by customized changes in the keyword advertisement content? While an emerging stream of empirical literature in sponsored search has looked at issues such as the effect of keyword rank on click-through and conversion rates (Rutz and Bucklin 2007, Ghose and Yang 2007, Agarwal et al. 2008), no prior work has empirically analyzed these questions.

Using a unique panel dataset of several hundred keywords collected from a large nationwide retailer that advertises on Google, we study the effect of consumer search on co-purchase interdependence across product categories. We propose a Hierarchical Bayesian modeling framework in which we model consumers’ cross category purchase behavior. Our data encompasses bedding, bath, home décor and kitchen items. Industry surveys have revealed that a vast majority of consumers (92%) use a search engine to help research and/or purchase appliances, kitchenware, furniture, and home improvement items (GMI 2005). Thus, our data is very representative of actual user behavior in the sponsored search context. In our data, 12.78% of the observations involved a joint purchase of products from two different categories in a session after a single keyword ad was clicked. Such evidence of cross category spillovers can have very useful profit implications since it suggests that some consumers click once on an ad but purchase multiple products once they visit the retailer’s website in the same session. Since retailers pay search engines on a per click basis for displaying keyword ads, such user behavior bodes very well for retailers. To the best of our knowledge, our paper is the first empirical study that documents the impact of search advertising on consumers’ cross-category purchase behavior in electronic markets. Our findings and contributions are as follows.

We present analysis with product and keyword-level variables to explore the extent of cross-category spillover opportunities across different categories from a given sponsored keyword advertisement. By examining cross-category purchases, we find that there exists potential for cross-category purchases or horizontal spillovers through paid search advertisements. A common industry practice to study cross-category purchasing behavior is to cross-tabulate joint purchases across multiple categories. However, this approach can-
not distinguish between consumer intrinsic preference for a product category and extrinsic preference from joint purchases of that one with other product categories. The choice of one category may affect the selection of another category due to the inter-dependence between the two categories. Alternatively, two categories may co-occur in a shopping basket not because of the extrinsic inter-dependencies but because of common intrinsic utility from these two categories. Our objective is to develop a general model of multi-category choice that explicitly allows for interdependence across categories, separating the latent utility of one product category into the intrinsic utility from that category and the extrinsic utility from the joint purchase of that category with other ones, after accounting for unobserved heterogeneity across keywords. To be clear, our goal is not to identify what drives this sort of interdependence in cross-category purchase behavior. Since our data is at the keyword level (and not at the consumer level), we are only able to provide evidence of the statistical relationship of cross-category purchase interdependence in the context of sponsored search advertising but are unable to infer what drives this relationship.

With respect to the intrinsic utility of purchase, we find that there is a considerable amount of spillovers between the initial search and the final purchase behavior across categories. Specifically, consumers who start a search for a product in one category eventually purchases products from a different category, in addition to purchases from the original category they searched for. However, such cross-category spillover effects accruing from a given keyword advertisement are not necessarily symmetric between any two given product categories. With respect to the extrinsic utility of purchase, we find evidence of positive cross-category interdependence for retailer-specific keywords. In addition, brand-specific and generic keywords are less likely to induce cross-category purchases compared to retailer keywords.

Overall, while prior stream of work in the offline context suggests that there appears to be an intrinsic propensity for any pair of product categories to co-occur within a household’s shopping basket, our results highlight that there could be both positive and negative extrinsic inter-dependence between category-level purchases in the context of online sponsored search advertising after controlling for intrinsic similarities among categories. Moreover, the presence of a brand name in the advertisement plays an important role in influencing the extent to which consumers spend on different product categories in a given shopping session.
Our research thus extends the existing literature by investigating consumers’ acquisition decisions for multiple products when exposed to advertising in the online context.

Based on our empirical estimates, we conduct counterfactual experiments to examine changes in profitability if firms were to engage in customized price discounts via sponsored search ads. We find that depending on the type of keyword clicked during the search process, a price discount on any given category or a combination of multiple categories leads to wide variation in profits. In general, price discounts applied to retailer–specific keywords increase profits while discounts applied to brand-specific or generic keywords decrease profits. Our analyses suggest that retailers can benefit from real-time targeting (for example, through coupons) based on consumer search patterns that provide information on their purchase intention.

The remainder of this paper is organized as follows. Section 2 gives an overview of the different streams of literature. In Section 3, we present a general model of multi-category choice that explicitly allows for interdependence across the four possible chosen categories, separates the effects of intrinsic utility of purchase from the extrinsic utilities, and accounts for unobserved heterogeneity across keywords. Section 4 describes the data and gives a brief background into some aspects of sponsored search advertising that are useful to understand before the empirical analyses. In Section 5, we discuss our estimation and conduct some counterfactual experiments. In Section 6, we discuss implications of our findings and conclude the paper.

2. Prior Literature

Our paper is related to two main streams of research – sponsored search advertising and multi-category purchase behavior. First, we discuss the work in online advertising. Much of the existing academic (e.g., Ilfeld and Winer 2002, Dreze and Hussherr 2003) on advertising in online world has focused on measuring changes in brand awareness, brand attitudes, and purchase intentions as a function of exposure, primarily through banner advertising. Manchanda et al. (2006) found that banner advertising actually increases purchasing behavior, in contrast to conventional wisdom. However, to the best of our knowledge none of these studies examined cross-category purchases through exposure to a single advertisement.
There is some empirical work in online sponsored search advertising. Existing research on search engines has analyzed search engine visits (Telang et. al. 2004) and the effectiveness of search engines (Bradlow and Schmittlein 2000) in information retrieval. Goldfarb and Tucker (2007) examine the factors that drive variation in prices for advertising legal services on Google. Animesh et al. (2008) look at how the creative content of ads affects click-through rates in search engines. More closely related to this paper is the emerging stream of research that models consumer click and purchase behavior (Rutz and Bucklin 2007, Ghose and Yang 2007, Ghose and Yang 2008, Rutz and Bucklin 2008, Agarwal et al. 2008) as well as advertiser bidding and search engine ranking behavior (Ghose and Yang 2007, 2008). Rutz and Bucklin (2007a) study hotel marketing keywords to analyze the profitability of different campaign management strategies.

Ghose and Yang (2007) quantify the impact of keyword type and length, position of the advertisement and the landing page quality on consumer click and conversion behavior as well as on advertiser’s cost per click and the search engine’s ranking decision for different ads. Ghose and Yang (2008) compare these performance metrics from organic search listings with that from paid search. Agarwal et al. (2008) quantify the profits associated with differences in keyword position. Rutz and Bucklin (2008) showed that there are spillovers between search advertising on branded and generic keywords, as some customers may start with a generic search to gather information, but later use a branded search to complete their transaction.

Second, our paper is also related to prior work in multi-category purchases in the offline world that has recognized that a consumer’s purchase decisions across categories are not independent. In other words, a consumer’s decision of when (or whether) to buy in one category depends on her corresponding decision in the related category (for example, Chintagunta and Halder 1998, Manchanda et al. 1999, Russell and Peterson 2000, Niraj et al. 2007). In the literature on multi-category purchases, it is well known that consumers make decisions in a variety of contexts such as choice of multiple categories during a shopping trip or mail-order purchasing (Manchanda et al. 1999). Two categories may co-occur in a shopping basket not because of the extrinsic interdependences but because of similar intrinsic utilities. Our paper is also related to the work that shows correlations in consumer’s quality perceptions of brands across two product categories due to a common brand name (Seetharaman et al. 2005). Finally, our paper is also related to work in cross-selling.
(for example, Knott et al. 2002, Li et al. 2005) that model consumers’ sequential acquisition decisions for multiple products. We thus aim to contribute to the literature by demonstrating the cross-selling potential of advertising in an online context, thereby supplementing the work on cross-selling in the offline world.

To summarize, our research is distinct from extant research in the following ways. *First*, we extend the online advertisement literature by modeling and empirically estimating how sponsored keyword advertisements can lead to opportunities for cross-category purchases in the online channel. *Second*, we build a model to map consumers’ search-purchase relationship in the online world. A unique aspect of sponsored search advertising is that it allows firms to track what products consumers initially search for and what products they eventually buy after visiting their website – something that is typically hard to acquire in physical world settings. Hence, we focus on purchase spillovers within a keyword but across different categories, i.e., horizontal spillovers. *Third*, much of the prior work in shopping basket analysis was based on multivariate probit models, and focused on quantifying purchase complementarities by examining correlations in the error terms. This implies that the cross-category interdependence has typically not been modeled as varying across units in prior work. Hence, an important difference between prior work and ours is that we allow the cross-category interdependence parameters to vary across keywords. Further, our autologistic model has a structural flavor because we derive the joint distribution of a purchase involving multiple product categories from the conditional distribution of purchases in one category given purchases in other categories. This is consistent with the notion of cross-category complementarity and/or substitutability. *Finally*, the majority of the work in this stream that has examined co-purchase incidence across categories is based on household level data of offline purchases that occurred through exposures to traditional marketing mix media. In contrast our data of co-purchase analysis is from sales in the online channel from exposure to sponsored search.

### 3. Model

---

2 Note that our data is based on conversions through search engine advertising. Any conversion by definition implies that at least one product was bought by the user. Hence, using the multivariate probit specification would be less appropriate in our setting since we do not observe instances when there were no purchases made in any category.
Our model is based on random utility theory and allows for the simultaneous and interdependent choice of multiple items. Let $Z_{hkj}$ denote the purchase decision in product category $j$ for order $k$ and keyword $h$. $Z_{hkj}$ is a dummy variable where 1 indicates a purchase and 0 indicates no purchase. In our empirical analysis, we have four product categories. Let $Z_h^k = (Z_{h1k}, Z_{h2k}, Z_{h3k}, Z_{h4k})$ represent the purchased bundle at order $k$ for keyword $h$. We start with modeling the purchase probability of product category $j$ from order $k$ for keyword $h$, conditional on the purchase decisions on all other categories $(j', j' \neq j)$. This takes a logit form. That is,

$$P(Z_{hkj} | Z_{hj'}, j \neq j') = \frac{\exp(V_{hkj})^{Z_{kj}}}{1 + \exp(V_{hkj})}, \quad j = 1, 2, 3, 4$$

(1)

where

$$V_{hkj} = \begin{bmatrix} V_{h1k} \\ V_{h2k} \\ V_{h3k} \\ V_{h4k} \end{bmatrix} = \begin{bmatrix} \pi_{hk1} \\ \pi_{hk2} \\ \pi_{hk3} \\ \pi_{hk4} \end{bmatrix} + \begin{bmatrix} 0 & \theta_1 & \theta_2 & \theta_3 \\ \theta_1 & 0 & \theta_2 & \theta_3 \\ \theta_1 & \theta_2 & 0 & \theta_3 \\ \theta_1 & \theta_2 & \theta_3 & 0 \end{bmatrix} \begin{bmatrix} Z_{h1k} \\ Z_{h2k} \\ Z_{h3k} \\ Z_{h4k} \end{bmatrix}$$

(2)

Equation (2) implies that the utility of a purchase from category $j$ is influenced by the intrinsic utility $\pi_{hj}$ and the extrinsic utility from purchases of other product categories $Z_{hj'} \,(j \neq j')$. $\theta_{hj'}$ measures the degree of interdependence between category $j$ and category $j'$. A positive $\theta_{hj'}$ suggests a complementary effect between the two categories. That is, the presence of category $j$ ($j'$) tends to increase the utility from category $j'$ ($j$). Similarly, a negative $\theta_{hj'}$ suggests a substitution effect between the two categories. That is, the presence of category $j$ ($j'$) tends to decrease the utility from category $j'$ ($j$). Finally, a zero effect of $\theta_{hj'}$ suggests independence between the two categories.

Based on Brook’s (1964) theory, we can show that the conditional probabilities (equations 1-2) lead to a proper joint distribution of the purchase incidences across the four product categories. Such an approach has been adopted in prior literature (see for example, Moon and Russell 2008). Hence,
\[
P(Z_{hk} = Z_{hk}^b) = \frac{\exp(\pi_{hk} Z_{hk}^b + \frac{1}{2} Z_{hk}^b W_h Z_{hk}^b)}{\sum_{b'} \exp(\pi_{hk} Z_{hk}^{b'} + \frac{1}{2} Z_{hk}^{b'} W_h Z_{hk}^{b'})}
\]

where \( b^* \) stands for each of the 15 combinations of the purchase bundle across the four categories. Since we are interested in the cross-category purchase patterns given a conversion, we can modify equation (3a) and rewrite the joint distribution excluding the “no-purchase-at-all” case as follows:

\[
P(Z_{hk} = Z_{hk}^b, \sum_{j=1}^{4} Z_{hkj} \neq 0) = \frac{\exp(\pi_{hk} Z_{hk}^b + \frac{1}{2} Z_{hk}^b W_h Z_{hk}^b)}{\sum_{b'} \exp(\pi_{hk} Z_{hk}^{b'} + \frac{1}{2} Z_{hk}^{b'} W_h Z_{hk}^{b'})}
\]

Details of the derivation of equation (4) are provided in the Appendix. This concludes the basic structure of modeling cross-category purchase incidences.

We next discuss how we model the intrinsic utility \( \pi \), and the cross-category interdependence parameter, \( \theta \). The intrinsic utility is a function of the various covariates that affect purchase decisions. This includes both product and keyword level variables. The main product level variable is price. The keyword level variables that determine purchases include the rank of the keyword on the search engine results page. This is based on prior work that has shown empirical evidence that higher ranked ads lead to higher conversion rates (Ghose and Yang 2007). Furthermore, we also consider the extent of latency in the purchase as a factor that is associated with the utility from a purchase. Hence, we specify the following equations:

\[
\pi_{hkj} = \beta_{hj0} + \beta_{hj1} Price_{hkj} + \beta_{hj2} Rank_{hk} + \beta_{hj3} Latency_{hk}
\]

\[
\beta_{hj0} = \alpha_{j0} + \alpha_{j1} S_h + \alpha_{j2} S_h^2 + \alpha_{j3} S_h^3 + \alpha_{j4} S_h^4 + \alpha_{j5} Brand_h + \eta_{hj}
\]

In equation (6), \( S_h \) is a dummy variable indicating whether the keyword \( h \) refers to the product category \( j \). Similarly, \( Brand_h \) is a dummy variable indicating whether the keyword specifies a brand name or not.
As a result, $\alpha_{j0}$ can be interpreted as the baseline effect of retailer-specific keywords on purchase decisions. The unobservable effects are denoted by $\eta_{bj}$ and are normally distributed as follows:

$$\eta_{bj} \sim N(0, \sigma_j^2)$$  \hspace{1cm} (7a)

Finally, we specify that

$$\beta_{hjm} \sim N(\bar{\beta}_{jm}, \tau_{jm}^2) \quad \text{for } j=1,2,3,4 \text{ and } m=1,2,3$$  \hspace{1cm} (7b)

where $\tau_{jm}^2$ captures the covariance in consumers’ response sensitivities to the marketing-mix variables such as price and keyword rank. We further model the heterogeneity in the degree of cross-category interdependence as follows:

$$\gamma_{j'0}^{j'} + \gamma_{j'1}^{j'} Brand_h + \gamma_{j'2}^{j'} Product_h + \xi_{h}^{j'} \quad \text{for } j=1,2,3,4 \text{ and } j' > j$$  \hspace{1cm} (8a)

$$\xi_{h}^{j'} \sim N(0, \delta_{j'}^2)$$  \hspace{1cm} (8b)

In equation (9a), Product is a dummy variable indicating whether the keyword specifies a particular product or not. As before, Brand is a dummy variable indicating whether the keyword specifies a brand name or not. Similar to equation (6), $\gamma_{j'0}^{j'}$ can be interpreted as the degree of interdependence between category $j$ and $j'$ for retailer specific keywords.

This model is specified in a hierarchical form, and this hierarchy consists of three levels. The first level captures the choice of items for the shopping basket during a shopping session, given a conversion. The second level captures both observed and unobserved differences on the model parameters across keywords. The third level specifies the priors for the unknown parameters.

4. Data

We first describe the data generation process for paid search advertisements since it differs on many dimensions from traditional offline advertisements. Once a keyword ad gets ranked by the search engine, these sponsored ads show up on the top left, and right of the computer screen in response to a query that a
consumer types on the search engine. The textual ad typically consists of headline, a limited number of
words describing the product or service and a hyperlink that refers the consumer to the advertiser’s website
after a click. The serving of an ad in response to a query containing a keyword is denoted as an impression. If
the consumer clicks on the ad, she is led to the landing page of the advertiser’s website. In the event that the
consumer ends up purchasing a product from the advertiser, this is recorded as a conversion. The time be-
tween a click on an ad and an actual purchase (conversion) from that ad is known as latency. This is usually
measured in days. In the majority of our data the value of this variable is 0, suggesting that the consumer
placed an order on the same day as when she clicked on the firm’s ad on the search engine page.

Our data contains weekly information on paid search advertising from a large nationwide retail
chain, which advertises on Google. The data span all keyword advertisements by the company during a pe-
riod of three months in the first quarter of 2007, specifically for the 24 calendar weeks from January 1 to
June 31. Most datasets that have been used in prior work to investigate consumer behavior in on-line envi-
ronments usually comprise of browsing behavior only and associated click-stream data. In contrast, our data
are unique in that we have individual ad-level stimulus (keyword ad) and response (purchase incidence).

We aim to investigate the impact of sponsored search advertising in a given category on consumers’
propensity to buy products from that as well as other categories. Our dataset has detailed information on the
various categories of products that were eventually purchased by consumers after they had clicked on any
given paid advertisement. Our dataset includes 8085 observations from a total of 278 unique keywords. Each
observation corresponds to a unique order (conversion). The data consists of keyword ads from 4 categories
of products that this nationwide chain retailer sells (bath, bedding, home décor and kitchen). Given that sur-
veys have revealed that a vast majority of consumers (92%) use a search engine to help research and/or pur-
chase appliances, kitchenware, furniture, and home improvement items (GMI 2005), our data is thus very
representative of actual user behavior in the sponsored search context. We have information on two impor-
tant keyword-specific characteristics, which we next discuss.

3 The firm is a large Fortune-500 retail store chain with several hundred retail stores in the US but due to the nature of the data shar-
ing agreement between the firm and us, we are unable to reveal the name of the firm.
In recognition of online consumer search behavior, search engines not only sell non-branded, generic keywords as advertisements, but also well-known brand names that can be purchased by advertisers in order to refer consumers to its Web site.\(^4\) Hence, we focus on the two important keyword specific characteristics for a firm (the advertiser) when it advertises on a search engine. This includes whether the keyword has (i) retailer-specific information (for example, “Retailername”, “Retailer Name”, “RetailerName.com”), and (ii) brand-specific information (for example, a product or manufacturer brand name). We enhanced the dataset by introducing the two keyword-specific characteristics Brand and Retailer. To be precise, for creating the variable in (i), we looked for the presence of the advertising retailer’s name in the keyword, and then labeled the dummy as 1 or 0, with 1 indicating the presence of the retailer’s name. For (ii), we looked for the presence of a brand name (either a product-specific or a company specific) in the keyword, and labeled the dummy as 1 or 0, with 1 indicating the presence of a brand name. We have several unique brands represented by these keywords and for the same advertiser we have 63 different combinations of its name, each represented by a unique keyword. This kind of keyword categorization is consistent with prior work such as Rutz and Bucklin (2007) and Ghose and Yang (2007).

As shown in Table 1a, about 23% of the keyword advertisements in our data include the retailer’s name, and approximately 36% include a brand name. And finally 77% of the keywords have product category information embedded in them which enables us to identify which category a keyword belongs to. By distinguishing between retailer information and brand information in keywords, we gain insights into the implications of searches originating from consumers who are aware of the advertiser and are likely to buy from that specific firm (retailer-specific keywords) relative to those consumers who are aware of a well known manufacturer brand (brand-specific keywords), and are hence likely to be more vulnerable to competition from other retailers who sell the same brand. We discuss further implications in Section 6.

\[^{4}\text{For example, a consumer seeking to purchase a digital camera is as likely to search for a popular brand name such as NIKON, CANON or KODAK on a search engine as searching for the generic phrase “digital camera” on the same search engine. Similarly, the same consumer may search for a retailer such as “BEST BUY” or “CIRCUIT CITY” on the search engine.}\]
Each order can lead to a purchase from the searched product category and/or from any of the other
three non-searched product categories. The joint pair-wise purchase incidence across all four categories is
given in Table 1a. For example, bedding and bath were bought 99 times in the same shopping session. The
other elements in this table give the instances when there was a purchase only in that category. For example,
a bedding item was bought on 527 shopping sessions. In sum, 12.78% of the observations involved a joint
purchase of products from two different categories in a session after a single keyword ad was clicked.

Table 1b reports the summary statistics of the data. There is considerable variation across the four
categories. The average price in Category 2 is 92.64 dollars (most expensive category) while the average
price of products in Category 1 is 34.77 dollars (least expensive category). The average keyword rank is
about 1.72 and the average latency is about 2.83 days. These statistics provide evidence suggesting that key-
word advertising can lead to purchases on a non-searched product category in addition to the searched cate-
gory, and that consumers may wait for a while after clicking on an ad to complete a purchase order.

5. Estimation and Results

We used the Markov Chain Monte Carlo (MCMC) methods to make inferences on model parame-
ters. In the Bayesian estimation, repeated draws were made from the series of full conditionals to arrive at the
joint posterior density of the unknown quantities using the MCMC Gibbs Sampling algorithm coupled with
Metropolis Hasting algorithm. We ran the Markov Chain for 10,000 iterations. Convergence was ensured by
monitoring the time-series of the draws from the full conditional distributions (Rossi and Allenby 2003). Ini-
tial iterations reflect a "burn-in" period where the chain may not have converged. We chose a burn-in length
of 5,000 iterations. Therefore, we retained only the last 5,000 draws from the posterior distributions for infe-
rence purposes. Since sequential draws from the joint posterior may be highly correlated, we adopt the "thin-
ning the chain" (Geyer 1992) procedure whereby we retain every 10\textsuperscript{th} draw from the retained part of the
chain to make inferences. Details of the algorithm are in the Online Appendix.
Compared with the classical estimation approach such as maximum likelihood estimation, the Bayesian estimation approach has several advantages. \textit{First}, unlike the MLE estimates which are not guaranteed to be global optimal, estimates from the Bayesian approach are usually global optimal. \textit{Second}, unlike the MLE estimates, the Bayesian estimation results are usually not sensitive to starting values. Finally and most importantly, the Bayesian method treats the individual level (keyword specific) parameters as model parameters and estimates them, rather than integrating them out as done in the maximum likelihood approach. As a result, we are able to obtain estimates (such as price sensitivity, rank sensitivity, and latency sensitivity) at the individual keyword level, which can be used for segmentation and targeting analysis.

5.1 Model Comparison

To establish the validity of our proposed model, we compare our model with four alternative models. Model 1 is the simplest model that does not account for cross-category interdependence, unobserved heterogeneity, and the effect of search characteristics on baseline utility as well as cross-category interdependence. In Model 2, we add unobserved heterogeneity through random coefficient estimation to Model 1, but do not account for the other effects. In Model 3, we add cross-category interdependence to Model 2, but do not account for the other effects. In Model 4, we then add the effect of search characteristics on the baseline utility to Model 3, but do not account for the last remaining effect of search characteristics on cross-category interdependence. Model 5 is the full model as proposed. Table 2 reports the fit statistics of the five models. As shown in table 2, the full model (model 5) outperforms the other four benchmark models, which highlights the importance of accounting for unobserved heterogeneity, the cross-category interdependence, and the effect of search characteristics on baseline utility and cross-category interdependence.

5.2 Empirical Results

We first discuss how search behavior in a given category affects the purchases in that category as well as spillovers into other categories. These estimates are reported in Table 3.
5.2.1 Keyword type and purchases

We highlight four main findings in this sub-section.

• **Within Category Search and Purchase Incidence:** On an average when users search for a keyword belonging to the bath category, the intention of purchasing from that category is the same compared to the purchase intention from a search using a “retailer” keyword. However, when users search for a keyword in the bedding, home décor, or kitchen categories, the intention of purchasing from the corresponding category is higher compared to the purchase intention from a search using a “retailer” keyword. Further, the within category search and purchase relationship is strongest for the bedding category and weakest for the kitchen category.

• **Across Category Search and Purchase Incidence:** When users search for a keyword in the bedding category, the intention of purchasing from the other categories (bath, home decor, or kitchen respectively) is lower compared to the purchase intention while searching using a “retailer” keyword, and is actually lowest for the kitchen category. When users search for a keyword in the home decor category, the intention of purchasing from the corresponding categories (bath, bedding, or kitchen) is lower compared to the purchase intention while searching using a “retailer” keyword, and is actually lowest for the kitchen category. When users search for a keyword in the kitchen category, the intention of purchasing from the other categories (bath, bedding, or home decor, respectively) is lower compared to the purchase intention while searching using a “retailer” keyword, and is actually lowest for the bath category.

• **Asymmetric Effect:** On an average when users search for a keyword in the bedding, home décor or kitchen category, they still exhibit a positive propensity to buy from the bath category. However, when users search for a keyword in the bath category, they do not exhibit an intention for buying from the bedding, home décor or kitchen categories which is statistically different from their purchase intention of buying from these categories using a “retailer” keyword.

• **Retailer and Brand:** When users search using a “retailer” keyword, on an average they exhibit the highest likelihood of buying from the bath category and the lowest likelihood of buying something from the
home décor category. For keywords containing “brand” information, on an average, consumers are less likely to buy products from categories 2 and 3 (bedding and home décor) but more likely to buy products from category 4 (kitchen) compared to their purchase intention in these categories when searching using a keyword containing the retailer information. Moreover, with respect to buying products from the bath category, there is no statistical difference in the likelihood of the consumer buying it from a “brand-specific” keyword and a “retailer-specific” keyword.

5.2.2 Impact of other covariates

Next we discuss how different covariates like price, keyword rank and latency affect product purchases within and across categories. The estimates from Table 4 suggest that as expected the category price is negatively associated with purchase intention in any category. In fact, category 1 (bath items) exhibits the highest price sensitivity while category 3 (home décor) has the lowest price sensitivity. Moreover, in terms of price sensitivity, one can see that categories 1 and 4 (bath and kitchen) are more similar to each other while categories 2 and 3 (bedding and home décor) are more similar to each other.

There is a statistically significant and as expected, negative effect of advertisement Rank on purchase incidence for the bath, bedding and kitchen categories. In other words, keyword ads that are placed higher up on the search engine’s results screen are more likely to lead to purchases in these categories. Moreover, categories 1 and 4 (bath and kitchen) are more sensitive to ad position on the search engine results page than category 2 (bedding). These findings have important implications since they show that the value per click on a given ad is not uniform across different slots on the search engine results page.

Latency does not seem to affect purchase incidence in any category except in category 3 (home décor) where it has a negative relationship with purchase incidence. This suggests that a longer duration of time between click and purchase is associated with a lower propensity to buy a product. This may be surprising at first but recall that these are average estimates aggregated across all keywords within each category. In order to better examine the effect of latency on purchase intentions, we plot the individual keyword level es-
estimates for latency for each of the four product categories. Note that for many keywords in each of the categories, the impact of latency is positive although for the majority of keywords it is negative.

Table 5 show the statistical significance of variance of the random coefficient parameters, implying that there is unobserved variance on the response parameters of price, rank and latency across keywords.

5.2.3 Cross-Category Interdependence

Next, we discuss the estimates of cross-category interdependence. The results from the estimation of the auto-logistic model suggest that “branded” keywords are less likely to induce cross-category purchases between categories 1 and 2 (bath and bedding), and between categories 3 and 4 (home décor and kitchen), compared to retailer keywords. Furthermore, generic keywords are less likely to exhibit positive interdependence in purchases from categories 1 and 4 (bath and kitchen), and purchases from categories 2 and 3 (bedding and home décor), compared to retailer keywords. Finally, the estimates on the intercept demonstrate that for retailer keywords in 5 out of 6 cases, we see strong evidence of positive cross-category interdependence. There is only one instance where two categories exhibit negative cross-category interdependence or a substitution-like pattern, and this is seen in the case when consumers purchase both categories 2 and 4 (bedding and kitchen items). Finally, in terms of absolute magnitude, the cross-category interdependence is strongest for the bath and home décor categories, and weakest for the bedding and kitchen categories.

Note that the results from our autologistic model are different from much of prior work which has shown that cross-category correlations among all possible pairs among different product categories are always positive, if they are different from zero (Seetharaman et al. 2005). As Seetharaman et al. (2005) discuss, one of the common empirical findings in the literature is that there is a base level of complementarity estimated among all pairs of product categories. A plausible reason for this could be the large number of “no-purchase” outcomes that typically characterizes any offline scanner panel data. Because our interest is
in examining consumer cross-category purchase patterns given a conversion, such results are less likely to be reflected in our estimates. Thus, the prior stream of work suggests that there appears to be an intrinsic propensity for any pair of product categories to co-occur within a household’s shopping basket. The results from our autologistic model wherein we find negative inter-dependence between the bedding and kitchen categories, highlight that this may not necessarily the case, especially when one examines data on final purchases.

To illustrate an example of cross-category purchase spillovers, we plot individual parameter estimates for two specific keywords. These are in the Online Appendix in Figure 2. One is a generic keyword, “Dinnerware”, and the other is a branded keyword “Nautica Havana Bedsheets”. In our data, we observe that a search session that started using the first keyword led to a conversion in both the home décor and kitchen categories. Similarly, a search session that started with the second keyword led to a conversion in both the bedding and bath categories.

Table 5 reports the estimated variance of the interdependence parameters. They all turn out to be highly significant, suggesting that the degree of cross-category interdependence is different across keywords, and driven by unobserved factors other than the two covariates of “brand” and “product”.

Finally, to alleviate any concerns in price estimation, we re-estimate our model by using an instrument variable approach. Due to the limited supply of available instruments, we follow prior work and use lagged value of price as an instrument (Villas-Boas and Winer 1999). Our estimation results remain highly consistent with what we obtained before, which suggests that price endogeneity is not an issue in this context.\(^5\)

5.3 Counterfactuals
In this section, we conduct two sets of counterfactual experiments to derive some insights from our empirical analyses. In particular, we analyze how profits change if firms were able to (i) induce changes in consumers’ cross-category interdependence parameter ($\Theta_{ij}$), and (ii) advertise a uniform price discount on products in one or multiple categories. The former strategy can be adopted by engaging in product-level or category-level promotional advertisements that increase consumer utility from buying multiple categories in a given search session. Coupons, for example, are a useful tool in this regard. The latter strategy can be adopted by changing the extent of price discount information provided in the ad creative that appears on the search engine results page (we expand on this later). 6

5.3.1 Changes in Cross-Category Interdependence

Towards the first analyses, we compute total profits from the four categories with the current estimates and use equations (4), (5), and (6) along with data on profit margins within each category. We then compare these profits with the cases when $\Theta_{ij}$ takes the four hypothetical values 1.5, 1, 0 and -1 for all pairs of categories (i.e., the parameter estimates from equation (3b) are modified with the specific value of $\Theta_{ij}$). These values are chosen simply for illustrative purposes to examine the impact on profits if firms are able to induce similar levels of purchase co-incidence between any two pairs of categories. Table 8a summarizes the results. There is a considerable change in profits with changes in the value of the cross-category interdependence parameter. These analyses also highlight that a strategy of equalizing the cross-category interdependence in all categories (for example, making $\Theta_{ij} = 1$) may reduce profits compared to the current situation. Thus, retailers can benefit from a more targeted approach for inducing category-level purchase spillovers.

---

3 Further, because the quality of the landing pages can also influence conversions (see for example, Ghose and Yang 2007), we also estimated the model with a variable that indicates the landing page quality score associated with a given keyword. The scoring details are provided in Ghose and Yang (2007). Again, there was no qualitative change to the nature of our results.

6 A typical sponsored search ad contains a title line and some description of the offering followed by the URL of the advertiser. This text is referred to as the “creative” and plays an important role in conveying information about the firm and its offerings.
5.3.2 Price Discount by Keyword Category

A common issue that is often raised in trade press reports is the extent to which information about price discounts should be explicitly mentioned in the ad creative of the sponsored search ad. This becomes important because there are trade-offs between having a quality-based versus price-based USP (Unique Selling Proposition) in these ads. This is discussed in Animesh et al. (2008) who examine how a firm’s positioning strategy as highlighted in its ad-creative influences the effectiveness of its listing in sponsored search. They find that under some conditions the click-through rate of a firm is influenced by the firm’s positioning strategy as highlighted by the price or quality USP in its ad creative. The theoretical reason behind this argument is that while all consumers may search sequentially through the ordered listings of search engine ads, a consumer who is quality seeking has a higher likelihood of clicking on ads that signal higher product quality, while a price sensitive consumer has a higher likelihood of clicking on ads that signal lower prices.

One heuristic that firms can use in this context is to advertise a price discount in all keywords that are associated with their product categories and highlight that in the ad creative. That is, any keyword that contains product information can also have information about the price discount stated in its ad text. A relatively more sophisticated approach would be to select those keywords that belong to a certain category and advertise the price discount in the creative of those keywords. Towards examining how the firms’ profit changes when one were to adopt such different heuristics, we conduct a number of different counterfactuals. We compute profits with the current estimates in tables 3 and 4, and use equations (4), (5), and (6) along with data on profit margins within each category. Our baseline case is the profit with the current estimates. We compare the current profits to the case when the firm engages in advertising a 10% discount in keyword ads that represent certain categories. We conduct the counterfactuals for different categories and with different combinations of categories as shown in Table 8b.7

A useful implication from this analysis is that advertising a uniform price discount in all keywords associated with a given product category is not always profitable. For example, while in general profits tend

---

7 The use of 10% discount is simply for illustrative purposes. The qualitative nature of our results is not sensitive to the actual magnitude of the discount used in these analyses.
to increase when a flat 10% discount is applied to a given product category, there are exceptions to this rule. Specifically, there is actually a decrease in profits when such a uniform discount is applied to keywords representing products in category 4 (kitchen items). These differences are driven by the variations in price sensitivities across categories. So retailers need to be careful about advertising a uniform price discount in the sponsored keyword ads. This calls attention for the need to customize the content of the text in the ad creative. Indeed, this becomes especially relevant because of the evidence of spillover effects accruing from searches in one category to purchases in other categories.

5.3.3 Price Discount by Keyword Category and Type

Another heuristic firms could adopt is to select keywords based on their type: retailer, brand or product-specific, and customize the price discount in the ad creative accordingly. Recall that a keyword is categorized as being “retailer-specific” if it includes the advertiser’s name. Such search queries typically imply that the user is already aware of the advertiser. That prior knowledge may have come from a variety of other marketing interactions, and in many cases, denote existing customers who are returning to buy from that firm again. A substantial percentage of sponsored search ads do not bring in new prospects—they simply deliver people who are already actively looking for the firm (Atlas 2008). Thus, depending on the keyword used to arrive at the website, different consumers may have a different purchase intention. Some may be much closer to the purchase point than others. Hence, it makes sense for firms to customize the price promotion offer to reap the benefits of differences in consumer intent during the search process. In practice, this is implemented by referrals to different landing pages within the company’s website based on the type of keyword.

Table 8b shows the results from offering a price discount in a given category based on the keyword that was clicked during the search process. For example, for consumers who arrive at the advertiser’s website after clicking on “retailer-specific” keywords, it seems more profitable for firms to give a price discount on all categories. However, for consumers who arrive at the advertiser’s website after clicking on “brand-specific” keywords, it seems most profitable for firms to give a price discount on specific categories (categories 2 and 3). Finally, for consumers who arrive at the advertiser’s website after clicking on the product-
specific keyword, it seems most profitable for firms to give a price discount on a combination of categories 1, 2 and 3. These analyses highlight the importance of customizing the price promotion by selecting the appropriate sample of keywords based on consumers’ initial search patterns.

6. Discussion and Managerial Implications

Although the average click-through and conversion rates are typically low in online advertising, there can be other benefits from advertising in such sponsored search media. Specifically, retailers can set up relevant cross-selling opportunities on their own websites by advertising ‘brand-specific’ and ‘product-specific’ keywords. The strategy is that when a consumer searches for a specific product and visits the retailer’s website by clicking on its advertisement, the retailer can pair that product with other products whose sales are also known to have been associated with that keyword from prior purchase incidences. These bundles of products can then be prominently featured on its website. This provides a retailer with an opportunity to not only convert someone on the product they had searched for, but also get other opportunities for cross-selling in a sponsored advertising environment. From an economic perspective, such within category purchases and cross-category purchase spillovers can be especially beneficial since a single click on an ad may lead to multiple purchases and thus significantly lower customer acquisition costs.

Our analysis suggests that retailers could use such insights to design explicit recommendation systems for promoting co-purchase networks like what Amazon does. The difference would be that while co-purchase networks in Amazon are mostly geared towards up-selling or vertical spillovers, the information gained from our model and framework is mostly suited for cross-selling or horizontal spillovers. From the retailer’s perspective, there could be synergies in promoting multiple categories simultaneously rather than separately. Indeed anecdotal evidence suggests that retailers are taking cross-selling reports from other marketing mix campaigns and putting up the top cross-selling product for the searched product on the same page (Squire 2003). Furthermore, consumers who display high cross-selling potential during paid search advertising can also be targeted with coupons customized to induce such co-purchases, not only in the online world
but also in the offline world. This becomes important in light of studies that have shown that 79% of users who search on Google end up purchasing offline at a retail store location (GMI 2005).

Interestingly, we find that latency in purchases is not detrimental for a firm that is sponsoring the online advertisement. We find that latency is associated with an increase in consumers’ spending in other products. This effect seems particularly strong in keywords that have a brand name in it, since consumers who click on branded keywords typically spend more on other categories than the one they were originally searching for. Thus, online advertisers can focus on investing more in such keywords relative to generic keywords, especially if the cannibalization effect of drawing out consumers from one category is smaller relative to revenue expansion effect. From the point of view of a manufacturer, such dependencies across categories may be exploited by running cooperative promotions within brands but across categories. To be clear, such decisions would need a detailed profitability analysis based not only on the potential from cross-selling in other product categories but also the revenue performance of the keyword in its own category.

Further, online advertisers know that in sponsored search, they can customize the ad creative and thus differentiate themselves from their competitors. For example, when advertising price discounts, they could either adopt a blanket advertising strategy in which all keywords have price discount information versus a more selective advertising strategy in which only certain keywords have price discount information embedded in the ad creative. An understanding of these issues in the sponsored search context requires an in-depth analysis of the consumer search to purchase conversion process. In light of differences in search-to-purchase correlations between retailer, brand and generic keywords, analyzing changes in profits when a price discount is advertised in each of these three types of ads lends useful insights. Indeed, depending on the type of keyword clicked during the search process, a price discount on a given category leads to wide variation in profits. Our analyses suggest that retailers can benefit from real-time price targeting based on consumer search patterns. The firm can create customized landing pages and then point external links to different landing pages in order to better target its customers based on which keyword they were searching.

Even though, millions of keywords are searched every month in major search engines, only a small portion of that are purchased because advertisers usually think of sponsoring those keywords that are directly
related to their product. They are less inclined to bid on keywords that have an indirect impact and can thus be very relevant for sales. Based on some of the insights from this paper which tie user search behavior to actual purchase behavior, search engine marketing firms (SEMs) can develop a keyword cross-sell system that automatically selects commercially valuable keywords from search queries, and then actively suggests those keywords to advertisers who might be interested in purchasing them. Newer keyword generation tools can be designed to match keywords with advertisers. This consists of the following three steps: (1) finding commercially valuable keywords to utilize as “seed keywords” and then expanding the seed keywords to search sequentially and/or conceptually related keywords which are denoted as “expanded keywords”; (2) locating potential advertisers based on the expanded keywords, and (3) filtering out inappropriate cross-selling by analyzing attitude and/or cause-effect relationships (Nong et al. 2007). The keywords are then automatically revealed to the advertiser for consideration as relevant terms for their advertisements. This kind of matching can substantially increase the efficiency of the sponsored search market.

Our paper has several limitations. These limitations arise primarily from the lack of information in our data. For example, we do not have precise data on competition since our data is limited to one firm. That is, we do not know the keyword ranks or other performance metrics such as conversion rates of the keyword advertisements of the competitors of the firm whose data we have used in this paper. Further, we do not have any knowledge of other information that was mentioned in the textual description in the space following a paid advertisement during consumers’ queries. Future work could integrate that information with our modeling approach to have more precise estimates. In addition, future work could examine product-specific characteristics to see how different kind of products affects conversions. This will help firms analyze which brands or products have higher conversion rates and lower costs per conversion. We also do not have information on if users had been exposed to the different products in prior visits to the firm’s website. It is possible that some users may have searched for a given keyword and visited the firm’s website in a previous session. Despite these limitations, we hope that this study will generate further interest in exploring this emerging area.
Appendix: Derivation of the Joint Probability

Let us first define the probability of the case where no product from any of the four categories is purchased. This can be written as follows:

\[ P(Z_{hk1}, Z_{hk2}, Z_{hk3}, Z_{hk4}) = P(Z_{hk1} = 0, Z_{hk2} = 0, Z_{hk3} = 0, Z_{hk4} = 0) \]  

(A1)

Then

\[ P(Z_{hk1}, Z_{hk2}, Z_{hk3}, Z_{hk4}) = P(Z_{hk1} \mid Z_{hk2}, Z_{hk3}, Z_{hk4})P(Z_{hk2}, Z_{hk3}, Z_{hk4}) \]  

(A2)

\[ P(Z_{hk2}, Z_{hk3}, Z_{hk4}) = \frac{P(Z_{hk10}, Z_{hk2}, Z_{hk3}, Z_{hk4})}{P(Z_{hk10} \mid Z_{hk2}, Z_{hk3}, Z_{hk4})} \]  

(A3)

After substituting equation (A3) to equation (A2), we have

\[ P(Z_{hk1}, Z_{hk2}, Z_{hk3}, Z_{hk4}) = P(Z_{hk1} \mid Z_{hk2}, Z_{hk3}, Z_{hk4})P(Z_{hk2}, Z_{hk3}, Z_{hk4}) \frac{P(Z_{hk10}, Z_{hk2}, Z_{hk3}, Z_{hk4})}{P(Z_{hk10} \mid Z_{hk2}, Z_{hk3}, Z_{hk4})} \]  

(A4)

Following the same logic, we can obtain the other three counterparts:

\[ P(Z_{hk10}, Z_{hk2}, Z_{hk3}, Z_{hk4}) = P(Z_{hk2} \mid Z_{hk10}, Z_{hk3}, Z_{hk4}) \frac{P(Z_{hk10}, Z_{hk20}, Z_{hk3}, Z_{hk4})}{P(Z_{hk10} \mid Z_{hk2}, Z_{hk3}, Z_{hk4})} \]  

(A5)

\[ P(Z_{hk10}, Z_{hk20}, Z_{hk3}, Z_{hk4}) = P(Z_{hk3} \mid Z_{hk10}, Z_{hk20}, Z_{hk4}) \frac{P(Z_{hk10}, Z_{hk20}, Z_{hk30}, Z_{hk4})}{P(Z_{hk10} \mid Z_{hk2}, Z_{hk3}, Z_{hk4})} \]  

(A6)

\[ P(Z_{hk10}, Z_{hk20}, Z_{hk30}, Z_{hk4}) = P(Z_{hk4} \mid Z_{hk10}, Z_{hk20}, Z_{hk30}) \frac{P(Z_{hk10}, Z_{hk20}, Z_{hk30}, Z_{hk40})}{P(Z_{hk40} \mid Z_{hk10}, Z_{hk20}, Z_{hk30})} \]  

(A7)

After combining equations (A4)-(A7), we have

\[
\begin{align*}
\frac{P(Z_{hk1}, Z_{hk2}, Z_{hk3}, Z_{hk4})}{P(Z_{hk10}, Z_{hk20}, Z_{hk30}, Z_{hk40})} &= \frac{P(Z_{hk1} \mid Z_{hk2}, Z_{hk3}, Z_{hk4})}{P(Z_{hk10} \mid Z_{hk2}, Z_{hk3}, Z_{hk4})} \times \frac{P(Z_{hk2}, Z_{hk3}, Z_{hk4})}{P(Z_{hk10}, Z_{hk20}, Z_{hk30}, Z_{hk4})} \\
\frac{P(Z_{hk3} \mid Z_{hk10}, Z_{hk20}, Z_{hk4})}{P(Z_{hk30} \mid Z_{hk10}, Z_{hk20}, Z_{hk4})} &= \frac{P(Z_{hk3} \mid Z_{hk10}, Z_{hk20}, Z_{hk30})}{P(Z_{hk30} \mid Z_{hk10}, Z_{hk20}, Z_{hk4})} \\
\frac{P(Z_{hk4} \mid Z_{hk10}, Z_{hk20}, Z_{hk30})}{P(Z_{hk40} \mid Z_{hk10}, Z_{hk20}, Z_{hk30})} &= \frac{P(Z_{hk4} \mid Z_{hk10}, Z_{hk20}, Z_{hk30})}{P(Z_{hk40} \mid Z_{hk10}, Z_{hk20}, Z_{hk30})}
\end{align*}
\]  

(A8)
Substituting equation (1) into equation (A8), and we can write down the joint probability of the purchase on the four categories as,

\[
P(P(Z_{hk1}, Z_{hk2}, Z_{hk3}, Z_{hk4}) = P(Z_{hk10}, Z_{hk20}, Z_{hk30}, Z_{hk40}),
\]

\[
\exp[Z_{hk1}(\pi_{hk1} + W_{h1}'Z_{hk}^b)].\exp[Z_{hk2}(\pi_{hk2} + W_{h2}'Z_{hk}^{(0234)})].\exp[Z_{hk3}(\pi_{hk3} + W_{h3}'Z_{hk}^{(0034)})].\exp[Z_{hk4}(\pi_{hk4} + W_{h4}'Z_{hk}^{(0004)})]
\]

where \(Z_{hk}^{(0234)} = (Z_{hk1} = 0, Z_{hk2}, Z_{hk3}, Z_{hk4})\), \(Z_{hk}^{(0034)} = (Z_{hk1}, Z_{hk2} = 0, Z_{hk3}, Z_{hk4})\) and \(Z_{hk}^{(0004)} = (Z_{hk1}, Z_{hk2} = 0, Z_{hk3} = 0, Z_{hk4})\)

Equation (A9) can be further written as,

\[
P(P(Z_{hk1}, Z_{hk2}, Z_{hk3}, Z_{hk4}) = P(Z_{hk10}, Z_{hk20}, Z_{hk30}, Z_{hk40}) \exp[\pi_{hk} Z_{hk}^b + \sum_{i<j} \theta_{h}^{i,j} Z_{hk}^i Z_{hk}^j])
\]

Since

\[
Z_{hk}^b W_{h} Z_{hk}^b = 2 \sum_{i<j} \theta_{h}^{i,j} Z_{hk}^i Z_{hk}^j,
\]

the joint probability can then be written as

\[
P(Z_{hk} = Z_{hk}^b) = \exp(\pi_{hk} Z_{hk}^b + \frac{1}{2} Z_{hk}^b W_{h} Z_{hk}^b) P(Z_{hk10}, Z_{hk20}, Z_{hk30}, Z_{hk40})
\]

Since our data do not include observations with no purchase on any of the four categories, and the probability of different combinations of purchases across the four categories adds up to 1, we then have,

\[
P(Z_{hk} = Z_{hk}^b, \sum_{j=1}^{4} Z_{hkj} \neq 0) = \frac{\exp(\pi_{hk} Z_{hk}^b + \frac{1}{2} Z_{hk}^b W_{h} Z_{hk}^b)}{\sum_{b^*} \exp(\pi_{hk} Z_{hk}^{b^*} + \frac{1}{2} Z_{hk}^{b^*} W_{h} Z_{hk}^{b^*})}
\]

where \(b^*\) stands for each of the 15 purchase combinations on the four product categories excluding the no-purchase-at-all case.
References


Table 1a: Descriptive Statistics (N=8085)

<table>
<thead>
<tr>
<th>Variable</th>
<th>Frequency</th>
</tr>
</thead>
<tbody>
<tr>
<td>Purchase of Category 1 alone</td>
<td>527</td>
</tr>
<tr>
<td>Purchase of Category 2 alone</td>
<td>2240</td>
</tr>
<tr>
<td>Purchase of Category 3 alone</td>
<td>2091</td>
</tr>
<tr>
<td>Purchase of Category 4 alone</td>
<td>2193</td>
</tr>
<tr>
<td>Purchase of Category 1 and 2</td>
<td>99</td>
</tr>
<tr>
<td>Purchase of Category 1 and 3</td>
<td>183</td>
</tr>
<tr>
<td>Purchase of Category 1 and 4</td>
<td>115</td>
</tr>
<tr>
<td>Purchase of Category 2 and 3</td>
<td>203</td>
</tr>
<tr>
<td>Purchase of Category 2 and 4</td>
<td>124</td>
</tr>
<tr>
<td>Purchase of Category 3 and 4</td>
<td>310</td>
</tr>
<tr>
<td>Keywords referring to retailer</td>
<td>63</td>
</tr>
<tr>
<td>Keywords referring to brand</td>
<td>100</td>
</tr>
<tr>
<td>Keywords referring to product</td>
<td>215</td>
</tr>
</tbody>
</table>

Table 1b: Descriptive Statistics

<table>
<thead>
<tr>
<th>Variable</th>
<th>Mean</th>
<th>Std. Dev.</th>
<th>Min</th>
<th>Max</th>
</tr>
</thead>
<tbody>
<tr>
<td>Price of Category 1 (dollars)</td>
<td>34.77</td>
<td>10.97</td>
<td>1.59</td>
<td>238.92</td>
</tr>
<tr>
<td>Price of Category 2 (dollars)</td>
<td>92.64</td>
<td>51.10</td>
<td>1.59</td>
<td>711.88</td>
</tr>
<tr>
<td>Price of Category 3 (dollars)</td>
<td>43.40</td>
<td>27.63</td>
<td>0.99</td>
<td>799.98</td>
</tr>
<tr>
<td>Price of Category 4 (dollars)</td>
<td>71.15</td>
<td>52.12</td>
<td>0.99</td>
<td>1599.96</td>
</tr>
<tr>
<td>Rank</td>
<td>1.72</td>
<td>2.10</td>
<td>1.00</td>
<td>41.00</td>
</tr>
<tr>
<td>Latency</td>
<td>2.83</td>
<td>6.28</td>
<td>0.00</td>
<td>30.00</td>
</tr>
</tbody>
</table>

Table 2: Fit Comparison across Alternative Models

<table>
<thead>
<tr>
<th>Model 1</th>
<th>Model 2</th>
<th>Model 3</th>
<th>Model 4</th>
<th>Model 5</th>
</tr>
</thead>
<tbody>
<tr>
<td>Random coefficients</td>
<td>No</td>
<td>Yes</td>
<td>Yes</td>
<td>Yes</td>
</tr>
<tr>
<td>Cross category interdependence</td>
<td>No</td>
<td>No</td>
<td>Yes</td>
<td>Yes</td>
</tr>
<tr>
<td>Effect of search characteristics on the baseline utility</td>
<td>No</td>
<td>No</td>
<td>No</td>
<td>Yes</td>
</tr>
<tr>
<td>Effect of search characteristics on the cross- category interdependence</td>
<td>No</td>
<td>No</td>
<td>No</td>
<td>No</td>
</tr>
<tr>
<td>Model Fit Statistics (Log Marginal Density⁸)</td>
<td>-18190</td>
<td>-12649</td>
<td>-12218</td>
<td>-11457</td>
</tr>
</tbody>
</table>

³ Log Marginal Density is a measure of model fit used in Bayesian estimation, and it is asymptotically consistent with the BIC measure used in classical estimation.
Table 3: The Effect of Search Characteristics on Purchase Intention

<table>
<thead>
<tr>
<th>Category</th>
<th>Search_C1</th>
<th>Search_C2</th>
<th>Search_C3</th>
<th>Search_C4</th>
<th>Brand</th>
</tr>
</thead>
<tbody>
<tr>
<td>Category 1 Purchase</td>
<td>-5.157 (0.095)</td>
<td>0.648 (0.772)</td>
<td>-1.108 (0.265)</td>
<td>-2.141 (0.165)</td>
<td>-3.619 (0.231)</td>
</tr>
<tr>
<td>Category 2 Purchase</td>
<td>-4.028 (0.194)</td>
<td>0.903 (1.023)</td>
<td>3.280 (0.372)</td>
<td>-1.673 (0.253)</td>
<td>-2.365 (0.350)</td>
</tr>
<tr>
<td>Category 3 Purchase</td>
<td>-3.535 (0.184)</td>
<td>0.479 (0.986)</td>
<td>-1.070 (0.396)</td>
<td>2.824 (0.270)</td>
<td>-2.330 (0.428)</td>
</tr>
<tr>
<td>Category 4 Purchase</td>
<td>-3.759 (0.165)</td>
<td>-0.180 (0.831)</td>
<td>-2.010 (0.286)</td>
<td>-2.782 (0.208)</td>
<td>0.722 (0.302)</td>
</tr>
</tbody>
</table>

Note: Posterior means and posterior standard deviations (in the parenthesis) are reported, and estimates that are significant at 95% are bolded in Tables 3 - 7.

Table 4: Effect of Other Variables on Purchase Intention

<table>
<thead>
<tr>
<th>Category</th>
<th>Price</th>
<th>Rank</th>
<th>Latency</th>
</tr>
</thead>
<tbody>
<tr>
<td>Category 1 Purchase</td>
<td>-0.157 (0.038)</td>
<td>-0.317 (0.043)</td>
<td>-0.069 (0.042)</td>
</tr>
<tr>
<td>Category 2 Purchase</td>
<td>-0.077 (0.030)</td>
<td>-0.154 (0.042)</td>
<td>-0.036 (0.031)</td>
</tr>
<tr>
<td>Category 3 Purchase</td>
<td>-0.060 (0.027)</td>
<td>-0.039 (0.054)</td>
<td>-0.112 (0.038)</td>
</tr>
<tr>
<td>Category 4 Purchase</td>
<td>-0.151 (0.044)</td>
<td>-0.308 (0.065)</td>
<td>-0.068 (0.044)</td>
</tr>
</tbody>
</table>

Table 5: Random Coefficients Variances of Parameters in Purchase Intention

<table>
<thead>
<tr>
<th>Category</th>
<th>Intercept</th>
<th>Price</th>
<th>Rank</th>
<th>Latency</th>
</tr>
</thead>
<tbody>
<tr>
<td>Category 1 Purchase</td>
<td>0.498 (0.096)</td>
<td>0.184 (0.025)</td>
<td>0.233 (0.030)</td>
<td>0.148 (0.025)</td>
</tr>
<tr>
<td>Category 2 Purchase</td>
<td>1.659 (0.198)</td>
<td>0.101 (0.014)</td>
<td>0.233 (0.029)</td>
<td>0.134 (0.023)</td>
</tr>
<tr>
<td>Category 3 Purchase</td>
<td>1.664 (0.166)</td>
<td>0.115 (0.015)</td>
<td>0.330 (0.049)</td>
<td>0.148 (0.019)</td>
</tr>
<tr>
<td>Category 4 Purchase</td>
<td>1.085 (0.114)</td>
<td>0.111 (0.014)</td>
<td>0.251 (0.027)</td>
<td>0.131 (0.022)</td>
</tr>
</tbody>
</table>

Notes: This table shows statistical significance of the variances and highlights that a random coefficients model is more appropriate than a keyword level fixed effects model.
### Table 6: Estimates on the Cross-Category Interdependence

<table>
<thead>
<tr>
<th>Category</th>
<th>Intercept</th>
<th>Brand</th>
<th>Product</th>
</tr>
</thead>
<tbody>
<tr>
<td>1 and 2</td>
<td>0.742</td>
<td>-0.118</td>
<td>-0.026</td>
</tr>
<tr>
<td></td>
<td>(0.123)</td>
<td>(0.047)</td>
<td>(0.158)</td>
</tr>
<tr>
<td>1 and 3</td>
<td>1.287</td>
<td>-0.006</td>
<td>-0.158</td>
</tr>
<tr>
<td></td>
<td>(0.141)</td>
<td>(0.151)</td>
<td>(0.177)</td>
</tr>
<tr>
<td>1 and 4</td>
<td>0.839</td>
<td>0.071</td>
<td>-0.292</td>
</tr>
<tr>
<td></td>
<td>(0.145)</td>
<td>(0.142)</td>
<td>(0.127)</td>
</tr>
<tr>
<td>2 and 3</td>
<td>0.621</td>
<td>0.120</td>
<td>-0.283</td>
</tr>
<tr>
<td></td>
<td>(0.131)</td>
<td>(0.118)</td>
<td>(0.128)</td>
</tr>
<tr>
<td>2 and 4</td>
<td>-0.272</td>
<td>0.044</td>
<td>0.173</td>
</tr>
<tr>
<td></td>
<td>(0.119)</td>
<td>(0.138)</td>
<td>(0.156)</td>
</tr>
<tr>
<td>3 and 4</td>
<td>0.886</td>
<td>-0.303</td>
<td>-0.217</td>
</tr>
<tr>
<td></td>
<td>(0.115)</td>
<td>(0.139)</td>
<td>(0.101)</td>
</tr>
</tbody>
</table>

### Table 7: Random Coefficients Variances of the Cross-Category Interdependence

<table>
<thead>
<tr>
<th>Category</th>
<th>Intercept</th>
</tr>
</thead>
<tbody>
<tr>
<td>1 and 2</td>
<td>0.874</td>
</tr>
<tr>
<td></td>
<td>(0.096)</td>
</tr>
<tr>
<td>1 and 3</td>
<td>1.036</td>
</tr>
<tr>
<td></td>
<td>(0.104)</td>
</tr>
<tr>
<td>1 and 4</td>
<td>0.857</td>
</tr>
<tr>
<td></td>
<td>(0.095)</td>
</tr>
<tr>
<td>2 and 3</td>
<td>0.605</td>
</tr>
<tr>
<td></td>
<td>(0.058)</td>
</tr>
<tr>
<td>2 and 4</td>
<td>1.054</td>
</tr>
<tr>
<td></td>
<td>(0.111)</td>
</tr>
<tr>
<td>3 and 4</td>
<td>0.708</td>
</tr>
<tr>
<td></td>
<td>(0.076)</td>
</tr>
</tbody>
</table>
### Table 8a: Impact of Changes in Interdependence on Total Profit

<table>
<thead>
<tr>
<th>Value of $\theta^\phi_h$</th>
<th>Percentage change in profits compared to baseline profits</th>
</tr>
</thead>
<tbody>
<tr>
<td>1.5</td>
<td>21.6 % increase</td>
</tr>
<tr>
<td>1</td>
<td>24.5 % decrease</td>
</tr>
<tr>
<td>0</td>
<td>70.6 % decrease</td>
</tr>
<tr>
<td>-1</td>
<td>89.5 % decrease</td>
</tr>
</tbody>
</table>

### Table 8b: Impact of Discount in Ad Creative on Total Profit

<table>
<thead>
<tr>
<th>Category on which price discount is applied</th>
<th>Whole Sample</th>
<th>Retailer Keywords Only</th>
<th>Brand Keywords Only</th>
<th>Product Keywords Only</th>
</tr>
</thead>
<tbody>
<tr>
<td>Category 1</td>
<td>0.12 % increase</td>
<td>0.25 % increase</td>
<td>0.22 % increase</td>
<td>0.34 % decrease</td>
</tr>
<tr>
<td>Category 2</td>
<td>0.99 % increase</td>
<td>7.2 % increase</td>
<td>2.48 % increase</td>
<td>0.14 % decrease</td>
</tr>
<tr>
<td>Category 3</td>
<td>16.4 % increase</td>
<td>7.63 % increase</td>
<td>0.98 % increase</td>
<td>1.74 % increase</td>
</tr>
<tr>
<td>Category 4</td>
<td>9.79 % decrease</td>
<td>3.86 % increase</td>
<td>9.1 % decrease</td>
<td>5.66 % decrease</td>
</tr>
<tr>
<td>Category 1 and 2</td>
<td>0.98 % increase</td>
<td>6.52 % increase</td>
<td>2.47 % increase</td>
<td>0.12 % decrease</td>
</tr>
<tr>
<td>Category 1 and 3</td>
<td>17 % increase</td>
<td>6.78 % increase</td>
<td>0.95 % increase</td>
<td>1.67 % increase</td>
</tr>
<tr>
<td>Category 1 and 4</td>
<td>0.01 % increase</td>
<td>3.36 % increase</td>
<td>8.39 % decrease</td>
<td>6.31 % decrease</td>
</tr>
<tr>
<td>Category 2 and 3</td>
<td>17.12 % increase</td>
<td>13.66 % increase</td>
<td>4.12 % increase</td>
<td>1.87 % increase</td>
</tr>
<tr>
<td>Category 2 and 4</td>
<td>0.08 % increase</td>
<td>10.39 % increase</td>
<td>5.96 % decrease</td>
<td>6.21 % decrease</td>
</tr>
<tr>
<td>Category 3 and 4</td>
<td>16.2 % increase</td>
<td>10.61 % increase</td>
<td>7.51% decrease</td>
<td>4.33 % decrease</td>
</tr>
<tr>
<td>Category 1, 2, and 3</td>
<td>17.2 % increase</td>
<td>13.91 % increase</td>
<td>3.37 % increase</td>
<td>1.89 % increase</td>
</tr>
<tr>
<td>Category 1, 2, and 4</td>
<td>0.99 % increase</td>
<td>10.33 % increase</td>
<td>5.93 % decrease</td>
<td>6.19 % decrease</td>
</tr>
<tr>
<td>Category 1, 3, and 4</td>
<td>16.27 % increase</td>
<td>10.95 % increase</td>
<td>7.48 % decrease</td>
<td>4.31 % decrease</td>
</tr>
<tr>
<td>Category 2, 3, and 4</td>
<td>17.15 % increase</td>
<td>17.47 % increase</td>
<td>0.1 % increase</td>
<td>4.11 % decrease</td>
</tr>
<tr>
<td>Category 1, 2, 3, and 4</td>
<td>17.23 % increase</td>
<td>17.71 % increase</td>
<td>5.03 % decrease</td>
<td>4.1 % decrease</td>
</tr>
</tbody>
</table>

**Notes:** Each column shows the percentage change in profits compared to baseline profits.
Online Appendix: The MCMC Algorithm

1. Draw \( \theta_h = (\theta_h^{12}, \theta_h^{13}, \theta_h^{14}, \theta_h^{23}, \theta_h^{24}, \theta_h^{34})' \)

We use Metropolis-Hastings algorithm with a random walk chain to generate draws of \( \theta_h \). Let \( \theta_h^{(p)} \) denote the previous draw, and then the next draw \( \theta_h^{(n)} \) is given by:

\[
\theta_h^{(n)} = \theta_h^{(p)} + \Delta
\]

Where \( \Delta \) is a 6x1 vector and each element is drawn from \( N(0, 0.05) \), with the accepting probability \( \alpha \) given by:

\[
\alpha = \min \left[ \frac{\exp[-1/2(\theta_h^{(a)} - m_h)'D^{-1}(\theta_h^{(a)} - m_h)]l(\theta_h^{(a)})}{\exp[-1/2(\theta_h^{(p)} - m_h)'D^{-1}(\theta_h^{(p)} - m_h)]l(\theta_h^{(p)})} , 1 \right]
\]

Where

\[
m_{h1} = \gamma_{120} + \gamma_{121} \text{Brand}_h + \gamma_{122} \text{Product}_h
\]

\[
m_{h2} = \gamma_{130} + \gamma_{131} \text{Brand}_h + \gamma_{132} \text{Product}_h
\]

\[
... \quad m_{h6} = \gamma_{340} + \gamma_{341} \text{Brand}_h + \gamma_{342} \text{Product}_h
\]

\( D \) is a diagonal matrix with the diagonal elements equal to \((\delta_{12}^2, \delta_{13}^2, \delta_{14}^2, \delta_{23}^2, \delta_{24}^2, \delta_{34}^2)\)

\[
l(\theta_h) = \prod_{k} P(Z_{hk} = Z_h^b, \sum_{j=1}^{4} Z_{hki} \neq 0)
\]

2. Draw \( \beta_h = (\beta_h^{12}, \beta_h^{13}, \beta_h^{14}, \beta_h^{23}, \beta_h^{24}, \beta_h^{34})' \) similarly as step 1

3. Draw \( \gamma_{j} \) for \( j=1,2,3,4 \) and \( j' > j \)

Define \( x = (1, \text{Brand}, \text{Product}) \) then

\( \gamma_{j} \sim \text{MVN}(A_{j}, B_{j}) \); \( A_{j} = B_{j}(\Sigma_0^{-1} \gamma + x' \theta_{j} / \delta_{j}^2)^{-1}, \quad B_{j} = (\Sigma_0^{-1} + x_j' x_j / \delta_{j}^2)^{-1}, \quad \gamma = 0, \quad \Sigma_0 = 100I \)

4. Draw \( \delta_{j}^2 \) for \( j=1,2,3,4 \) and \( j' > j \)

\( \delta_{j} \sim \text{Inverted Gamma} (A_{j}, B_{j}) \);

\[
A_{j} = s_0 + N / 2 \quad \text{and} \quad B_{j} = \frac{2}{\sum_{h=1}^{N} (\theta_h^{j} - \gamma_{j0} - \gamma_{j0} \text{Brand}_h - \gamma_{j2} \text{Product}_h)^2 / q_0} \quad (q_0 = 1)
\]

5. Draw \( \alpha_{j} \) (\( j=1,2,3,4 \)) similarly as step 3

6. Draw \( \beta_{j} \) (\( j=1,2,3,4 \)) similarly as step 3

7. Draw \( \sigma_{j}^2 \) (\( j=1,2,3,4 \)) similarly as step 4

8. Draw \( \tau_{jm}^2 \) (\( j=1,2,3,4; \; m=1,2,3 \)) similarly as step 4
Figure 1a: Plot showing individual estimates for latency for keywords in categories 1 and 2

Figure 1b: Plot showing individual estimates for latency for keywords in categories 3 and 4
Figure 2: Plots showing the individual estimates for the cross-category interdependence parameters for two representative generic and branded keywords.