Working Paper #08-16

September 2008

A Dynamic Model of Sponsored Search Advertising

Song Yao and Carl F. Mela

Duke University

A Dynamic Model of Sponsored Search Advertising
Preliminary and Incomplete

Song Yao
Carl F. Mela[1]

September 2008

[1]Song Yao (email: song.yao@duke.edu, phone: 919-660-2882) is a Ph.D. Candidate at Duke University, Durham, North Carolina, 27708. Carl F. Mela (email: mela@duke.edu, phone: 919-660-7767, fax: 919-681-6245) is a Professor of Marketing, The Fuqua School of Business, Duke University, Durham, North Carolina, 27708. We gratefully acknowledge financial support from the NET Institute (www.netinst.org) and the Kauffman Foundation.

Abstract

Sponsored search advertising is ascendant—Jupiter Research reports expenditures rose 28% in 2007 to $8.9B and will continue to rise at a 15% CAGR, making it one of the major trends to affect the marketing landscape. Yet little, if any empirical research focuses upon search engine marketing strategy by integrating the behavior of various agents in sponsored search advertising (i.e., searchers, advertisers, and the search engine platform). The dynamic structural model we propose serves as a foundation to explore these and other sponsored search advertising phenomena.

Fitting the model to a proprietary data set provided by an anonymous search engine, we conduct several policy simulations to illustrate the benefits of our approach. First, we explore how information asymmetries between search engines and advertisers can be exploited to enhance platform revenues. This has consequences for the pricing of market intelligence. Second, we assess the effect of allowing advertisers to bid not only on key words, but also by consumers searching histories and demographics thereby creating a more targeted model of advertising. Third, we explore several different auction pricing mechanisms and assess the role of each on engine and advertiser profits and revenues. Finally, we consider the role of consumer search tools such as sorting on consumer and advertiser behavior and engine revenues.

One key finding is that the estimated advertiser value for a click on its sponsored link averages about 24 cents. Given the typical $22 retail price of the software products advertised on the considered search engine, this implies a conversion rate (sales per click) of about 1.1%, well within common estimates of 1-2% (gamedaily.com). Hence our approach appears to yield valid estimates of advertiser click valuations. Another finding is that customers appear to be segmented by their clicking frequency, with frequent clickers placing a greater emphasis on the position of the sponsored advertising link. Estimation of the policy simulations is in progress.
Keywords: Sponsored Search Advertising, Two-sided Market, Dynamic Game, Structural Models, Empirical IO, Customization, Auctions

1 Introduction

1.1 Sponsored Search Marketing

1.1.1 Growth in Sponsored Search

Sponsored search on sites such as `Google`, `Yahoo`, `Sidestep`, `Kayak`, `Bookfinder`, `MSN`, etc. is one of the largest and fastest growing advertising channels. In the United States alone, 2007 annual expenditures on sponsored search advertising increased 28% to $8.9B and the number of firms using sponsored search advertising rose from 29% to 41%.[1] Hence, the tactic is becoming a central component of the promotional mix in many organizations. By contrast, overall 2007 advertising spending across all channels in the United States is estimated to be $283.8B, an increase of only 0.7%.[2]

The growth of this new medium arise in part due to the increasing popularity of search engine sites relative to other media among consumers. In April of 2008, American Internet users conducted 10.6B searches on the 5 leading search engines.[3] By comparison, a top rated TV show such as "Desperate Housewives" only has about 25M viewers (IRI, 2007); and the growing popularity of DVR services offered by TiVo and cable companies have and will further decrease the audience base of traditional TV advertising. Moreover, Qiu et al. (2005) estimate that more than 13.6% of the web traffic is affected by search engines. Since more and more consumers use the Internet for their transactions (Ansari et al. (2008)), Internet search is an especially efficient way to promote online channels. Not only does search advertising have expanding reach, but it often targets consumers who are actively seeking information related to the advertisers' products. For example, a search of "sedan" and "automotive dealer" might signal an active purchase state. As a result of these various factors, Jupiter Research reports that 82% of advertisers were satisfied or extremely satisfied with search marketing ROI and 65% planned to increase search spending in 2007.[4]

Given the increasing ubiquity of sponsored search advertising, the topic has seen increased attention in marketing as of late (Ghose and Yang (2007); Rutz (2007); Rutz and Bucklin (2007); Goldfarb and Tucker (2007)). Yet most empirical work on the topic remains focused on the advertiser. To date, empirical research on key word search has largely ignored the perspective of the search engine. Given that the search engine interacts with advertisers to determine the price of the advertising (and hence its efficacy), our objective is to broaden this stream of research to incorporate the role of the search engine and its users. This exercise enables us to determine the role of search engine marketing strategy on the behavior of advertisers and consumers as well as the attendant implications for search engine revenues.

[1] "US Paid Search Forecast, 2007 to 2012", Jupiter Research, 2007.
[2] "Insider's Report", 2007, McCann WorldGroup, Inc..
[3] "April 2008 U.S. Search Engine Rankings," comScore, Inc. (`http://www.comscore.com/press/release.asp?press=2230`).
[4] "US Paid Search Forecast, 2007 to 2012", Jupiter Research, 2007.

Our key contributions are as follows. From a theoretical perspective, we conceptualize and develop an integrated model of web searcher, advertiser and search engine behavior. To our knowledge, this is the first empirical paper focusing on the marketing strategy of the search engine. From a substantive point of view, our contribution is to offer concrete marketing policy recommendations to the search engine including its i) pricing (for both the key words and the clickstream data it collects), ii) key word auction design (such as the pricing mechanism and whether advertiser bidding should be targeted by segment as well as key word), and iii) web page design (e.g., should features like sort or filter be added or dropped). From a methodological view, we develop a dynamic structural model of key word advertising. The dynamic aspect of the problem requires the use of some recent innovations pertaining to the estimation of dynamic games in economics (e.g., Bajari et al. (2007), Pesendorfer and Schmidt-Dengler (2008)). We extend this work to be Bayesian in implementation and apply it to wholly new context.

One notable finding is that advertisers in our application have an average value per click of \$0.24. Given the average price of software products advertised on this site is about \$22, this implies these advertisers expect about 1.1% (i.e., \$0.24/\$22) of clicks will lead to a purchase. This is consistent with the industry average of 1-2% reported by GameDaily.com, suggesting good face validity for our model. In addition, we find considerable heterogeneity in consumer response to sponsored search advertising. Frequent link clickers, who represent 10% of the population but 90% of the clicks tend to be more sensitive to slot order – in part because slot position can signal product quality. These insights represent central inputs into our yet to be completed policy simulations alluded to above.

1.1.2 Sponsored Search Advertising

The Internet contains an estimated 155 million sites and Internet search engines wade through this information to return relevant results in response to users' search queries.[5] These "organic" search results are often displayed as a list of links sorted by their relevance to the search query (Bradlow and Schmittlein (2000)). Search engines range from the quite general type (e.g., `Google.com` searches encompass most of the Internet) to the more focused ones (`DealTime.com` searches Internet stores, `hotels.com` searches travel products, `www.addall.com` searches books, etc..). Sponsored search involves advertisements placed above or along side the organic search results (See Figure 1 and Figure 2). Given that users are inclined to view the topmost slots in the page (Ansari and Mela (2003)), advertisers are willing to pay a premium for these more prominent slots (Goldfarb and Tucker (2007)).

To capitalize on this premium, advertising slots are auctioned off by search engines. Advertisers specify bids on a per-click basis for a respective search term. Advertisers consider potential competition, the cost of bidding, and the expected revenue accruing from the advertisement when

[5] "January 2008 Web Server Survey," Netcraft Company (`http://news.netcraft.com/archives/2008/01/28/january_2008_web_server_survey.html`).

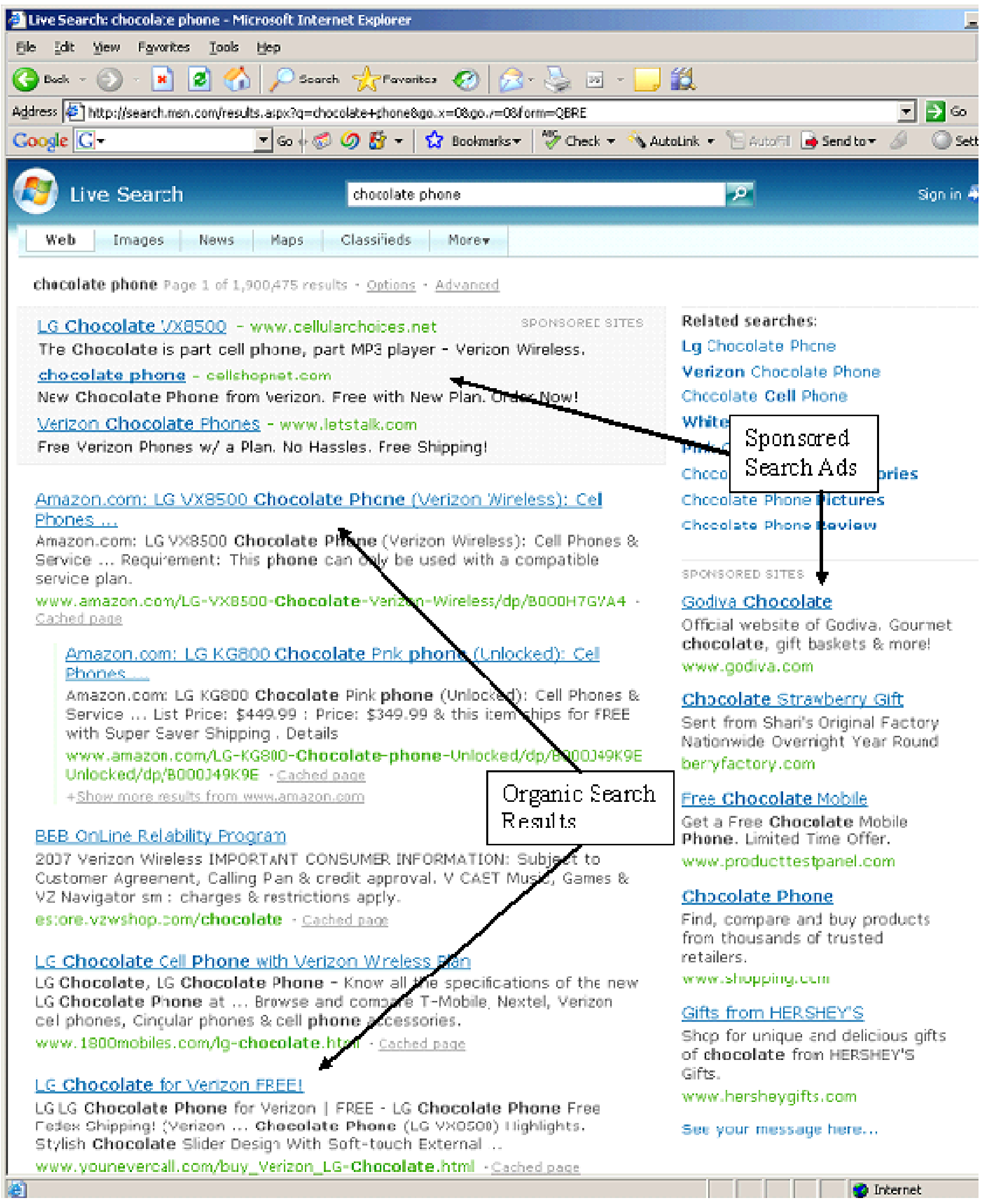

Figure 1: Searching "chocolate phone" Using A Generic Search Engine

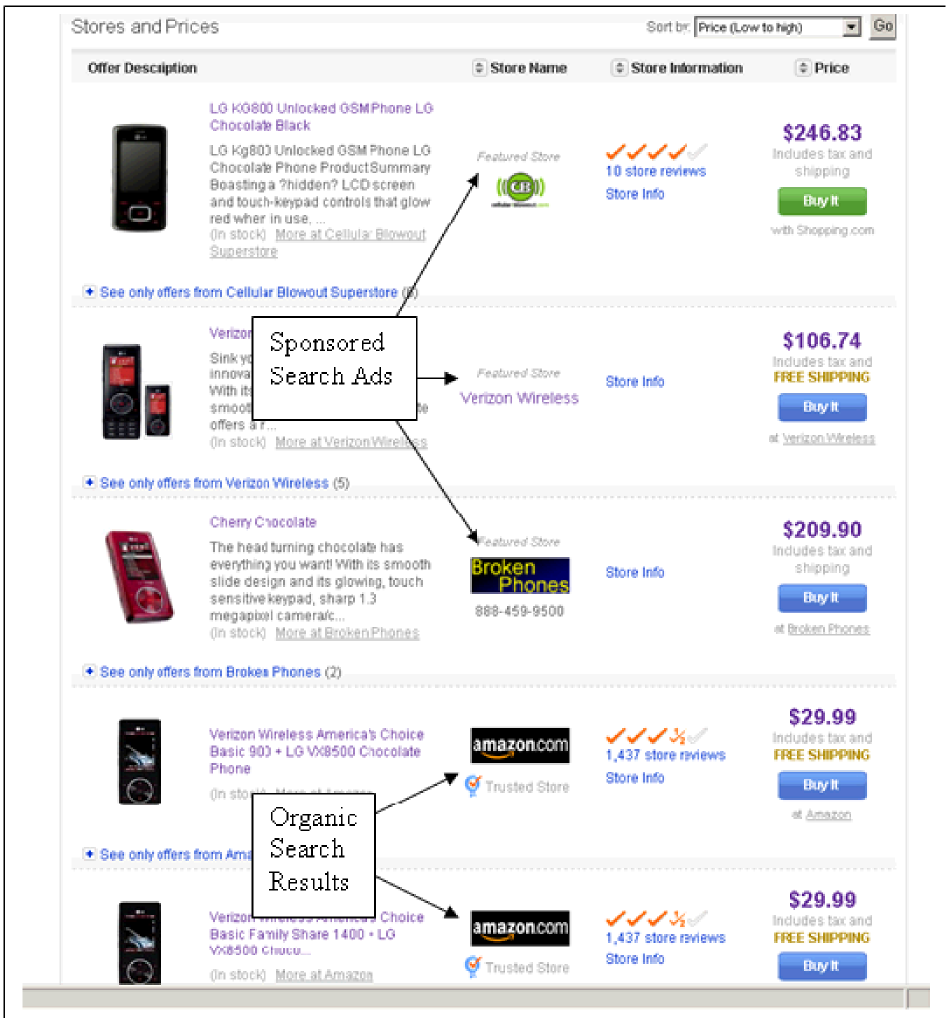

Figure 2: Searching "chocolate phone" Using A Specialized Search Engine

deciding how to bid. Though most search engines use auctions to price advertisements, there is considerable variation in the nature of the auctions they use. For example `Overture.com` (who pioneered Internet search auctions and is now a part of `Yahoo!`) adopted a first price auction wherein the advertiser bidding the highest amount per click received the most prominent placement at the cost of its own bid for each click.[6] First price auctions are still used by `Shopping.com` and a number of other Internet properties. However, because bids are priced on a per click basis, search engines should be not only concerned with the magnitude of bid but also the likely number of clicks that an advertisement will generate. For example, a high per-click bid with few or no clicks may generate less revenue for a search engine than a low per-click bid advertisement with a high click-through rate. Hence `Google` has developed an algorithm which factors in not only the level of the bid, but the expected click-through rate of the advertiser. Another distinction of the `Google` practice is that advertisers pay the next bidder's bid (adjusted for click-through rates) as opposed to their own bids.[7] Moreover, `Google` and `MSN` recently enabled advertisers to bid by demographics or the browsing history of the users, thus enabling even more precise targeting.

In light of the increased use of search engines by consumers, the attendant rise in search engine advertising, and the resulting interest in pricing mechanisms on the part of the search engine, we model the behavior of consumers and advertisers in order to obtain insights into the policy of the search engine platform. Much like Yao and Mela (2007), we construct an *empirical* model of a two-sided network in an auction context. One side of the network includes the searchers who generate revenue for the advertiser. On the other side of the two-sided network are advertisers whose bidding behavior determines the revenue of the search engine. In the middle lies the search engine. The goal of the search engine is to price consumer information, set auction mechanisms, design web page to elucidate product information so as to maximize its profits. By integrating these agents in a single model, it becomes possible to explore the effect of search engine strategy on the demand and pricing for search engine advertising as well as the revenues of the search engine. In particular, we consider the following policy simulations:

- Mechanism Design. The wide array of search pricing mechanisms raises the question of which auction mechanism is the best in the sense of incenting advertisers to bid more aggressively thereby yielding maximum returns for the search engine. We contrast the two most common designs and their attendant revenue implications.

[6] In the economics literature, such an auction with multiple items (slots) where bidders pay what they bid is sometimes termed as discriminatory auction (Krishna (2002)).

[7] With a simplified setting, Edelman et al. (2007) show that the Google practice may result in an equilibrium with bidders' payoffs equivalent to the Vickrey-Clarke-Groves (VCG) auction, whereas VCG auction has been proved to be maximizing total payoffs to bidders (Groves (1979)). Iyengar and Kumar (2006) further show that under some conditions the Google practice induces VCG auction's dominant "truth-telling" bidding strategy, i.e., bidders will bid their own valuations.

- Market Intelligence. Advertisers' knowledge about consumers changes if search engines sell consumer demographic and behavioral information to advertisers. The bidding strategy of advertisers is likely altered by the change of their information state. This raises the question of how information asymmetries between the engine and advertisers affect bidding behavior and how the consumer information should be priced.

- Customization and Targeting. Most search engines auction key words across all market segments. However, it is possible to auction key words by segment. We assess the potential revenue implications of this strategy.

- Search Tools. Many search engines, especially specialized ones such as `Shopping.com`, provide options to sort/filter search results using certain criteria such as product prices. On one hand, sort/filter may intensify competition among advertisers by mitigating the perceived difference across goods (Diehl et al. (2003)). On the other hand, these tools can induce consumers to focus on quality differentiation thereby attenuating the competition (Lynch and Ariely (2000)). This leads to the question of how such an easy access to products information impacts consumer searching behavior and hence firms' advertising decisions.

Though we cast our model in the context of sponsored search, we note that the problem, and hence the conceptualization is even more general. Any interactive, addressable media format (e.g., DVR, satellite digital radio) can be utilized to implement similar auctions for advertising. For example, with the convergence in media between computers and television in DVRs, simple channel or show queries can be accompanied by sponsored search and this medium may help to offset advertising losses arising from ads skipping by DVR users. In such a notion, the research literature on sponsored search auctions generalizes to a much broader context and our model serves as a basis for exploring search based advertising.

1.2 Recent Literature

Research on sponsored search, commensurate with the topic it seeks to address, is nascent and growing. This literature can be characterized along two distinct dimensions: theoretical and empirical. The theoretical literature details how agents are likely to react to different pricing mechanisms. One major conclusion of this literature is that the optimal pricing mechanism is incumbent upon the behavior of the various agents. However, there is little attention directed to the issue of asymmetries in information states between the advertiser and the platform. Moreover, the theoretical literature does not measure how agents actually behave in a given market so it can not speak to how changes in platform marketing strategy will manifest in a given market. In contrast, the empirical literature measures the effect of advertising on consumer response in a given market but not the reaction of these agents to changes in the platform environment (e.g., advertising pricing, information state or the webpage design of the platform). Further, the empirical literature typically focuses on the behavior of the advertiser but not that of the searchers or the marketing actions of the search platform. Next, empirical work to date is largely silent on competitive interactions between various advertisers. In sum, by integrating the theoretical and empirical research streams we seek to develop a more complete representation of the role of pricing and information in the context of key word search. To elaborate on these points, we begin by surveying theoretical work on sponsored search and then proceed to discuss some recent empirical research.

Foundational theoretical analyses of sponsored search include Edelman et al. (2007) and Varian (2006) who examine the bidding behaviors of advertisers in this auction game. The authors assume the auction game as a complete information and simultaneous-move static game, in which exogenous advertising click through rates increase with better placements. In equilibrium advertiser bidding behavior has the same payoff structure as a Vickrey-Clarke-Groves auction, where a winner's payment to the seller equals to those losing bidders' potential payoffs (opportunity costs) were the winner absent (Groves (1979)). Extending this work, Chen and He (2006) incorporate clicking behavior into their model and show that, under the `Google` bidding mechanism, consumers clicking behavior is affected by the easy access to product information. In particular, they make inference about product quality based on the ranking presented by the platform and search sequentially according to the ranking. As an equilibrium response advertisers submit bids equal to their true values for the advertising. Katona (2007) further extends the analysis by relaxing several key assumptions such as the competition for traffic between sponsored links and organic links, the heterogeneity of advertisers in term of their inherent attractiveness to consumers. The author shows multiple equilibria in this auction which do not have closed form solutions. Additional work by Iyengar and Kumar (2006), Feng (2008), and Garg et al. (2006) explicitly consider the effect of the various auction mechanisms on search engine profits. In particular, Iyengar and Kumar (2006) show that the `Google` pricing mechanism maximizes neither the search engine's revenue nor the efficiency of the auction suggesting the potential to improve on this mechanism as we seek to do. Further, they show that the optimal mechanism is incumbent upon the characteristics of the market thereby making it imperative to estimate market response as we intend to do in order to improve on pricing mechanisms.

In sum, the key insights from this stream of work are that i) there are three key sets of agents interacting in the sponsored search context, persons that engage in key word search, advertisers that bid for key words, and the search platform, ii) one can characterize how advertisers and searchers will react to changes in the auction mechanisms employed by the search engine, iii) searchers will react to the search engine's web page design, which in turn will affect advertisers bidding behavior and iv) changes in advertiser behavior are incumbent upon the parameters of the system; given these are not estimated it is hard to characterize precisely how these agents will behave. Additionally, we note that the oft invoked assumption of a static advertiser game over bidding periods is inconsistent with the pricing practices used by search engines. Search engines typically use the preceding period's click throughs together with current bids to determine advertising placement, making this an inherently dynamic game. Finally, this research typically assumes no asymmetry in information states between the advertiser and the search engine even though the search engine knows individual level clicking behaviors and the advertiser does not. We redress these issues in this paper.

Empirical research on sponsored search advertising is also proliferating. Notable among these papers, Rutz and Bucklin (2007) investigate the efficacy of different keyword choices by measuring the conversion rate from users' clicks on ads to actual sales for the advertiser. In a related paper, Rutz (2007) considers how advertiser revenue is affected by click throughs and exposures. This work is important because it demonstrates that advertiser valuations differ for various placements and key words and that the bids are likely to be related to placements. Ghose and Yang (2007) further investigate the relationships among different metrics such as click-through rate, conversion rate, bid price and advertisement rank.

Overall, the empirical research on sponsored search establishes a firm link between advertising, slot position and revenues – and indicates that these effects can differ across advertisers. Yet some limitations of this stream of work include its emphasis on a single agent (the advertiser) and the lack of information on competing bidders, which make it difficult to predict how advertisers might react to a change in the auction mechanism, webpage design or information state regarding consumers. Yet these interactions are material to understanding the role of each agent in the context of sponsored search. For example, an advertiser's value to the search engine pertains not only to its payment to the search engine as is often assumed in past empirical work, but also the effect that advertiser has on the intensity of competition during bidding. The increased intensity of competition may serve to drive bids upward and hence increase search engine revenues. Related, advertisers' actions affect internet users. For example, with alternative advertisers being placed at premium slots on a search result page, it is likely that users' browsing behaviors will be different. Further, since advertisers make decisions with the consideration of users' reactions, any variations of users' behaviors have feedbacks on advertisers' actions and thus will ultimately affect the search engine's revenue. Hence when making policy prescriptions for the search engine, we believe that it would be more reasonable to incorporate the theoretical work on strategic interaction in the context of key word search into an empirical analysis of advertiser bidding behavior.

This suggests it is desirable to model and estimate the equilibrium behavior of all the agents in a network setting. In this regard, sponsored search advertising can be characterized as a two-sided market wherein searchers and advertisers interact on the platform of the search engine (Rochet and Tirole (2006); Tucker (2005)). This enables us to generalize a structural modeling approach advanced by Yao and Mela (2007) to study two-sided markets. These authors model bidder and seller behavior in the context of electronic auctions to explore the effect of auction house pricing on the equilibrium number of listings and closing prices. However, additional complexities exist in the key word search setting including i) the aforementioned information asymmetry between advertisers and the search engine and ii) the substantially more complex auction pricing mechanism used by search engines relative to the fixed fee auction house pricing considered in Yao and Mela (2007). Moreover, unlike the pricing problem addressed in Yao and Mela (2007), sponsored search bidding is inherently dynamic owing to the use of lagged advertising click rates to determine current period advertising placements. Hence we incorporate the growing literature of two-step dynamic game estimation (e.g., Hotz and Miller (1993); Bajari et al. (2007); Bajari and Hong (2006)). Instead of explicitly solving for the equilibrium dynamic bidding strategies, the two-step estimation approach assumes that observed bids are generated by equilibrium play and then use the distribution of bids to infer underlying primitive variables of bidders (e.g., the advertiser's expectation about the return from advertising). Similar method is also used in an auction context in Jofre-Bonet and Pesendorfer (2003). However, our approach is unique inasmuch as it is Bayesian instantiation of these estimators, which leads to desirable small sample properties and enables considerable flexibility in modeling choices. Equipped with these advertiser primitives, we solve the dynamic game played by the advertiser to ascertain how changes in search engine policy affect equilibrium bidding behavior.

In sum, our goal is to develop an integrated model of key word search that incorporates the behavior of both searchers and advertisers. This approach enables us to investigate how the policies of the search engine affect its revenues. Such policies include the marketing of information, targeted bidding, pricing mechanisms and webpage design among others. This goal mandates the use of a dynamic structural model of key word search and, to our knowledge, this paper is the first to integrate empirical and theoretical work on key word search to develop such an approach and to provide some explicit prescriptions for the marketing policies of search engines.

The remainder of this paper proceeds as follows. Given the relatively novel research context, we begin by describing the data to help make the problem more concrete. We then outline the details of our model beginning with the clicking behavior of consumers and concluding with the advertiser bidding behavior. Next, we turn to estimation and present our results. We then explore the role of information asymmetry, targeted bidding, advertising pricing and webpage design by developing policy simulations which alter the search engine marketing strategies. We conclude with some future directions.

2 Empirical Context

The data underpinning our analysis is drawn from a major search engine for high technology consumer products. Within this broad search domain, we consider search for music management software because the category is relatively isolated in the sense that searches for this product do not compete with others on the site. The category is a sizable one as well for this search engine. Along with the increasing popularity of MP3 players, the use of music management PC software is increasing exponentially making this an important source of revenue. The goal of the search engine is to enable consumers to identify and then download trial versions of these software before their final purchase.[8] It should be noted that the search engine defines the music management broadly enough that an array of different search terms (e.g., MP3, iTunes, iPod, lyric, etc.) yield the same search results for the software products in this category. Hence we consider the consumer decision of whether to search for music software on the site and whether to download given a search.

Because consumers are far more likely to click on links near the top of the search results page, advertisers compete for these slots by attending the auction.[9] More successful bids lead to appearances closer to the top of the list. Winning bids are denoted as *sponsored search results* and the site flags these as sponsored links. The site affords up to five premium slots which is far less than the 400 or so products that would appear at the search engine. Losing bidders and non-bidders are listed beneath the top slots on the page and like previous literature we denote these listings as *organic search results*. We seek to model this bidding behavior.

The search engine captures data on advertisers (products attributes, products download history and bids from active bidders), consumers (their visitation log files and demographics), and relevant site characteristics from the search engine platform (such as page characteristics and link order). We detail these data next.

2.1 Data Description

The data are comprised of 3 files, including:

- Bidding file. Bidding is logged into a file containing the bidding history of all active bidders from January 2005 till August 2007. It records the exact bids submitted, the time of each bid submission and the resulting monthly allocation of slots. Hence, the unit of analysis is vendor-bid event. These data form the cornerstone of our bidding model.

- Products file. Product attributes are kept in a file that records, for each software firm in each month, the characteristics of the software they purvey. This file also indicates the download history of each product in each month.

- Consumers file. Consumer log files record each visit to the site and is used to infer whether downloads occur as well as browsing histories. A separate but related file includes registration

[8] A "click" and a "download" are essentially the same from the perspectives of the advertiser, the consumer as well as the search engine. In the "click" case, a consumers makes several clicks to investigate and compare products offered by different vendors and then make a final purchase. In the "download" case, a consumer downloads several products and makes the comparison before a final purchase. Hence there is no difference for a "click" and a "download" in the current context. We use "click" and "download" interchangeably throughout the paper.

[9] We detail the specific rules of the bidding process when describing the bidder model in section 3.2.

information and detailed demographics for those site visitors that are registered. These data are central to the bidding model in the context of complete information.

We detail each of these files in turn.

2.1.1 Bidding File

Table 1 reports summary statistics for the bidding files. At this search engine, bids were submitted on a monthly basis. Over the 32 months from January 2005 to August 2007, 322 bids (including zeros) were submitted by 21 software companies.[10] As indicated in Table 1, bidders on average submitted about 22 positive bids in this interval (slightly less than once per month). The average bid amount (conditioned on bidding) was \$0.20 with a large variance across bidders and time.

Table 1: Bids Summary Statistics

	Mean	Std. Dev.	Minimum	Maximum
Non-zero Bids (¢)	19.55	8.32	15	55
Non-zero Bids/Bidder	21.78	10.46	1	30
All Bids (¢)	8.14	11.04	0	55
Bids/Bidder	23.13	9.68	1	32

2.1.2 Product File

Searching for a key word results in a list of relevant software products and their respective attributes (which may vary over time). Attribute information is stored in a product file along with the download history of all products that appeared in this category from January 2005 to August 2007. In total, these data cover 394 products over 32 months. The attributes include the price of the non-trial version of a product, backward compatibility with preceding operating systems (e.g., Windows 98 and Windows Server 2003), expert ratings provided by the site and consumer ratings of the product.[11] Trial versions typically come with a 30-day license to use the product for free, after which consumers are expected to pay for its use. Expert ratings at the site are collected from several industrial experts of these products. The consumer rating is based on the average feedback score about the product from consumers. Tables 2 and 3 give summary statistics for all products as well as active bidders' products. Based on the compatibility information, we sum each product's compatibility dummies and define this summation as a measure for that product's compatibility with old and unpopular OS. This variable is later used in our estimation.

Overall, active bidders' products have higher prices, better ratings and more frequent updates.

[10] Since some products were launched after January 2005, they were not observed in all periods.

[11] We further considered file size but found many missing values. Moreover, in light of increased Internet speed, file size has become somewhat inconsequential in the download decision and is omitted from our analysis.

Table 2: Product Compatibility

	Percentage
All Products	
Windows NT 4.0	54
Windows 98	64
Windows Me	66
Windows 2000	91
Windows Server 2003	43
Bidders' Products	
Windows NT 4.0	67
Windows 98	67
Windows Me	71
Windows 2000	85
Windows Server 2003	57

Table 3: Product Attributes and Downloads

	Mean	Std. Dev.	Minimum	Maximum
All Products				
Non-trial Version Price $	16.65	20.43	0	150
Expert Rating (if rated)	3.87	0.81	2	5
Average Consumer Rating (if rated)	3.89	1.31	1	5
Months Lapse Since Last Update	15.31	9.88	1	31
Compatibility Index	3.29	1.47	0	5
Number of Downloads/(Product×Month)	1367.29	9257.16	0	184442
Bidders' Products				
Non-trial Version Price $	21.97	15.87	0	39.95
Expert Rating (if rated)	4	0.50	3	5
Average Consumer Rating (if rated)	4.06	0.91	2.5	5
Months Lapse Since Last Update	2.38	0.66	1	3
Compatibility Index	3.51	1.51	0	5
Number of Downloads/(Product×Month)	1992.12	6557.43	0	103454

2.1.3 Consumer File

This file contains the log file of consumers from May 2007 till August 2007. The consumers file contains each consumer's browsing log when they visit the search engine. It also has the registration information if a consumer has registered with the search engine before.

The browsing log of a consumer records the entry time, browsing path and duration of the visit. It also indicates whether the consumer made downloads and, if yes, which products she downloaded. Upon viewing the search results of software products, the search engine allowed the consumer to sort the results based on some attributes such as the ratings; consumers can also filter products based on some criteria such as whether a product's non-trial version is free. The browsing log records the sorting and filtering actions of each consumer.

Since the demographic information upon the registration is only optional, the dataset provides little if any reliable demographics of consumers. So we only focus on whether a consumer is a registered user of the search engine.

3 Model

The model must capture the behaviors of the two key agents interacting on the search engine platform: i) advertisers who bid to maximize their respective profits and ii) utility maximizing consumers who decide whether to click on the advertiser's link. For any given policy applied by the search engine, this integrated model enables us to predict equilibrium revenues for the search engine. Recognizing that the behavior of the bidder (advertiser) is conditional on the behavior of the consumer, we begin with the consumer model and then solve the bidder problem conditional on the consumer behavior.

3.1 Consumer Model

Advertiser profit (and therefore bidding strategy) is incumbent upon their forecast of consumer downloads for their products $d(k, X_j^t; \Omega_c)$, where k denotes the position of the advertisement on the search engine results page, X_j^t indicate the attributes of the advertiser j's product at time t and Ω_c are parameters to be estimated. Thus, we seek to develop a forecast for $d(k, X_j^t; \Omega_c)$ and the attendant consequences for bidding. To be consistent with the advertisers information set, we begin by basing these forecasts of consumer behavior solely on statistics observed by the advertiser: the aggregate download data and the distribution of consumers characteristics. Later, in the policy section of the paper, we assess what happens to bidding behavior and platform revenues when disaggregate information is revealed to advertisers by the platform. We begin by describing the consumer's download decision process and how it affects the overall number of downloads.

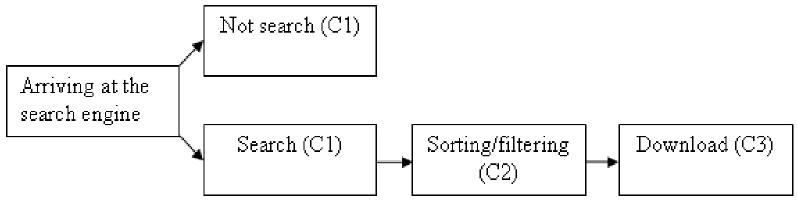

Figure 3: Consumer Decisions

3.1.1 The Consumer Decision Process

Figure 3 overviews the decisions made by consumers. In any given period t, the consumer's problem is whether and which software to select in order to maximize their utility. The resolution of this problem is addressed by a series of conditional decisions.

First, the consumer decides whether she should use search on the category considered in this analysis (C1). We presume that the consumer will search if it maximizes her expected utility to do so.

Conditioned upon engaging a search, the consumer next decides whether to sort and/or filter the results (C2). Sorting re-orders the search results by a specified criterion such as the rating of a software. Filtering excludes various products from consideration based on the product attributes (e.g., the price of the software). The two search options lead to the following 4 options for viewing the results: κ ={0 ≡neither, 1 ≡sorting but not filtering, 2 ≡not sorting but filtering, 3 ≡sorting and filtering}. For each option, the set of products returned by the search engine differs in terms of the number and the order of products. Consumers choose the sorting/filtering option that maximizes their expected utility.

Third, the consumer then chooses which, if any products to download (C3). We presume that consumers choose to download software if it maximizes their expected utility. We discuss the modeling details for this process in a backward induction manner (C3–C1).

Download We assume that consumers exhibit heterogeneous preferences for the products and these consumers choose products to download to maximize their expected payoffs. Consumer i of preference segment g ($g = 1, 2, ..., G$) has some underlying latent utility $u_{ijt}^{g\kappa}$ for downloading software j in period t, conditional on her sorting/filtering choice κ. A product will be downloaded if and only if $u_{ijt}^{g\kappa} \geq 0$. Let a index product attributes

$$u_{ijt}^{g\kappa} = \widetilde{\alpha}_j^g + \sum\nolimits_a x_{jat}^{\kappa} \widetilde{\beta}_a^g + \widetilde{\varepsilon}_{ijt}^{g\kappa} \tag{1}$$

where

- x^{κ}_{jat} is the observed attribute a of product j; product attributes also includes product j's slot k on the search page that may vary conditional on sorting/filtering choice κ (hence the superscript κ);
- $\widetilde{\beta}^{g}_{a}$ is consumer i's "taste" regarding product attribute a, which is segment specific;
- $\widetilde{\varepsilon}^{g\kappa}_{ijt}$'s are individual idiosyncratic preference shocks, realized *after* the sorting/filtering decision. They are independently and identically distributed over individuals, products and periods as zero mean normal random variables.

To allow the variance of the download ($\widetilde{\varepsilon}^{g\kappa}_{ijt}$) and sorting/filtering errors ($\xi^{g\kappa}_{it}$, which will be detailed below) to differ, both must be properly scaled (cf., Train (2003), Chapter 2). Hence we invoke the following assumption.

Assumption 1: $\widetilde{\varepsilon}^{g\kappa}_{ijt}$'s are independently and identically distributed normal random variables with mean 0 and variance normalized to $(\delta^g)^2 \cdot \pi^2/6$. $\xi^{g\kappa}_{it}$'s are independently and identically distributed Type I extreme value random variables with variance normalized to $\pi^2/6$.

Under assumption 1, we may re-define the utility in equation 1 as

$$u^{g\kappa}_{ijt} = \delta^g \pi/\sqrt{6}(\underbrace{\alpha^g_j + \sum\nolimits_a x^{\kappa}_{jat}\beta^g_a}_{\overline{u}^{g\kappa}_{ijt}} + \varepsilon^{g\kappa}_{ijt}) \tag{2}$$

where $\{\alpha^g_j, \beta^g_a, \varepsilon^{g\kappa}_{ijt}\} = \{\widetilde{\alpha}^g_j, \widetilde{\beta}^g_a, \widetilde{\varepsilon}^{g\kappa}_{ijt}\}/(\delta^g\pi/\sqrt{6})$; $\overline{u}^{g\kappa}_{ijt}$ is the scaled "mean" utility and $\varepsilon^{g\kappa}_{ijt} \sim N(0,1)$. The resulting choice process is a multivariate probit choice model.[12] Let $d_{ijt} = 1$ stand for downloading and $d_{ijt} = 0$ stand for not downloading. We have

$$d_{ijt} = \begin{cases} 1 & \text{if } u^{g\kappa}_{ijt} \geq 0 \\ 0 & \text{otherwise} \end{cases} \tag{3}$$

and the probability of downloading conditional on parameters $\{\alpha^g_j, \beta^g_a\}$ is

$$\begin{aligned} \Pr(d_{ijt} = 1) &= \Pr(u^{g\kappa}_{ijt} \geq 0) \\ &= \Pr(\delta^g\pi/\sqrt{6}(\overline{u}^{g\kappa}_{ijt} + \varepsilon^{g\kappa}_{ijt}) \geq 0) \\ &= \Pr(-\varepsilon^{g\kappa}_{ijt} \leq \overline{u}^{g\kappa}_{ijt}) \\ &= \Phi(\overline{u}^{g\kappa}_{ijt}) \end{aligned} \tag{4}$$

where $\Phi(\cdot)$ is the standard normal distribution CDF.

[12] It can be shown that, under very weak assumptions, download decisions across multiple products with the purpose of maximizing total expected utility can be represented by a multivariate binary choice probit model.

Sorting and Filtering Prior to making a download decision, consumers face several filtering and sorting options which we index as $\kappa = 0, 1, 2, 3$. We expect consumers to choose the option that maximizes their expected download utility. Although consumers know the distribution of the product utility error terms $(\tilde{\varepsilon}_{ijt}^{g\kappa})$, these error terms do not realize before the sorting/filtering. Hence consumers can only form an expectation about the total utilities of all products under a given sorting/filtering option κ before choosing that option. Let $U_{it}^{g\kappa}$ denote the total expected utility from products under option κ, which can be calculated based on equation 1:

$$U_{it}^{g\kappa} = \sum_j E_\varepsilon(u_{ijt}^{g\kappa} | u_{ijt}^{g\kappa} \geq 0) \Pr(u_{ijt}^{g\kappa} \geq 0) \tag{5}$$

This definition reflects that a product's utility is realized only when it is downloaded. Hence, the expected utility $E_\varepsilon(u_{ijt}^{g\kappa} | u_{ijt}^{g\kappa} \geq 0)$ is weighted by the download likelihood, $\Pr(u_{ijt}^{g\kappa} \geq 0)$. The expectation, $E_\varepsilon(\cdot)$, is taken over the random preference shocks $\varepsilon_{ijt}^{g\kappa}$.

In addition to $U_{it}^{g\kappa}$, individuals may accrue additional benefits or costs for using sorting/filtering option κ. These benefits or costs may arise from individual differences of efficiency or experience in terms of engaging the various options for ordering products. We denote such benefits or costs by random terms $\xi_{it}^{g\kappa}$'s. As indicated in assumption 1, $\xi_{it}^{g\kappa}$'s are i.i.d. Type I extreme value. $\xi_{it}^{g\kappa}$ is not observed by researchers but known to individual i. Note that these sorting/filtering benefits or costs do not materialize during the consumption of the products. Therefore they do not enter the latent utility in equation (1). The total utility of search option κ is thus given by

$$z_{it}^{g\kappa} = U_{it}^{g\kappa} + \xi_{it}^{g\kappa} \tag{6}$$

Consumers choose the option of sorting/filtering that leads to the highest total utility $z_{it}^{g\kappa}$.

With $\xi_{it}^{g\kappa}$ following a Type I extreme value distribution, the choice of sorting/filtering becomes a logit model such that

$$\Pr(\kappa)_{it}^{g} = \frac{\exp(U_{it}^{g\kappa})}{\sum_{\kappa'=0}^{3} \exp(U_{it}^{g\kappa'})} \tag{7}$$

To better appreciate the properties of this model, note that $U_{it}^{g\kappa}$ in equation 5 can be written in a closed form.[13]

$$\begin{aligned} U_{it}^{g\kappa} &= \sum_j E_\varepsilon(u_{ijt}^{g\kappa}|u_{ijt}^{g\kappa} \geq 0) \cdot \Pr(u_{ijt}^{g\kappa} \geq 0) \\ &= \delta^g \pi/\sqrt{6} \sum_j \left(\overline{u}_{ijt}^{g\kappa} + \frac{\phi(\overline{u}_{ijt}^{g\kappa})}{\Phi(\overline{u}_{ijt}^{g\kappa})} \right) \cdot \Phi(\overline{u}_{ijt}^{g\kappa}) \end{aligned} \tag{8}$$

With such a formulation, the factors driving the person's choice of filtering or sorting become more apparent:

- Filtering eliminates options with negative utility, such as highly priced products (because consumer price sensitivity is negative). As a result, the summation in 8 for the filter option will increase as the negative $\overline{u}_{ijt}^{g\kappa}$ are removed. This raises the value of the filter option suggesting that price sensitive people are more likely to filter on price.
- Sorting re-orders products by their attribute levels. Products that appear low on a page will typically have lower utility regardless of their product content (because consumer slot rank sensitivity is negative). For example, suppose a consumer relies more on product ratings. By moving more desirable items that have high ratings up the list, sorting can increase the $\overline{u}_{ijt}^{g\kappa}$ for these items, thereby increasing the resulting summation in 8 and the value of this sorting option.[14]

Keyword Search The conditional probability of keyword search takes the form

$$\Pr(search_i^g) = \frac{\exp(\lambda_0^g + \lambda_1^g IV_{it}^g)}{1 + \exp(\lambda_0^g + \lambda_1^g IV_{it}^g)} \tag{9}$$

[13] For a normal random variable x with mean μ, standard deviation σ and left truncated at a (Greene (2003)), $E(x|x \geq a) = \mu + \sigma\lambda(\frac{a-\mu}{\sigma})$, where $\lambda(\frac{a-\mu}{\sigma})$ is the hazard function such that $\lambda(\frac{a-\mu}{\sigma}) = \frac{\phi(\frac{a-\mu}{\sigma})}{1-\Phi(\frac{a-\mu}{\sigma})}$.

Hence with $u_{ijt}^{g\kappa} \sim N(\delta^g\pi/\sqrt{6}\overline{u}_{ijt}^{g\kappa}, (\delta^g)^2\pi^2/6)$, we have

$$\begin{aligned} & E(u_{ijt}^{g\kappa}|u_{ijt}^{g\kappa} \geq 0) \\ = \; & (\delta^g\pi/\sqrt{6} \cdot \overline{u}_{ijt}^{g\kappa} + \delta^g\pi/\sqrt{6} \cdot \frac{\phi(-\frac{\delta^g\pi/\sqrt{6}\cdot\overline{u}_{ijt}^{g\kappa}}{\delta^g\pi/\sqrt{6}})}{1 - \Phi(-\frac{\delta^g\pi/\sqrt{6}\cdot\overline{u}_{ijt}^{g\kappa}}{\delta^g\pi/\sqrt{6}})}) \\ = \; & \delta^g\pi/\sqrt{6}(\overline{u}_{ij}^{g\kappa} + \frac{\phi(\overline{u}_{ij}^{g\kappa})}{\Phi(\overline{u}_{ij}^{g\kappa})}) \end{aligned}$$

[14] In particular, in the data over 80% consumers who used sorting option chose ratings to re-order products. Thus, we suspect consumers who rely on ratings are more likely to use the sorting option to see which items are the most popular ones.

where IV_i^g is the inclusive value for searching conditional on the segment membership. IV_{it}^g is defined as

$$IV_{it}^g = \log[\sum\nolimits_\kappa \exp(U_{it}^{g\kappa})] \tag{10}$$

This specification can be interpreted as the consumer making decision to use a key word search based on the rational behavior of utility maximization (cf. McFadden (1977); Ben-Akiva and Lerman (1985)).[15] A search term is more likely to be invoked if it yields higher expected utility. In our data, we focus on a single search term.

3.1.2 Segment Membership

Recognizing that consumers are heterogeneous in their behaviors described above, we apply a latent class model in the spirit of Kamakura and Russell (1989) to capture heterogeneity in consumer preferences. Heterogeneity in preference can arise, for example, when some consumers prefer some features more than others. We assume G exogenously determined segments. Consumer decisions vary across segments. Consumers are homogeneous within the same segment. Segment-specific heterogeneity is stable across time.[16]

The prior probability for user i being a member of segment g is defined as

$$pg_{it}^g = \exp\left(\gamma_0^g + Demo_{it}'\gamma^g\right) / \Sigma_{g'=1}^G \exp\left(\gamma_0^{g'} + Demo_{it}'\gamma^{g'}\right) \tag{11}$$

where $Demo_{it}'$ is a vector of attributes of user i such as demographics and past browsing history; vector $\{\gamma_0^g, \gamma^g\}_{\forall g}$ contains parameters to be estimated. For the purpose of identification, one segment's parameters are normalized to zero.

In light of the foregoing model, the probability of user i downloading product j in period t is

$$\begin{aligned} & P_{ijt} \\ = \; & \int_{Demo_{it}} \sum\nolimits_g \sum\nolimits_\kappa [\Phi(\overline{u}_{ijt}^{g\kappa}) \frac{\exp(U_{it}^{g\kappa})}{\sum\limits_{\kappa'=0}^{3} \exp(U_{it}^{g\kappa'})}] \Pr(search_{it}^g) pg_{it}^g d\mathcal{D}(Demo_{it}) \end{aligned} \tag{12}$$

where the first term in the brackets captures the download likelihood, the second term captures the search strategy likelihood, and the first term outside the brackets captures the likelihood of search. $\mathcal{D}(Demo_{it})$ is the distribution of demographics. Since advertisers only know the distribution of demographics, the resulting probability must integrate over the demographics.

[15] This specification is consistent with the consumer information structure such that $\xi_i^{g\kappa}$ is not observed by researchers but known to consumer i.

[16] It is possible to allow for continous mixtures of heterogeneity as well. In our application, many consumers enter only once making it difficult to identify a consumer specific term for them.

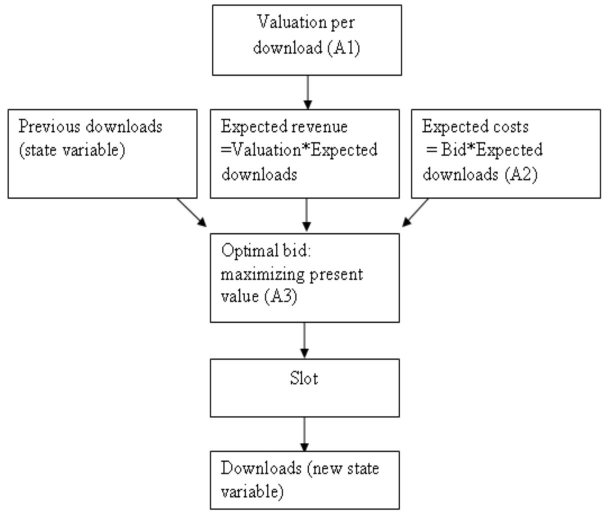

Figure 4: Advertiser Decisions

Correspondingly, the advertiser has an expected number of downloads for appearing in slot k, $d(k, X_j^t; \Omega_c)$, which can be computed as follows

$$d(k, X_j^t; \Omega_c) = M_t P_{ijt} \tag{13}$$

where Ω_c is the set of consumer preference parameters; M_t is the market size in period t.

3.2 Advertiser Model

Figure 4 overviews the dynamic game played by the advertiser. Advertiser j's problem is to decide the optimal bid amount b_j^t with the objective of maximizing discounted present value of payoffs. Higher bids lead to greater revenues because they yield more favorable positions on the search engine, thereby yielding more click-throughs for the advertiser. However, higher bids also increase costs (payments) leading to a trade-off between costs and revenues. The optimal decision of whether and how much to bid is incumbent upon the bidding mechanism, the characteristics of the advertiser, the information available at the time of bidding (including the state variables), and the nature of competitive interactions.

An advertiser's period profit for a download is the value it receives from the download less the costs (payments) of the download. Though we do not observe the value of a download, we infer this value by noting the observed bid can be rationalized only for a particular value accrued by the advertiser. We presume this value is drawn from a distribution known to all firms. The total period revenue for the advertiser is then the value per download times the expected number of downloads.[17] The total period payment upon winning is the number of downloads times the advertiser's bid. Hence, the total expected period profit is the number of downloads times the profit per download (i.e., the value per downloads less the payment per download).

Of course, the bid levels and expected download rates are affected by rules of the auction. Though we elaborate in further details on the specific rules of bidding below, at this point we simply note that the rules of the auction favor advertisers whose products were downloaded more frequently in the past since such products are more likely to lead to higher revenues for the platform.[18] Current period downloads are, in turn, affected by the position of the advertisement on the search engine. Because past downloads affect current placement, and thus current downloads, the advertiser problem is inherently dynamic; and past downloads are treated as a state variable.

Finally, given the rules of the auction, we note that all advertisers move simultaneously. While we presume a firm knows its own value, we assume competing firms know only the distribution of this value.

The process is depicted in Figure 4. We describe the process with more details as follows: Section 3.2.1 details the rules of the auction that affect the seller costs (A2), section 3.2.2 details the advertisers' value distribution (A1) and section 3.2.3 indicates how period values and costs translate to discounted profits and the resulting optimal bidding strategy (A3).

3.2.1 Seller Costs and the Bidding Mechanism

We begin by discussing how slot positions are allocated with respect to bids and the effect of these slot positions on consumer downloads (and thus advertiser revenue).

Upon a consumer completing a query, the search engine returns $k = 1, 2, ...K, ..., \overline{N}$ slots covering the products of all firms. Only the top $K = 5$ slots are considered as premium slots. Auctions for these K premium slots are held every period ($t = 1, 2, ...$). An advertiser seeks to appear in a more prominent slot because this may increase demand for the advertiser's product. Slots $K + 1$ to $\overline{N}$ are non-premium slots which compose a section called organic search section.

In order to procure a more favorable placement, advertiser j submits bid b_j^t in period t. These bids, submitted simultaneously, are summarized by the vector $\mathbf{b}^t = \{b_1^t, b_2^t, ..., b_N^t\}$.[19] Should an advertiser win slot k, the realized number of downloads d_j^t is a random draw from the distribution with the expectation $d(k, X_j^t; \Omega_c)$. The placement of advertisers into the K premium slots is determined by the ranking of their $\{b_j^t d_j^{t-1}\}_{\forall j}$, i.e. the product of current bid and last period realized downloads; the topmost bidder gets the best premium slot; the second bidder gets the second best

[17] The expected number of downloads is inferred form the consumer model and we have derived this expression in section 3.1.2.

[18] This is because the payment made to the search engine by an advertiser is the advertiser's bid times its total downloads.

[19] For the purpose of a clear exposition, we sometimes use boldface notations or pairs of braces to indicate vectors whose elements are variables across all bidders. For example, $\mathbf{d}^t = \{d_j^t\}_{\forall j}$ is a vector whose elements are $d_j^t, \forall j$.

premium slot and so on. A winner of one premium slot pays its own bid b_j^t for each download in current period. Hence the total payment for winning the auction is $b_j^t d_j^t$.

If an advertiser is not placed at one of the K premium slots, it will appear in the organic section; advertisers placed in the organic section *do not* pay for downloads from consumers. The ranking in the organic search section is determined by product update recency at period t, which is a component of the attribute vector of each product, X_j^t. Other attributes include price, consumer ratings and so on. For our purpose, we assume X_j^t is exogenously determined. These attributes are posted on the search engine and are common knowledge.

Given that the winners are determined in part by the previous period's downloads, the auction game is inherently dynamic. Before submitting a bid, the commonly observed state variables at time t are the realized past downloads of all bidders from period $t-1$.[20]

$$\mathbf{s}^t = \mathbf{d}^{t-1} = \{d_1^{t-1}, d_2^{t-1}, ..., d_N^{t-1}\} \tag{14}$$

3.2.2 Seller Value

The advertiser's bid determines the cost of advertising and must be weighed against the potential return when deciding how much to bid. We denote advertiser j's valuation regarding one download of its product in period t as v_j^t. We assume that this valuation is private information but drawn from a normal distribution that is commonly known to all advertisers. Specifically,

$$\begin{aligned} v_j^t &= v(X_j^t;\theta) + R_j^t \\ &= v(X_j^t;\theta) + r_j + r_j^t \\ &= X_j^t\theta + r_j + r_j^t \end{aligned} \tag{15}$$

where θ are parameters to be estimated and reflect the effect of product attributes on valuation; and $R_j^t = r_j + r_j^t$. r_j and r_j^t are independent random terms. $r_j \sim N(0, \psi_1^2)$ is a random effect term assumed to be identically and independently distributed *across advertisers*. This random term captures heterogeneity in valuations that may arise from unobserved firm specific effects such as more efficient operations. $r_j^t \sim N(0, \psi_2^2)$ is a private shock to an advertiser's valuation in period t, assumed to be identically and independently distributed *across advertisers and periods*. The sources of this private shock may include: (1) temporary increases in the advertiser's valuation due to some events such as a promotion campaign; (2) unexpected shocks to the advertiser's budget for financing the payments of the auction; (3) temporary production capacity constraint for delivering

[20] Though state variables can be categorized as endogeneous (past downloads) and exogenous (product attributes), our exposition characterizes only downloads as state variables because these are the only states whose evolution is subject to a dynamic constraint.

the product to users and so on. Given the distributions of r_j and r_j^t, the distribution of R_j^t is $N(0, \psi_1^2 + \psi_2^2)$.[21]

3.2.3 Seller Profits: A Markov Perfect Equilibrium (MPE)

Given state variable $\mathbf{s}^t$, v_j^t, predicted downloads and search engine's auction rules, bidder j decides the optimal bid amount b_j^t with the objective of maximizing discounted present value of payoffs. In light of this, every advertiser has an expected period payoff, which is a function of $\mathbf{s}^t$, $\mathbf{X}^t$, R_j^t and all advertisers' bids $\mathbf{b}^t$

$$\begin{aligned} & \mathbf{E}\pi_j\left(\mathbf{b}^t, \mathbf{s}^t, \mathbf{X}^t, R_j^t; \theta\right) \\ = & \sum\nolimits_{k=1}^{K} \Pr\left(k|b_j^t, \mathbf{b}_{-j}^t, \mathbf{s}^t, \mathbf{X}^t\right) \cdot (v_j^t - b_j^t) \cdot d(k, X_j^t; \Omega_c) \\ & + \sum\nolimits_{k=K+1}^{\overline{N}} \Pr\left(k|b_j^t, \mathbf{b}_{-j}^t, \mathbf{s}^t, \mathbf{X}^t\right) \cdot v_j^t \cdot d(k, X_j^t; \Omega_c) \\ = & \sum\nolimits_{k=1}^{K} \Pr\left(k|b_j^t, \mathbf{b}_{-j}^t, \mathbf{s}^t, \mathbf{X}^t\right) \cdot (X_j^t\theta + r_j + r_j^t - b_j^t) \cdot d(k, X_j^t; \Omega_c) \\ & + \sum\nolimits_{k=K+1}^{\overline{N}} \Pr\left(k|b_j^t, \mathbf{b}_{-j}^t, \mathbf{s}^t, \mathbf{X}^t\right) \cdot (X_j^t\theta + r_j + r_j^t) \cdot d(k, X_j^t; \Omega_c) \end{aligned} \tag{16}$$

where the expectation for profits is taken over other advertisers' bids $\mathbf{b}_{-j}^t$. $\Pr(k|\cdot)$ is the conditional probability of advertiser j getting slot k, $k = 1, 2, ..., \overline{N}$. $\Pr(k|\cdot)$ depends on not only bids, but also states $\mathbf{s}^t$ (the previous period's downloads) and product attributes $\mathbf{X}^t$.[22] This is because: i) the premium slot allocation is determined by the ranking of $\{b_j^t d_j^{t-1}\}_{\forall j}$, where $\mathbf{d}^{t-1}$ are the state variables and ii) the organic slot allocation is determined by product update recency, an element of $\mathbf{X}^t$.

In addition to the current period profit, an advertiser also takes its expected future payoffs into account when making decisions. In period t, given the state vector $\mathbf{s}^t$, advertiser j's discounted expected future payoffs evaluated prior to the realization of the private shock R_j^t is given by

$$\mathbf{E}\left[\sum\nolimits_{\tau=t}^{\infty} \rho^{\tau-t}\pi_j\left(\mathbf{b}^\tau, \mathbf{s}^\tau, \mathbf{X}^\tau, R_j^\tau; \Omega_a\right)|\mathbf{s}^t\right] \tag{17}$$

where $\Omega_a = \{\theta, \psi'\}$ and $\psi' = \{\psi_1, \psi_2\}$. ρ is a common discount factor. The expectation is taken over the random term R_j^t, bids in period t as well as all future realization of $\mathbf{X}^t$, shocks, bids and state variables. The state variables $\mathbf{s}^{t+1}$ in period $t+1$ is drawn from a probability distribution $P\left(\mathbf{s}^{t+1}|\mathbf{b}^t, \mathbf{s}^t, \mathbf{X}^t\right)$.

[21] The random shock r_j^t is realized at the beginning of period t. Although r_j^t is private knowledge, the distribution of $r_j^t \sim N(0, \psi_2^2)$ is common knowledge among bidders. Further, the random effect r_j of bidder j are known to all bidders but not to researchers. Given bidders may observe opponents' actions for many periods, the random effect can be perfectly inferred (Greene (2003)).

[22] Note that $\mathbf{s}^t$ and $\mathbf{X}^t$ are observed by all bidders before bidding.

We use the concept of a pure strategy Markov perfect equilibrium (MPE) to model the bidder's problem of whether and how much to bid in order to maximize the discounted expected future profits (Bajari et al. (2007); Ryan and Tucker (2007); Dubé et al. (2008); and others). The MPE implies that each bidder's bidding strategy only depends on the then-current profit-related information, including state, $\mathbf{X}^t$ and its private shock R_j^t. Hence we can describe the equilibrium bidding strategy of bidder j as a function $\sigma_j\left(\mathbf{s}^t, \mathbf{X}^t, R_j^t\right) = b_j^t$. Given a state vector $\mathbf{s}$ and product attributes $\mathbf{X}$ and prior to the realization of current R_j (with the time index t suppressed), bidder j's expected payoff under the equilibrium strategy profile $\boldsymbol{\sigma} = \{\sigma_1, \sigma_2, ..., \sigma_N\}$ can be expressed recursively as:

$$V_j\left(\mathbf{s}, X; \boldsymbol{\sigma}\right) = \mathbf{E}\left[\pi_j\left(\boldsymbol{\sigma}, \mathbf{s}, \mathbf{X}, R_j; \Omega_a\right) + \rho \int_{\mathbf{s}'} V_j\left(\mathbf{s}', \mathbf{X}'; \boldsymbol{\sigma}\right) dP\left(\mathbf{s}'|\mathbf{b}, \mathbf{s}, \mathbf{X}\right)|\mathbf{s}\right] \tag{18}$$

where the expectation is taken over current and future realizations of random terms R and $\mathbf{X}$.

The advertiser model can then be used in conjunction with the consumer model to forecast advertiser behavior as we shall discuss in the policy simulation section. In a nutshell, we presume advertiser will choose bids to maximize their expected profits. A change in information states, bidding mechanisms or webpage design will lead to an attendant change in bids conditioned on the advertisers value function which we estimate as described next.

4 Estimation

4.1 An Overview

Though it is standard to estimate dynamic MPE models via a dynamic programming approach such as a nested fixed point estimator (Rust (1994)), this requires one to repetitively evaluate the value function (18) through dynamic programming for each instance in which the parameters of the value function are updated. Even when feasible, it is computationally demanding to implement this approach. Instead, we consider the class of two-step estimators. The two-step estimators are predicated upon the notion that the dynamic program can be estimated in two steps that dramatically simplify the estimation process by facilitating the computation of the value function. Specifically, in this application we implement the two-step estimator proposed by Bajari et al. (2007) (BBL henceforth).

As can be seen in equation 18, the value function is parameterized by the primitives of the value distribution $\Omega_a = \{\theta, \psi'\}$. Under the assumption that advertisers are behaving rationally, these advertiser private values for clicks should be consistent with observed bidding strategies. Therefore, in the second step estimation, values of $\Omega_a = \{\theta, \psi'\}$ are chosen so as to make the observed bidding strategies to be congruent with rational behavior. We detail this step in Section 4.3 below.

However, as can be observed in equations 18 and 16, computation of the value function is also incumbent upon i) the bidding policy function that maps bids to the states (downloads), product attributes, and private shocks $\sigma_j\left(\mathbf{s}^t, \mathbf{X}^t, R_j^t\right) = b_j^t$; ii) the expected downloads $d(k, X_j^t; \Omega_c)$, and iii) a function that maps the likelihood of future states as a function of current states and actions $P\left(\mathbf{s}^{t+1}|\mathbf{b}^t, \mathbf{s}^t, \mathbf{X}^t\right)$. These are estimated in the first step as detailed in Section 4.2 below and then substituted into the value function used in the second step estimation.

4.2 First Step Estimation

In the first step of the estimation we seek to obtain:

1. A "partial" policy function $\widetilde{\sigma}_j\left(\mathbf{s}, \mathbf{X}\right)$ describing the equilibrium bidding strategies as a function of the observed state variables and product attributes, $\mathbf{X}$. We estimate the policy function by noting that players adopt equilibrium strategies (or decision rules) and that behaviors generated from these decision rules lead to correlations between i) the observed states (i.e., past period downloads) and product characteristics and ii) advertiser decisions (i.e., bids). The partial policy function captures this correlation. In our case, we use a random effects Tobit model to link bids to states and product characteristics as described in Section A.1.1 of the Appendix. Subsequently, the full policy function $\sigma_j\left(\mathbf{s}, \mathbf{X}, R_j\right)$ can be inferred based on $\widetilde{\sigma}_j\left(\mathbf{s}, \mathbf{X}\right)$ and the distribution of private random shocks R_j. The partial policy function can be thought of as the marginal distribution of the full policy function. Inferences regarding the parameters of the full policy function can be made by finding the distribution of R_j that, when "integrated out," leads to the best rationalization for the observed bids. We discuss our approach to infer the full policy functions from the partial policy function in Appendix A.1.1.

2. The expected downloads for a given firm at a given slot, $d(k, X_j; \Omega_c)$. The $d(k, X_j; \Omega_c)$ follows directly from the consumer model. Hence, the first step estimation involves i) estimating the parameters of the consumer model and then ii) using these estimates to compute the expected number of downloads. The expected total number of downloads as a function of slot position and product attributes is obtained by using the results of the consumer model to forecast the likelihood of each person downloading the software and then summing these probabilities across persons.[23] We discuss our approach for determining the expected downloads in Section A.1.2 of the Appendix.

[23] As an aside we note that advertisers have limited information from which to form expectations about total downloads because they observe only the aggregate information of downloads but not the individual specific download decisions. Hence advertisers must infer the distribution of consumer preferences from these aggregate statistics. In a subsequent policy simulation we allow the search engine to provide individual level information to advertisers in order to assess how it affects advertiser behavior and therefore search engine revenues.

3. The state transition probability $P(\mathbf{s}'|\mathbf{b}, \mathbf{s}, \mathbf{X})$ which describes the distribution of future states (current period downloads) given observations of the current state (past downloads), product attributes and actions (current period bids). These state transitions can be derived by i) using the policy function to predict bids as a function of past downloads, ii) determining the slot ranking as a function of these bids, past downloads and product attributes, and then iii) using the consumer model to predict the number of current downloads as a function of slot position. Details regarding our approach to determining the state transition probabilities is outlined in Section A.1.3 of the Appendix.

With the first step estimates of $\sigma_j(\mathbf{s}, \mathbf{X}, R_j)$, $d(k, X_j; \Omega_c)$ and $P(\mathbf{s}'|\mathbf{b}, \mathbf{s}, \mathbf{X})$, we can compute the value function in 18 as a function with only $\Omega_a = \{\theta, \psi'\}$ unknown. In the second step we estimate these parameters.

4.3 Second Step Estimation

The goal of the second step estimation is to recover the primitives of the bidder value function, $\Omega_a = \{\theta, \psi'\}$. The intuition behind how the second-stage estimation works is that true parameters should rationalize the observed data. For bidders' data to be generated by rational plays, we need

$$V_j(\mathbf{s}, \mathbf{X}; \sigma_j, \boldsymbol{\sigma}_{-j}; \Omega_a) \geq V_j\left(\mathbf{s}, \mathbf{X}; \sigma'_j, \boldsymbol{\sigma}_{-j}; \Omega_a\right), \forall \sigma'_j \neq \sigma_j \tag{19}$$

where σ_j is the equilibrium policy function; σ'_j is some deviations from σ_j. This equation means that any deviations from the observed equilibrium bidding strategy will not result in more profits. Otherwise, the strategy would not be optimal. Hence, we first simulate the value functions under the equilibrium policy σ_j and the deviated policy σ'_j (i.e., the left hand side and the right hand side of equation 19). Then we try to choose $\Omega_a = \{\theta, \psi'\}$ to maximize the likelihood that equation 19 holds. We describe the details of this second step estimation in Appendix A.2.

4.4 Sampling Chain

With the posterior distributions for the advertiser and consumer models established, we estimate the models using MCMC approach as detailed in Appendix B. This is a notable deviation from prior research that uses a gradient based technique. The advantage of using a Bayesian approach, as long as suitable parametric assumptions can be invoked, is that it facilitates model convergence, has desirable small sample properties, increases statistical efficiency, and enables the estimation of a wide array of functional forms (Rossi et al. (2005)). Hence we seek to make a methodological contribution to the burgeoning literature on two-step estimators for dynamic games.

5 Results

5.1 First Step Estimation Results

Recall, the goal of the first step estimation is to determine the policy function, $\sigma_j\left(\mathbf{s}^t, \mathbf{X}^t, R_j^t\right)$, the expected downloads $d(k, X_j^t; \Omega_c)$, and the state transition probabilities $P\left(\mathbf{s}^{t+1}|\mathbf{b}^t, \mathbf{s}^t, \mathbf{X}^t\right)$. To determine $\sigma_j\left(\mathbf{s}^t, \mathbf{X}^t, R_j^t\right)$, we first estimate the partial policy function $\widetilde{\sigma}_j\left(\mathbf{s}^t, \mathbf{X}^t\right)$ and then compute the full policy function. To determine $d(k, X_j^t; \Omega_c)$, we first estimate the consumer model and then compute the expected downloads. Last $P\left(\mathbf{s}^{t+1}|\mathbf{b}^t, \mathbf{s}^t, \mathbf{X}^t\right)$ is derived from the consumer model and the partial policy function. Thus, in the first stage we need only to estimate the partial policy function and the consumer model. With these estimates in hand we compute $\sigma_j\left(\mathbf{s}^t, \mathbf{X}^t, R_j^t\right)$, $d(k, X_j^t; \Omega_c)$, and $P\left(\mathbf{s}^{t+1}|\mathbf{b}^t, \mathbf{s}^t, \mathbf{X}^t\right)$ for use in the second step. Thus, below, we report the estimates for the partial policy function and the consumer model on which these functions are all based.

5.1.1 Partial Policy Function $\widetilde{\sigma}_j(\mathbf{s}, \mathbf{X})$

The vector of independent variables $(\mathbf{s}, \mathbf{X})$ for the partial policy function (i.e., the Tobit model of advertiser behavior that captures their bidding policy as outlined in Appendix section A.1.1) contains the following variables:

- Product j's state variable, last period download d_j^{t-1}. We reason that high past downloads increase the likelihood of a favorable placement and therefore affect bids.

- Two market level variables, the sum of last period downloads from all bidders and the number of bidders in last period. Since we only have 322 observations of bids, it is infeasible to estimate a parameter to reflect the effect of each opponent's state (i.e., competition) on the optimal bid. Moreover, it is unlikely a bidder can monitor every opponent's state in each period before bidding because such a strategy carries high cognitive and time costs. Hence, summary measures provide a reasonable approximation of competing states in a limited information context. Others in the literature who have invoked a similar approach include Jofre-Bonet and Pesendorfer (2003) and Ryan (2006). Like them, we find this provides a fair model fit. Another measure of competitive intensity is the number of opponents. Given bidders cannot directly observe the number of competitors in the current period, we used a lagged measure of the number of bidders.

- Product j's attributes in period t (X_j^t), including its non-trial version price, expert rating, consumer rating, update recency and compatibility with old/unpopular OS. We expect that a higher quality product will yield greater downloads thereby affecting the bidding strategy.

- To control the possible effect of the growth of ownership of MP3 players, we also collect the average lagged price of all new MP3 players in the market from a major online retailing platform (`www.pricegrabber.com`).

Table 4 reports the estimation results for the Tobit model. As a measure of fit of the model, we simulated 10,000 bids from the estimated distribution. The probability of observing a positive simulated bid is 41.0%; the probability of observing a positive bid in the real data is 41.6%. Conditional on observing a positive simulated bid, those bids have a mean of \$0.20 with a standard deviation of \$0.07. In the data the mean of observed positive bids is \$0.20 and the standard deviation is \$0.08. We also estimate the same model only using 70% (227/322) of the observations and use the left 30% as a holdout sample. The estimates have minimal changes. We then use the holdout to simulate 10,000 bids. The probability of observing a positive bid is 40.2% while there are 41.1% positive bids in the holdout sample. Among the positive simulated bids, the mean is \$0.22 and the standard deviation is \$0.09. The corresponding statistics in the holdout is \$0.21 and \$0.11. Overall, the fit is good.

Table 4: Bidding Function Estimates

	Median	95% Interval
φ		
Constant	-10.55^*	$(-15.23, -6.78)$
Lagged Downloads$_{jt}/10^3$	-0.12^*	$(-0.16, -0.08)$
Total Lagged Downloads$_t/10^3$	0.04^*	$(0.01, 0.08)$
Lagged Number of Bidders$_t$	0.04	$(-0.55, 0.26)$
Lapse Since Last Update$_{jt}$	-0.41^*	$(-0.85, -0.05)$
Non-trial Version Price$_{jt}$	0.37^*	$(0.31, 0.39)$
Expert Ratings$_{jt}$	0.44	$(-6.30, 2.77)$
Consumer Ratings$_{jt}$	0.92^*	$(0.09, 1.66)$
Compatibility Index$_{jt}$	-1.73^*	$(-2.79, -0.75)$
Lagged MP3 Player Price$_t$	0.03^*	$(0.02, 0.03)$
τ	7.51^*	$(7.02, 7.99)$
τ_{re}	14.75^*	$(14.55, 14.94)$
Log Marginal Likelihood	-1148.05	

The estimates yield several insights into the observed bidding strategy. First, the bidder's state variable (d_j^{t-1}) is negatively correlated with its bid amount b_j^t because the ranking of the auction is determined by the product of b_j^t and d_j^{t-1}. All else equal, a higher number of lagged downloads means a bidder can bid less to obtain the same slot. Second, the total number of lagged downloads in the previous period ($\sum_{j'} d_{j'}^{t-1}$) and the lagged number of bidders both have positive impact on a bidder's bid. We take this to mean increased competition leads to higher bids. Third, bids are increasing in the product price. One possible explanation is that a high priced product yields more value to the firm for each download and hence the firm competes more aggressively for a top slot.

Similarly and fourth, a high price for MP3 players reflects greater value for the downloads also leading to a positive effect on bids. Fifth, "Lapse Since Last Update" has a negative effect on bids. Older products are more likely obsolete, thereby generating lower value for consumers. If this is the case, firms can reasonably expect fewer final purchases after downloads and bid less for these products. Likewise and sixth, higher compatibility with prior software versions reflects product age leading to a negative estimate for this variable. Finally, ratings from consumers and experts (albeit not significant for experts) have a positive correlation with bid amounts – these again imply greater consumer value for the goods making it more profitable to advertise them.

5.1.2 Consumer Model

The consumer model is estimated using MCMC approach based on the posterior distribution described in Appendix A.1.2. We consider the download decisions for each of the 21 products who entered auctions plus the top 3 products who did not. Together these firms constitute over 80% of all downloads. The remaining number of downloads are scattered across 370 other firms, each of whom has negligible share. Hence we exclude them from our analysis.

Table 5: Alternative Numbers of Latent Segments

	Log Marginal Likelihood
1 Segment	−12769.3
2 Segments*	−12511.9
3 Segments	−12571.1
4 Segments	−12551.4

We calibrate the model by estimating an increasing number of latent segments until there is no significant improvement in model fit. We use log marginal likelihood as the measurement for model fit. In Table 5 we report the comparison of the log marginal likelihoods for models with up to 4 segments. The model with 2 segments gives the best result.

Table 6 presents the estimates of the model with 2 segments. Conditional on the estimated segment parameters and demographic distribution, we calculate the segment sizes as 89.5% and 10.5%, respectively. Based on the parameter estimates in Table 6, Segment 1 is less likely to initiate a search (low λ_0^g). Moreover, upon engaging a search, this segment appears to be less sensitive to slot ranking but more sensitive to consumer and expert ratings than Segment 2. Segment 2, who searches more frequently, relies more heavily on the slot order when downloading. Overall, we speculate that segment 1 are the occasional downloaders who base their download decisions on others' ratings and tend not to exclude goods of high price. In contrast, segment 2 are the "experts," who tend to rely on their own assessments when downloading.

Table 6: Consumer Model Estimates

	Segment 1 (89.5%) (Infrequent searcher)	Segment 2 (10.5%) (Frequent searcher and slot sensitive)
	Median (95%Interval)	Median (95%Interval)
β^g (utility parameters)		
Constant	-0.09 $(-0.11, 0.001)$	0.35 $(0.31, 0.38)$
Slot Rank	-0.08 $(-0.06, -0.09)$	-0.51 $(-0.52, -0.50)$
Non-trial Version Price	0.03 $(0.03, 0.04)$	-0.04 $(-0.04, -0.03)$
Expert Ratings	0.16 $(0.15, 0.17)$	0.06 $(0.06, 0.07)$
Consumer Ratings	0.11 $(0.11, 0.12)$	0.03 $(0.03, 0.05)$
Compatibility Index	-0.08 $(-0.09, -0.07)$	0.16 $(0.16, 0.17)$
Total Download Percentage	0.01 $(-0.02, 0.05)$	0.09 $(0.08, 0.10)$
δ^g (sorting/filtering scaling)	1.52 $(1.48, 1.55)$	1.87 $(1.78, 1.99)$
λ^g (search probability)		
λ_0^g (base)	-10.22 $(-10.75, -9.60)$	-0.78 $(-1.21, -0.54)$
λ_1^g (1-correlation)	0.02 $(0.01, 0.02)$	0.03 $(0.01, 0.04)$
γ^g (segment parameters)		
Constant	–	-4.01 $(-4.74, -2.87)$
Music Site Visited	–	7.66 $(5.77, 10.18)$
Registration Status	–	-0.24 $(-1.91, 0.86)$
Product Downloaded in Last Month	–	-0.40 $(-1.57, -0.02)$

More insights on this contrast can be gleaned by determining the predicted probabilities of searching and sorting/filtering by computing $\Pr(search_i^g) = \frac{\exp(\lambda_0^g + \lambda_1^g IV_{it}^g)}{1+\exp(\lambda_0^g + \lambda_1^g IV_{it}^g)}$ and $\Pr(\kappa)_{it}^g = \frac{\exp(U_{it}^{g\kappa})}{\sum_{\kappa'=0}^{3} \exp(U_{it}^{g\kappa'})}$ in equation 9 and 7, respectively. Table 7 reports these probabilities for both segments.

Table 7: Searching Behavior of Consumers

	Segment 1	Segment 2
Searching	0.4%	60.8%
No sorting or filtering	78.7%	86.1%
Sorting but no filtering	21.3%	8.2%
No sorting but filtering	$\to 0$	5.5%
Sorting and filtering	$\to 0$	0.3%

Table 7 confirms the tendency of Segment 2 to be more likely to initiate a search in the focal category. Though comprising only 10.5% of all consumers, they represent 95% of all searches. The increased searching frequency suggests that members of Segment 2 are ideal customers to target because more searches lead to more downloads.

Moreover, Segment 2 is more likely to be influenced by sponsored advertising. To see this, note that Segment 1 consumers put more weights on ratings of products (e.g., expert and consumer ratings) than Segment 2 consumers do. As a consequence Segment 1 consumers engage in far more sorting. Sorting eliminates the advantage conferred by sponsored advertising because winners of the sponsored search auction may be sorted out of desirable slots on the page.

However, Table 7 indicates consumers in Segment 1 seldom filter. Filtering occurs when consumers seek to exclude negative utility options from the choice set. Given the high sensitivity to rank order, Segment 2 is more prone to eliminate options. We suspect this segment, by virtue of being a frequent visitor, searches for very specific products that conform to a particular need. This also increases the chance a sponsored link will be filtered. Overall, however, Segment 1 is more likely to sort and/or filter than Segment 2 (21.3% vs. 13.9%) suggesting that Segment 2 is more valuable to advertisers. We will explore this conjecture in more detail in our policy analysis.

5.2 Second Step Estimation Results

Table 8 shows the results of second step estimation.[24] We find that newer, more expensive and better rated products yield greater values to the advertiser. This is consistent with our conjecture in Section 5.1.1 that firms bid more aggressively when having higher values for downloads. We find that, after controlling for observed product characteristics, 95% of the variation in valuations

[24] We do not estimate the discount factor ρ. As shown in Rust (1994), the discount factor is usually unidentified. We fix $\rho = 0.95$ for our estimation. We also experiment $\rho = 0.90$ and see minimal difference in the results.

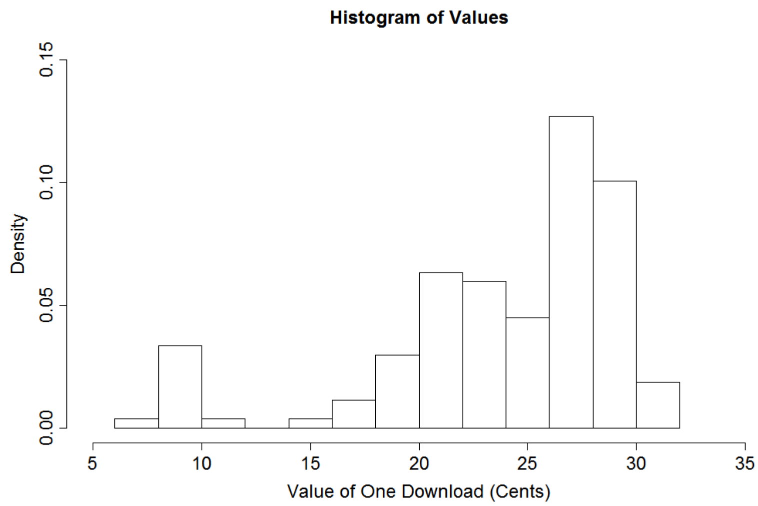

Figure 5: Distribution of Values per Download

across firms is on the order of \$0.02. We attribute this variation in part due to differences in the operating efficiency of the firms.

Table 8: Value per Click Parameter Estimates

	Median	95% Interval
θ		
Constant	5.55*	$(3.02, 6.34)$
Lapse Since Last Update$_{jt}$	-0.74*	$(-0.84, -0.59)$
Non-trial Version Price$_{jt}$	0.24*	$(0.15, 0.39)$
Expert Ratings$_{jt}$	0.50*	$(0.27, 0.59)$
Consumer Ratings$_{jt}$	1.23*	$(1.10, 1.41)$
Compatibility Index$_{jt}$	-0.31*	$(-0.43, -0.21)$
Lagged MP3 Player Price$_t$	0.03*	$(0.02, 0.04)$
ψ_1, random effect std. dev.	0.75*	$(0.04, 1.89)$
ψ_2, random shock std. dev.	0.45*	$(0.03, 0.99)$
Log Marginal Likelihood	-1683.41	

Given the second step results, we can further estimate the value of a download to a firm. We estimate the advertiser's value for a download in each period. The distribution of these estimates across time and advertisers is depicted in Figure 5. As indicated in the figure there is substantial variation in the valuation of downloads. Table 8 explains some of this variation as a function of the characteristics of the software and firm specific effects. Overall, the mean value of a download to these advertisers is \$0.24. This compares to an average bid of \$0.20 as indicated in Table 1. Hence, on average, each click implies an expected return to a firm of about \$0.04. To our knowledge, this is the first paper to impute the advertiser's return from a click in a key word search context. One way to interpret these results is to consider the firm's expected sales per download to rationalize the bid. The firm's profit per click is roughly $CR_j^t \cdot P_j^t - b_j^t$, where CR_j^t indicates the download-sale conversion rate (or sales per download) and P_j^t is the non-trial version price. Ignoring dynamic effects and setting this profit per click equal to $v_j^t - b_j^t$ yields a rough approximation of the conversion rate as $CR_j^t = v_j^t / P_j^t$. Viewed in this light, the effect of higher quality software, which raises v_j^t, leads to a higher implied conversion rate.[25] Noting that the average price of the software is \$22, this average per-click valuation implies that 1.1% of all clicks lead to a purchase (that is, the conversion rate is $0.24/22 = 1.1\%$). This estimate lies within the industry average conversion rate of $1-2\%$ reported by Gamedaily.com, suggesting our findings have high face validity.[26]

6 Policy Simulations

We describe four policy simulations: i) the value of disaggregate consumer data, ii) the value of targeting (i.e., not only allowing advertisers to bid on keywords but also on market segments.) iii) the effect of alternative pricing mechanisms on search engine revenue and iv) the effect of alternative webpage designs on search engine revenues.[27] Details regarding the implementation of the policy simulations are presented in Appendix C. Hence,we limit our discussion to the objectives and insights from these simulations.

6.1 Policy Simulation I: Incorporating Disaggregate-level Data

Advertisers and search engines are endowed with different levels of information. The search engine know all of the clicks made by visitors to its site. The advertiser knows only the total number of downloads all the advertisers received. Hence, there is an information asymmetry arising from the different level of market intelligence accruing to each respective agent. Given this disaggregate consumer information is owned by the search engine but not observed by advertisers, it is relevant to ask how the information revelation from the platform to advertisers will affect advertiser behavior and hence platform revenues. In practice, this means that the platform is interested in whether to sell or give this information to advertisers and how it should be priced. More generally, this

[25] Note we do not model the equilibrium pricing strategy of the firms. We conjecture that this pricing game is played across multiple markets and media as well as over a longer time horizon. The characteristics of the key word advertising problem likely have only a small effect on prices set by firms. Our treatment of prices as exogenous is consistent with all the prior research in key word search.

[26] "Casual Free to Pay Conversion Rate Too Low." Gamedaily.com (`http://www.gamedaily.com/articles/features/magid-casual-free-to-pay-conversion-rate-too-low/70943/?biz=1`).

[27] These policy simulations may involve explicitly solving the dynamic programming problem for advertisers. This is because the environment variables such as auction rules have been changed, which makes the estimated bidding policy function become inapplicable. The advance in research on approximate dynamic programming makes solving high-dimension DP problems become possible (Powell (2007)).

counterfactual exemplified the value of market intelligence and how it can be computed in the context of a structural model.

Accordingly, we implement a counterfactual scenario under which advertisers have access to the click histories of consumers. We then assess i) how bidding behavior and returns to advertisers change under this counterfactual information structure and ii) how the resulting revenues change at the search engine. By comparing these returns with those in the observed case of information asymmetry, we can obtain a measure for the value of the information. This value leads to managerial prescriptions for the search engine regarding the pricing strategy of such information.

6.2 Policy Simulation II: Segmentation and Targeting

It can be profitable for advertisers to target specific consumers. In this instance, instead of a single bid on a key word, an advertiser can vary its bids across market segments. For example, consider two segments A and B wherein segment B is more sensitive to product price and segment A is more sensitive to product quality. Consider further, two firms X and Y where firm X purveys a lower price, but lower quality, product. Intuitively, firm X should bid more aggressively for segment B because quality sensitive segment A will not likely buy the low quality good X. This should lead to higher revenues for the search engine. On the other hand, there is less bidding competition for firm X within segment B because Y finds this segment unattractive – this dearth of competition can drive the bid of X down for segment B. This would place a downward pressure on search engine profits. Hence, the optimal revenue outcome for the search engine is likely to be incumbent upon the distribution of consumer preferences and the characteristics of the goods being advertised. Our approach can assess these effects of segmentation and targeting strategy on the search engine's revenue.

6.3 Policy Simulation III: Alternative Auction Mechanisms

Auction mechanism design has established a rich literature body since the study by Vickrey (1961). With the purpose of revenue maximization,[28] the optimal design involves several aspects such as the determination of payments and winners as well as the choice of reserve price. We will focus on the payment rule in this investigation.

While the focal search engine currently charges winning advertisers their own bids, many major search engines such as `Google` and `Yahoo!` are applying a "generalized second-price auction" as termed in Edelman et al. (2007). Under the generalized second-price auction rules, winners are still determined by the ranking of $\{b_{j'}^{t} d_{j'}^{t-1}\}_{\forall j'}$. However, instead of paying its own bid amount,

[28] Sometimes efficiency of allocation is also an objective of the auction design. An efficient auction mechanism enables the bidder with the highest value to win the best slot, the second value bidder to get the second best slot and so on. While most search engines are profit-seeking firms, we will only focus on the revenue maximization objective.

the winner of a slot pays the highest losing bidder's bid adjusted by their last period downloads.[29] For example, suppose bidder j wins a slot with the bid of b_j^t and last period download d_j^{t-1}, its payment for each download will be $b_{j'}^t d_{j'}^{t-1}/d_j^{t-1}$, where j' is the highest losing bidders for the slot bidder j wins.

In the context of sponsored search auctions, although "generalized second-price auction" is widely adopted by major search engines, the optimality of such a mechanism is not confirmed (Iyengar and Kumar (2006); Katona (2007)). With the estimates from the model, we are able to implement a policy simulation by altering the payment rule of the game and compare the revenues of the search engine under the two different mechanisms. We intend to gain some empirical knowledge about the auction mechanism design in sponsored search auctions, which will also shed insights into future theoretical investigations.

6.4 Policy Simulation IV: Alternative Webpage Design

The goal of the search engine's sorting/filtering options is to provide consumers with easier access to price and rating information across different products. As shown in section 3.1 and evidenced by our results, sorting and filtering play a crucial role in consumer decision process. In light of this outcome, it is possible to consider an alternative webpage design of the search engine – eliminating the option of sorting and filtering for consumers – and assessing the resulting impact on consumer search, advertiser bidding, and the search engine's revenues. As we note below, the sorting and filtering options can have conflicting outcomes on the nature of competition and therefore advertiser bidding behavior.

One view is that sorting increases competition by making products more substitutable. Alba et al. (1997) hence express the concern faced by many online retailers: since online shopping reduces search costs, consumers increase their consideration sets which intensifies competition. Diehl et al. (2003) show that, based on a consumer's keyword query, search engines oversample products that match the consumer's interests; these sampled products are more likely to be close substitutes. Thus, there will be less product differentiation and more intensified competition. An implication for sponsored search advertising is that advertisers would bid more aggressively in the auctions to secure premium slots in order to differentiate their goods.

An alternative view, espoused by Lynch and Ariely (2000), proposes that sorting may actually decrease competition. Search engines not only lead to lower search costs, but they can also make quality information more salient. When making decisions, consumers place greater weight on attributes that are more convenient to process (cf. Russo (1977); Häubl and Murray (2003)). In our case, the consumer and expert ratings become more prominent to consumers. Thus, when consumers have the easier access to rating information, the product quality becomes more important to consumers and advertisers may have less incentive to bid.

[29] In the paper by Edelman et al. (2007), the adjustment using last period downloads is not considered.

Using our model, it can be tested which effect may have a greater impact on consumer behavior and therefore the advertisers' bidding incentives. Using our integrated approach to assessing advertiser and consumer behavior, we can further impute the consequences of a change in web page design on search engine revenues.

7 Conclusion

Given the \$9B firms annually spend on key word advertising and its rapid growth, we contend that the topic is of central concern to advertisers and platforms that host advertising alike. In light of this growth, it is surprising that there is little extant empirical research pertaining to modeling the demand and pricing for key word advertising. As a result, we develop a dynamic structural model of advertiser bidding behavior coupled with an attendant model of search behavior. The interplay of these two agents has a number of implications for the platform that hosts them. The model is dynamic because past clicks on the advertisers' links affect the search engine's current allocation of advertising slots. We adopt a structural approach in order to simulate the effect of various changes in the search engine's policy. In particular, we consider i) how the platform or search engine should price its advertising via alternative auction mechanisms, ii) whether the platform should accommodate targeted bidding wherein advertisers bid not only on key words, but also behavioral segments (e.g., those that purchase more often), iii) whether and how the search engine should sell information on individual clicking histories and iv) how an alternative webpage design of the search engine with less product information would affect bidding behavior and the engine's revenues.

Our model of bidding behavior is predicated on the advertiser choosing its bids to maximize the net present value of its discounted profits. The period profits contain two components – i) the advertiser's value for a given click times the number of clicks on the advertisement and ii) the payment in form of the bid per click times the number of clicks on the advertisement. Whereas the advertiser's costs are determined by their bids, we infer the advertiser's valuation for clicks. Specifically, we estimate valuations by choosing them such that, for an observed set of bids, the valuations rationalize the bidding strategy; that is, making profits as high as possible. In this sense, our structural model "backs out" the advertiser's expectation for the profit per click. Given an estimate of these valuations, it becomes possible to ascertain how advertiser profits are affected by a change in the rules of the auction, a change in the webpage design, or a change in the information state of the advertiser. As noted above, another central component to the calculation of advertiser profits is the expectation of the number of clicks on its advertisement received from consumers. This expectation of clicks is imputed from our consumer model.

Our consumer choice model follows from the standard random utility theory (McFadden (1977)) and is computed using traditional MCMC methods adapted to our context. The advertiser model is less straightforward because it is a dynamic program. We use recent advances in economics wherein a two-step estimator is applied to the problem (BBL). The first step is used to infer the bidding policy and consumer clicking behavior. The second step is used to infer the advertiser valuations conditioned on the bidding strategy and the consumer clicking decisions. Our approach departs from previous work on two-step estimators via our Bayesian instantiation. Like all MCMC approaches, this innovation enables one to estimate a broader set of models and does not rely on asymptotic for inference (Rossi et al. (2005)).

The estimates from our empirical model yield some insights into advertiser bidding behavior and consumer searching behavior. The estimates from advertiser bidding function indicates that bid amounts have positive correlations with product attributes that may enhance product quality. One possible explanation is that a higher quality leads to the advertiser's greater valuation about sponsored search advertising and hence the more aggressive bidding. Our consumer model indicates that consumers who engage in more search and clicking may also be more responsive to sponsored advertising than others. If so, these consumers should be the focus of advertisers and search engines' marketing campaigns.

Further policy insights will be drawn from the ongoing policy simulations.

Several extensions are possible. First, we use a two-step estimator to model the dynamic bidding behavior of advertisers without explicitly solving for the equilibrium bidding strategy. Solving explicitly for this strategy could provide more insights into bidder behavior in this new marketing phenomenon. For example, following the extant literature we assume that a bidder's return of the advertising only comes from consumers clicks. It is possible that advertisers also accrue some values from the exposures at the premium slots. A clear characterization of bidding strategy can better facilitate our understanding about how advertisers value sponsored advertising in term of clicks and exposures and hence present a better guideline for search engines to design their pricing schedule. Second, our analysis focuses upon a single category. Bidding across multiple keywords is an important direction for future research. In particular, the existence of multiple keywords auctions may present opportunities for collusions among bidders. For example, advertisers may collusively diverge their bids to different keywords. By doing so, they can find a more profitable trade-off between payments to the search engine and clicks across keywords. In a theoretical paper by Stryszowska (2005) the author shows an equilibrium where bidders implicitly collude across multiple auctions in the context of online auctions such as `eBay.com`. One managerial implication is how to detect and discourage collusions and reduce its negative impact on search engine revenues. Third, competition between search engines over advertisers is not modeled. Though our data provider has a dominant role in this specific category, inter-engine competition is unattended in the literature. To some extent, sponsored search advertising can be understood as advertisers purchasing products (media) from search engines through auctions. An advertiser makes discrete choice about search engines before entering auctions. Little research has been done on the advertiser's choice problem, even though there is abundant discrete choice research that can be applied (cf. Palma et al. (1992)). Accordingly, the inter-engine competition deserves future attention. Overall, we hope this study will inspire further work to enrich our knowledge of this new marketplace.

Appendix

A Two Step Estimator

A.1 First Step Estimation

A.1.1 Estimating the Advertiser's Policy Function

The Partial Policy Function The partial policy function links states ($\mathbf{s}$) and characteristics ($\mathbf{X}$) to decisions ($\mathbf{b}$). Ideally this relation can be captured by a relatively flexible parametric form and estimated via methods such as maximum likelihood or MCMC to obtain the partial policy function parameter estimates. The exact functional form is typically determined by model fit comparison among multiple specifications (e.g., Jofre-Bonet and Pesendorfer (2003)). We considered several different specifications for the distribution of bids and found the truncated normal distribution gives the best fit in terms of marginal likelihoods.[30] Specifically, we allow

$$\begin{aligned} b_j^t &= \begin{cases} y_j^{t*} & \text{if } y_j^{t*} \geq \chi \\ 0 & \text{otherwise} \end{cases} \\ y_j^{t*} &\sim N([\mathbf{s}^{t\prime}, \mathbf{X}_j^{t\prime}] \cdot \varphi + \varphi_j^{re}, \tau^2) \\ \varphi_j^{re} &\sim N(0, \tau_{re}^2) \end{aligned} \tag{A1}$$

where $[\mathbf{s}^{t\prime}, \mathbf{X}_j^{t\prime}]$ is the vector of independent variables; τ is the standard deviation of y_{jt}^*; φ_j^{re} is a bidder specific random effect whose distribution is $N(0, \tau_{re}^2)$; χ is the truncation point, which is set at 15 to be consistent with the 15¢ minimum bid requirement of the search engine.

One possible concern when estimating the partial policy function $\widetilde{\boldsymbol{\sigma}}(\mathbf{s}, \mathbf{X})$ (and the full policy function $\boldsymbol{\sigma}(\mathbf{s}, \mathbf{X}, R_j)$ next) is that there may be multiple equilibrium strategies; and the observed data are generated by multiple equilibria. If this were the case, the policy function would not lead to a unique decision and would be of limited use in predicting advertiser behavior. It is therefore necessary to invoke the following assumption (BBL).

Assumption 2 (Equilibrium Selection): The data are generated by a single Markov perfect equilibrium profile $\boldsymbol{\sigma}$.

Assumption 2 is relatively unrestrictive since our data is generated by auctions of *one* keyword and from *one* search engine. Given data are from a single market, the likelihood is diminished that different equilibria from different markets are confounded. We note that this assumption is often employed in such contexts (e.g., Dubé et al. (2008)).

[30] We experimented alternative specifications including a Beta distribution and a Weibull distribution whose scale, shape and location parameters are functions of $(\mathbf{s}, \mathbf{X})$. The current specification gives the best fit in terms of marginal likelihoods.

This partial policy function is then used to impute the full policy function $b_j = \sigma_j(\mathbf{s}, \mathbf{X}, R_j)$ as detailed below based on R_j's distribution parameters $\psi' = (\psi_1, \psi_2)$.

Full Policy Functions $\sigma_j^t\left(\mathbf{s}^t, \mathbf{X}^t, R_j^t\right)$ To evaluate the value function of this dynamic game, we need to calculate bids as a function of not only $(\mathbf{s}^t, \mathbf{X}^t)$ but also the unobserved shocks $R_j^t = r_j + r_j^t$ (see section 3.2.3). To infer this full policy function $\sigma_j\left(\mathbf{s}^t, \mathbf{X}^t, R_j^t\right)$ from the estimated partial policy function, $\widetilde{\sigma}_j(\mathbf{s}^t, \mathbf{X}^t)$, we introduce two additional assumptions.

Assumption 3 (Monotone Choice): For each bidder j, its equilibrium strategy $\sigma_j\left(\mathbf{s}^t, \mathbf{X}^t, R_j^t\right)$ is increasing in R_j^t (BBL).

Assumption 3 implies that bidders who draw higher private valuation shocks R_j^t will bid more aggressively.

Assumption 4: The ratio of standard deviations of r_j and r_j^t equals to the ratio of τ_{re} and τ, the standard deviations of random effects and shocks in the partial policy function.

Assumption 4 implies that the bidder's latent variable y_j^{t*} are affected proportionally by r_j and r_j^t in terms of magnitude.

To explore these two assumptions, note that the partial policy function $\widetilde{\boldsymbol{\sigma}}\left(\mathbf{s}^t, \mathbf{X}^t\right)$ presents distributions for bid b_j^t and the latent y_j^{t*}, whose CDF's we denote as $F_b\left(b_j^t|\mathbf{s}^t, \mathbf{X}^t\right)$ and $F\left(y_j^{t*}|\mathbf{s}^t, \mathbf{X}^t\right)$, respectively.[31] According to the model in equation A1, the population mean of y_j^{t*} across bidders and periods is $[\mathbf{s}^t, \mathbf{X}_j^t] \cdot \varphi$. Around this mean, the variation across *bidders and periods* can be decomposed into two parts: the one that varies across *both* bidders and periods and the one *only* varies across bidders. The former is captured by the variance term τ^2 and the latter is represented by the random effect variance τ_{re}^2. With the assumptions 3 and 4, we can attribute τ^2 to the random shocks r_j^t that vary across *both* bidders and periods and τ_{re}^2 to r_j that *only* vary across bidders.

Given the normal distribution assumption of the random shock $R_j^t = r_j + r_j^t \sim N\left(0, \psi_1^2 + \psi_2^2\right)$, we may impute the y_j^{t*} (and hence b_j^t) for each combination of $\left(\mathbf{s}^t, \mathbf{X}^t, R_j^t\right)$, i.e., the full policy function. To see this, note that since $\sigma_j\left(\mathbf{s}^t, \mathbf{X}^t, R_j^t\right)$ is increasing in R_j^t,[32]

$$F\left(y_j^{t*}|\mathbf{s}^t, \mathbf{X}^t\right) = \Pr\left(\sigma_j\left(\mathbf{s}^t, \mathbf{X}^t, R_j^t\right) \leq y_j^{t*}|\mathbf{s}^t, \mathbf{X}^t\right) = \Phi\left(\sigma_j^{-1}\left(y_j^{t*}, \mathbf{s}^t, \mathbf{X}^t\right) / \sqrt{\psi_1^2 + \psi_2^2}\right)$$

[31] To be more specific, we estimate a continuous distribution $F\left(y_j^{t*}|\mathbf{s}^t, \mathbf{X}^t\right)$ for y_j^{t*} from equation A1; then conditional on the trunction point χ, we can back out the (discontinuous) distribution $F_b\left(b_j^t|\mathbf{s}^t, \mathbf{X}^t\right)$ for b_j^t.

[32] In this Appendix, we are abusing the notation of $\sigma_j\left(\mathbf{s}^t, \mathbf{X}^t, R_j^t\right)$. For the purpose of a clear exposition, we define $\sigma_j\left(\mathbf{s}^t, \mathbf{X}^t, R_j^t\right) = b_j^t$ in the paper. To match the bidding function estimated in equation A1, the more accurate definition should be

$$\begin{aligned} b_j^t &= \begin{cases} y_j^{t*} & \text{if } y_j^{t*} \geq \chi \\ 0 & \text{otherwise} \end{cases} \\ y_j^{t*} &= \sigma_j\left(\mathbf{s}^t, \mathbf{X}^t, R_j^t\right) \end{aligned}$$

where $\sigma_j^{-1}\left(y_j^{t*}, \mathbf{s}^t, \mathbf{X}^t\right)$ is the inverse function of $\sigma_j\left(\mathbf{s}^t, \mathbf{X}^t, R_j^t\right)$ with respect to R_j^t and $\Phi(\cdot)$ is the CDF of standard normal distribution. In equilibrium, we have $\sigma_j\left(\mathbf{s}^t, \mathbf{X}^t, R_j^t\right) = y_j^{t*}$. By substitution and rearrangement we get

$$\begin{aligned} y_j^{t*} &= \sigma_j\left(\mathbf{s}^t, \mathbf{X}^t, R_j^t\right) \\ &= F^{-1}\left(\Phi\left(\sigma_j^{-1}\left(y_j^{t*}, \mathbf{s}^t, \mathbf{X}^t\right)/\sqrt{\psi_1^2+\psi_2^2}\right)|\mathbf{s}^t, \mathbf{X}^t\right) \\ &= F^{-1}\left(\Phi\left(R_j^t/\sqrt{\psi_1^2+\psi_2^2}\right)|\mathbf{s}^t, \mathbf{X}^t\right) \\ &= F^{-1}\left(\Phi\left(\widetilde{R}_j^t/\sqrt{1+\tau^2/\tau_{re}^2}\right)|\mathbf{s}^t, \mathbf{X}^t\right) \end{aligned} \tag{A2}$$

where $\sigma_j^{-1}\left(y_j^{t*}, \mathbf{s}^t, \mathbf{X}^t\right) = R_j^t$; $R_j^t/\sqrt{\psi_1^2+\psi_2^2}$ has a standard normal distribution.

Therefore there is a unique mapping between the likelihood of observing a given valuation shock R_j^t and the y_j^{t*}. Each R_j^t drawn by a firm implies a corresponding quantile on the R_j^t's distribution; this quantile in turn implies a y_j^{t*} from the distribution represented by that firm's partial bidding function $\widetilde{\sigma}_j(s^t, X^t)$. However, because we do not know ψ_1 and ψ_2 and thus the distribution of R_j^t, we have to make draws from an alternative distribution $\widetilde{R}_j^t = \widetilde{r}_j + \widetilde{r}_j^t$ that has a one-one quantile mapping to R_j^t. To do this, we first draw the random effect $\widetilde{r}_j$ from $N(0,1)$ and keep it fixed for bidder j across periods. Next for period t, we draw a random shock $\widetilde{r}_j^t$ from some scaled normal distribution. To construct a proper one-one mapping of quantiles between the two distribution of R_j^t and $\widetilde{R}_j^t$, we need to make sure that the distribution of $\widetilde{r}_j^t$ is properly scaled so that $var(\widetilde{r}_j^t)/var(\widetilde{r}_j) = var(r_j^t)/var(r_j) = \psi_2^2/\psi_1^2$. Because of assumption 4, we know that $\psi_2^2/\psi_1^2 = \tau^2/\tau_{re}^2$. Hence we should draw $\widetilde{r}_j^t \sim N(0, \tau^2/\tau_{re}^2)$. Note that now $\widetilde{R}_j^t = \widetilde{r}_j + \widetilde{r}_j^t$ is following a distribution $N(0, 1+\tau^2/\tau_{re}^2)$. Therefore $\widetilde{R}_j^t/\sqrt{1+\tau^2/\tau_{re}^2}$ has the same distribution of $N(0,1)$ as $R_j^t/\sqrt{\psi_1^2+\psi_2^2}$. Further, $\widetilde{R}_j^t$ is properly scaled such that the quantiles of $\widetilde{R}_j^t$ and R_j^t are uniquely mapped. We then compute the likelihood of $\widetilde{R}_j^t/\sqrt{1+\tau^2/\tau_{re}^2}$ as $\Phi(\widetilde{R}_j^t/\sqrt{1+\tau^2/\tau_{re}^2})$. Next, we determine $F\left(y_j^{t*}|\mathbf{s}^t, \mathbf{X}^t\right)$ using results estimated in A1 and looking at the distribution of its residuals to determine F. That is, for each value of y_j^{t*} we should be able to compute its probability for a given $\mathbf{s}^t$ and $\mathbf{X}^t$ using F. Accordingly, F^{-1} links probabilities to y_j^{t*} (therefore b_j^t) for a given $\mathbf{s}^t$ and $\mathbf{X}^t$. We then use F^{-1} to link the probability $\Phi(\widetilde{R}_j^t/\sqrt{1+\tau^2/\tau_{re}^2})$ to b_j^t for a particular $\mathbf{s}^t$ and $\mathbf{X}^t$. In this manner we ensure the bids and valuations in equation (A11) comport. In Appendix A.2.1, when evaluating the value function for a set of given parameter values of ψ_1 and ψ_2 in equation A11 or evaluating base functions defined in A12, we integrate out over the unobserved shocks R_j^t by drawing many $\widetilde{r}_j$ and $\widetilde{r}_j^t$ from $N(0,1)$ and $N(0, \tau^2/\tau_{re}^2)$, respectively.

A.1.2 Consumer Model Estimation

We derive the consumer model conditioned on the information state of the advertiser as described in section 3.1. Given advertisers do not observe what each person downloaded or the characteristics of these persons, it must infer consumer behavior from aggregate instead of individual level data.

Advertisers do observe the aggregate data in the form of download counts $\mathbf{d}^t = \{d_1^t, d_2^t, ..., d_N^t\}$ in period t. A single d_j^t follow a binomial distribution. Given the individual level download probabilities P_{ijt} in equation 12, a single d_j^t's probability mass function is $\begin{pmatrix} M_t \\ d_j^t \end{pmatrix} [P_{ijt}]^{d_j^t}[1-P_{ijt}]^{M_t - d_j^t}$, where M_t is the consumer population size in period t. Hence the likelihood of observing $\mathbf{d}^t$ is

$$L(\mathbf{d}^t|\Omega_c) = \prod\nolimits_j \begin{pmatrix} M_t \\ d_j^t \end{pmatrix} [P_{ijt}]^{d_j^t}[1-P_{ijt}]^{M_t - d_j^t}$$

where $\Omega_c \equiv \{\alpha^g, \boldsymbol{\beta}^g, \delta^g, \gamma^g, \lambda_0^g, \lambda_1^g\}_g$ are parameters to be estimated.

Naturally, the full posterior distribution of the model will be the product of $L(\mathbf{d}^t|\Omega_c)$ across periods and $p(\Omega_c)$, the prior distributions of parameters, i.e.,

$$p(\Omega_c|data) \propto \prod_t L(\mathbf{d}^t|\Omega_c) \cdot p(\Omega_c) \tag{A3}$$

An advertiser's predicted downloads $d(k, X_j^t; \Omega_c)$ can readily be constructed using the parameter estimates as shown in equation 13

$$d(k, X_j^t; \widehat{\Omega}_c) = M_t \widehat{P}_{ijt} \tag{A4}$$

This prediction is then used to forecast expectations of future downloads and slot positions in the firm's value function in the second step estimation.

A.1.3 State Transition Function $P\left(\mathbf{s}^{'}|b_j, \mathbf{b}_{-j}, \mathbf{s}, X\right)$

To compute the state transition, note that the marginal number of expected downloads is given by the expected downloads given a slot position multiplied by the probability of appearing in that slot position and then summed across all positions:

$$P\left(\mathbf{s}^{'}|b_j, \mathbf{b}_{-j}, \mathbf{s}, \mathbf{X}\right) = \sum\nolimits_k d(k, X; \Omega_c) \Pr\left(k|b_j, \mathbf{b}_{-j}, \mathbf{s}, \mathbf{X}\right) \tag{A5}$$

The expected downloads given a slot position in A5 is defined in 13. We can decompose the likelihood of appearing in slot k as follows

$$\begin{aligned} & \Pr(k|b_j, \mathbf{b}_{-j}, \mathbf{s}, \mathbf{X}) \\ = \ & \Pr_{\{k\leq K\}}(k|b_j, \mathbf{b}_{-j}, \mathbf{s}, \mathbf{X})\, I\{k \leq K\} + \Pr_{\{k>K\}}(k|b_j, \mathbf{b}_{-j}, \mathbf{s}, \mathbf{X}) I\{k > K\} \end{aligned} \tag{A6}$$

where $\Pr_{\{k\leq K\}}(k|b_j, \mathbf{b}_{-j}, \mathbf{s}, \mathbf{X})$ is the probability of appearing in slot k of the sponsored search section (i.e., $k \leq K$); and $\Pr_{\{k>K\}}(k|b_j, \mathbf{b}_{-j}, \mathbf{s}, \mathbf{X})$ is the likelihood of appearing in slot k of the organic search section (i.e., $k > K$). We discuss these two probabilities next.

Likelihood of Premium Slot $k \leq K$ Let us first consider the likelihood of winning one of the premium slots k ($k \leq K$), $\Pr_{\{k\leq K\}}(k|b_j, \mathbf{b}_{-j}, \mathbf{s}, \mathbf{X})$ as an order statistic reflecting the relative quality of the advertiser's bid, which is defined as $b_j d_j^{(-1)}$. Higher quality bids are more likely to be assigned to better slots. Denote $\Psi_{bd}(b_{j'} d_{j'}^{(-1)}|\mathbf{s}, \mathbf{X})$ as the distribution CDF of $b_{j'} d_{j'}^{(-1)}, \forall j'$, where $d_{j'}^{(-1)}$ is from the state vector and $b_{j'}$ has a distribution depending on the strategy profile $\boldsymbol{\sigma}(\cdot)$.[33] For bidder j to win a premium slot k by bidding b_j, it implies that (1) among all of the other $N-1$ competing bidders, there are $k-1$ bidders have a higher ranking than j in terms of $b_{j'} d_{j'}^{(-1)}$ and (2) the other ones have a lower ranking than j. The probability of having a higher ranking than j is $[1 - \Psi_{bd}(b_j d_j^{(-1)}|s, X)]$. Thus the probability of bidder j winning slot k by bidding b_j is simply an order statistics as shown below; note that the combination $\begin{pmatrix} N-1 \\ k-1 \end{pmatrix}$ in the equation is because any $(k-1)$ out of the $(N-1)$ competing bidders can have a higher ranking than j.[34]

$$\begin{aligned} & \Pr_{\{k\leq K\}}(k|b_j, \mathbf{b}_{-j}, \mathbf{s}, \mathbf{X}) \\ = \ & \begin{pmatrix} N-1 \\ k-1 \end{pmatrix} [1 - \Psi_{bd}(b_j d_j^{(-1)}|\mathbf{s}, \mathbf{X})]^{k-1} [\Psi_{bd}(b_j d_j^{(-1)}|\mathbf{s}, \mathbf{X})]^{(N-1)-(k-1)} \\ = \ & \begin{pmatrix} N-1 \\ k-1 \end{pmatrix} [1 - \Psi_{bd}(b_j d_j^{(-1)}|\mathbf{s}, \mathbf{X})]^{k-1} [\Psi_{bd}(b_j d_j^{(-1)}|\mathbf{s}, \mathbf{X})]^{N-k} \end{aligned} \tag{A7}$$

Likelihood of Organic Slot $k > K$ Next we consider what happens when an advertiser does not win this auction and is placed in the organic search section. In this case, by the rules of the auction, the bidder's slot is determined by its update recency compared to all products in the organic search section. For bidder j to be placed in organic slot $k > K$ it implies that (1) there are

[33] It is difficult to write a closed form solution for Ψ_{bd} but we may use the sample population distribution to approximate Ψ_{bd}.

[34] An alternative interpretation of equation A7 is the probability mass function (PMF) of a binomial distribution. Among $N-1$ competing bidders, there are $k-1$ higher than bidder j and $(N-1)-(k-1)$ lower than j; and the probability of higher than j is $[1 - \Psi_{bd}(b_j d_j^{(-1)}|\mathbf{s}, \mathbf{X})]$. Hence we may consider the expression in A7 as the PMF of a binomial distribution.

K bidders have a higher ranking of $b_{j'}d_{j'}^{(-1)}$ than bidder j (i.e., j loses the auction) and (2) among the other $\overline{N}-K-1$ products (i.e., all products at the search engine less those who win premium slots and j itself), there are $k-K-1$ products have a higher update recency than j and (3) the other ones have a lower ranking than j. Hence,

$$\begin{aligned} & \Pr_{\{k>K\}}(k|b_j,\mathbf{b}_{-j},\mathbf{s},\mathbf{X}) \\ = \ & \Pr(k>K|b_j,\mathbf{b}_{-j},\mathbf{s},\mathbf{X})\cdot\Pr(k|b_j,\mathbf{b}_{-j},\mathbf{s},\mathbf{X},k>K) \end{aligned} \tag{A8}$$

where the first term is the probability of losing the auction (condition 1) and the second term denotes the likelihood of appearing in position $k>K$ (condition 2 and 3). Note that the main reason for the difference between A7 and A8 is the change of ranking mechanisms. The ranking is based on $b_{j'}d_{j'}^{(-1)}$ for $k\leq K$ and update recency when $k>K$. The first term in A8 does not appear as an order statistics (as shown below) since when $k>K$ the order of $b_{j'}d_{j'}^{(-1)}$ becomes meaningless. Instead, the update recency is affecting the ranking. The two terms in A8 can be expressed as follows.

Losing the auction implies that among j's $N-1$ opponents, there are K bidders have a higher ranking than j in terms of $b_{j'}d_{j'}^{(-1)}$. Hence,

$$\Pr(k>K|b_j,\mathbf{b}_{-j},\mathbf{s},\mathbf{X})=\begin{pmatrix} N-1 \\ K \end{pmatrix}[1-\Psi_{bd}(b_jd_j^{(-1)}|\mathbf{s},\mathbf{X})]^K \tag{A9}$$

The conditional probability of being placed in an organic slot $k>K$ (condition 2 and 3) is again an order statistics.[35] This distribution is incumbent upon the update recency of all $\overline{N}$ products exclusive of the K winners in the sponsored search section. Denoting the distribution of update recency of all products as Ψ_{up}, which can be approximated from the sample population distribution observed in the data, we obtain the following:

$$\begin{aligned} & \Pr(k|b_j,\mathbf{b}_{-j},\mathbf{s},\mathbf{X},k>K) \\ = & \begin{pmatrix} \overline{N}-K-1 \\ k-K-1 \end{pmatrix}[1-\Psi_{up}]^{k-K-1}[\Psi_{up}]^{(\overline{N}-K-1)-(k-K-1)} \\ = & \begin{pmatrix} \overline{N}-K-1 \\ k-K-1 \end{pmatrix}[1-\Psi_{up}]^{k-K-1}[\Psi_{up}]^{\overline{N}-k} \end{aligned} \tag{A10}$$

Combining A10 and A9 into A8, and then A8 and A7 into A6 yields the state transition equation.

Given that we have detailed the estimation of the first step functions ($\sigma_j(\mathbf{s},\mathbf{X},R_j)$, $d(k,X_j^t;\Omega_c)$, $P(\mathbf{s}'|\mathbf{b},\mathbf{s},\mathbf{X})$), we now turn to the second step estimator, which is incumbent upon these first step functions.

[35] This order statistics can again be interpreted as the PMF of a binomial distribution similar to A7.

A.2 Second Step Estimation of Bidder Model

In this Appendix we detail how to estimate the parameters in the value function. This is done in two phases; first we simulate the value function conditioned on Ω_a and second we construct the likelihood using the simulated value function conditioned on Ω_a.

A.2.1 Phase 1: Simulation of Value Functions Given Ω_a

To construct the value function we first simplify its computation by linearization and, second, using this simplification we simulate the expected value function conditioned on Ω_a by integrating out over draws for $\mathbf{s}^t$, $\mathbf{X}^t$, and $(\widetilde{r}_j + \widetilde{r}_j^t)$.

Linearize the Value Function We simplify the estimation procedure by relying on the fact that equation 16 is linear in the parameters Ω_a. We can rewrite equation 16 by factoring out Ω_a.

$$\begin{aligned}
& \mathbf{E}\pi_j\left(\mathbf{b}^t, \mathbf{s}^t, \mathbf{X}^t, R_j^t; \theta\right) \qquad \text{(A11)} \\
= \ & \sum\nolimits_{k=1}^{K} \Pr\left(k | b_j^t, \mathbf{b}_{-j}^t, \mathbf{s}^t, \mathbf{X}^t\right) \cdot (v(X_j^t;\theta) + r_j + r_j^t - b_j^t) \cdot d(k, X_j^t; \Omega_c) \\
& + \sum\nolimits_{k=K+1}^{\overline{N}} \Pr\left(k | b_j^t, \mathbf{b}_{-j}^t, \mathbf{s}^t, \mathbf{X}^t\right) \cdot (v(X_j^t;\theta) + r_j + r_j^t) \cdot d(k, X_j^t; \Omega_c) \\
= \ & \left[\sum\nolimits_{k=1}^{\overline{N}} \Pr\left(k | b_j^t, \mathbf{b}_{-j}^t, \mathbf{s}^t, \mathbf{X}^t\right) \cdot d(k, X_j^t; \Omega_c) \cdot X_j^t\right] \cdot \theta \\
& + \left[\sum\nolimits_{k=1}^{\overline{N}} \Pr\left(k | b_j^t, \mathbf{b}_{-j}^t, \mathbf{s}^t, \mathbf{X}^t\right) \cdot d(k, X_j^t; \Omega_c) \cdot \widetilde{r}_j\right] \cdot \psi_1 \\
& + \left[\sum\nolimits_{k=1}^{\overline{N}} \Pr\left(k | b_j^t, \mathbf{b}_{-j}^t, \mathbf{s}^t, \mathbf{X}^t\right) \cdot d(k, X_j^t; \Omega_c) \cdot \widetilde{r}_j^t\right] \cdot \psi_2 \\
& - b_j^t \sum\nolimits_{k=1}^{K} \Pr\left(k | b_j^t, \mathbf{b}_{-j}^t, \mathbf{s}^t, \mathbf{X}^t\right) d(k, X_j^t; \Omega_c) \\
= \ & Base_{j1}^t \theta + Base_{j2}^t \psi - Base_{j3}^t
\end{aligned}$$

where

$$\begin{aligned}
Base_{j1}^t \ & \equiv \ \left[\sum\nolimits_{k=1}^{\overline{N}} \Pr\left(k | b_j^t, \mathbf{b}_{-j}^t, \mathbf{s}^t, \mathbf{X}^t\right) \cdot d(k, X_j^t; \Omega_c) \cdot X_j^t\right] \qquad \text{(A12)} \\
Base_{j2}^t \ & \equiv \ \begin{bmatrix} \sum_{k=1}^{\overline{N}} \Pr\left(k | b_j^t, \mathbf{b}_{-j}^t, \mathbf{s}^t, \mathbf{X}^t\right) \cdot d(k, X_j^t; \Omega_c) \cdot \widetilde{r}_j \\ \sum_{k=1}^{\overline{N}} \Pr\left(k | b_j^t, \mathbf{b}_{-j}^t, \mathbf{s}^t, \mathbf{X}^t\right) \cdot d(k, X_j^t; \Omega_c) \cdot \widetilde{r}_j^t \end{bmatrix}' \\
Base_{j3}^t \ & \equiv \ b_j^t \sum\nolimits_{k=1}^{K} \Pr\left(k | b_j^t, \mathbf{b}_{-j}^t, \mathbf{s}^t, \mathbf{X}^t\right) d(k, X_j^t; \Omega_c) \\
\widetilde{r}_j \ & \sim \ N(0,1) \quad \widetilde{r}_j^t \sim N(0, \tau^2/\tau_{re}^2)
\end{aligned}$$

Note that the values of $\left\{Base_{j1}^t, Base_{j2}^t, Base_{j3}^t\right\}_{\forall t}$ are conditionally independent of θ and ψ. This enables us to first evaluate $\left\{Base_{j1}^t, Base_{j2}^t, Base_{j3}^t\right\}_{\forall t}$ and keep them constant when drawing θ and ψ from their posterior distributions. By doing so, we reduce the computational burden of estimation as described next.

Simulate the Value Functions Given Ω_a After the linearization, given a set of advertiser parameters $\Omega_a = \{\theta, \psi_1, \psi_2\}$ and equation A11, the value function depicted in equation 18 can also be written as the following with period index t invoked:

$$\begin{aligned} V_j\left(\mathbf{s}^0, \mathbf{X}^0; \boldsymbol{\sigma}; \Omega_a\right) &= \mathbf{E}_{\mathbf{s},\mathbf{X},R}\left[\sum_{t=0}^{\infty} \rho^t \pi_j\left(\boldsymbol{\sigma}, \mathbf{s}^t, \mathbf{X}^t, R_j^t; \Omega_a\right)\right] \\ &= \mathbf{E}[\sum_{t=0}^{\infty}(\rho^t Base_{j1}^t \theta + Base_{j2}^t \psi - Base_{j3}^t)] \\ &= [\mathbf{E}\sum_{t=0}^{\infty} \rho^t Base_{j1}^t]\theta + [\mathbf{E}\sum_{t=0}^{\infty} \rho^t Base_{j2}^t]\psi - [\mathbf{E}\sum_{t=0}^{\infty} \rho^t Base_{j3}^t] \end{aligned} \tag{A13}$$

where the expectation is taken over current and future private shocks, future states $\mathbf{s}^t$, future $\mathbf{X}^t$ and R^t.

An estimated value function $\widehat{V}_j\left(\mathbf{s}^0, \mathbf{X}^0; \boldsymbol{\sigma}; \Omega_a\right)$ can then be obtained by the following steps:

1. Draw $\widetilde{r}_j$ from $N(0,1)$ for all bidder j and keep $\widetilde{r}_j$ fixed for all periods;

2. Draw private shocks $\widetilde{r}_j^t$ from $N(0, \tau^2/\tau_{re}^2)$ for all bidders j in period 0; draw initial choice of $\mathbf{s}^0$ from the distribution of state variables derived from the observed data; draw $\mathbf{X}^0$ from the observed distribution of product attributes.

3. Starting with the initial state $\mathbf{s}^0$, $\mathbf{X}^0$ and the $(\widetilde{r}_j + \widetilde{r}_j^0)$ from step 1-2, calculate $\widehat{b}_j^0$ for all bidders using the inversion (equation A2) described in Appendix A.1.1.

4. Use $\mathbf{s}^0$, $\mathbf{X}^0$ and $\widehat{\mathbf{b}}^0$ to determine the slot ranking, whose distribution is $\Pr\left(k|b_j^t, \mathbf{b}_{-j}^t, \mathbf{s}^t, \mathbf{X}^t\right)$ in equation A6 in Appendix A.1.3; using $d(k, X_j^0; \Omega_c)$ in equation (13), obtain a new state vector $\mathbf{s}^1$, whose distribution is $P(\mathbf{s}^1|\widehat{\mathbf{b}}^0, \mathbf{s}^0, \mathbf{X}^0)$ in equation A5 in Appendix A.1.3; draw $\mathbf{X}^1$ from the observed distribution of product attributes.

5. Repeat step 2-4 for T periods for all bidders to compute all $\mathbf{s}^t$, $\mathbf{X}^t$, $(\widetilde{r}_j + \widetilde{r}_j^t)_{\forall j}$, and $\mathbf{b}^t$ for all periods; T is large enough so that the discount factor ρ^T approaches 0.

6. Using $\mathbf{s}^t$, $\mathbf{X}^t$, $(\widetilde{r}_j + \widetilde{r}_j^t)_{\forall j}$, $d(k, X_j^t; \Omega_c)$ and $\mathbf{b}^t$, evaluate $\left\{Base_{j1}^t, Base_{j2}^t, Base_{j3}^t\right\}_{t=0,\ldots,T}$ and $\left\{[\sum_{t=0}^{T} \rho^t Base_{j1}^t], [\sum_{t=0}^{T} \rho^t Base_{j2}^t], [\sum_{t=0}^{T} \rho^t Base_{j3}^t]\right\}$.

7. The resulting values of

$$\left\{Base^t_{j1}, Base^t_{j2}, Base^t_{j3}\right\}_{t=0,\dots,T}$$

and

$$\left\{[\sum_{t=0}^{T}\rho^t Base^t_{j1}], [\sum_{t=0}^{T}\rho^t Base^t_{j2}], [\sum_{t=0}^{T}\rho^t Base^t_{j3}]\right\}$$

depend on the random draws of $\mathbf{s}^t, \mathbf{X}^t, R^t$. To compute

$$\left\{[\mathbf{E}\sum_{t=0}^{\infty}\rho^t Base^t_{j1}], [\mathbf{E}\sum_{t=0}^{\infty}\rho^t Base^t_{j2}], [\mathbf{E}\sum_{t=0}^{\infty}\rho^t Base^t_{j3}]\right\},$$

repeat step 1-6 for NR times so as to integrate out over the draws. Note that when T is large enough $[E\sum_{t=0}^{T}\rho^t Base^t_{j.}]$ is a good approximation of $[E\sum_{t=0}^{\infty}\rho^t Base^t_{j.}]$ since ρ^T approaches 0.

8. Conditional on a set of parameters $\Omega_a = \{\theta, \psi'\}$ and

$$\left\{[\mathbf{E}\sum_{t=0}^{\infty}\rho^t Base^t_{j1}], [\mathbf{E}\sum_{t=0}^{\infty}\rho^t Base^t_{j2}], [\mathbf{E}\sum_{t=0}^{\infty}\rho^t Base^t_{j3}]\right\},$$

we may evaluate $\widehat{V}_j\left(\mathbf{s}^0, \mathbf{X}^0; \boldsymbol{\sigma}; \Omega_a\right)$ from (A13).

An estimated deviation value function $\widehat{V}_j\left(\mathbf{s}^0, \mathbf{X}^0; \sigma'_j, \boldsymbol{\sigma}_{-j}; \Omega_a\right)$ with an alternative strategy σ'_j other than σ_j can be constructed by following the same procedure. We draw a deviated strategy σ'_j by adding disturbance to the estimated policy function from Step 1. In particular, we add a normally distributed random variable ($mean = 0$; $s.d. = 0.3$) to each parameter.

We implement this process by first drawing $NS = 10$ initial states for each bidder and $\{X^t\}_{t=0,1,\dots,T}$ of all $T = 200$ periods. Then for each combination of bidder and initial state, we use this process to compute the base value functions and $ND = 100$ perturbed base functions. In Step 6, we use $NR = 100$. The discount factor ρ is fixed as 0.95.

The computational burden is reduced tremendously since we have linearized the value functions and factored out the parameters Ω_a. We do not need to re-evaluate the value functions for each set of parameters Ω_a. Instead, we only evaluate the base functions in equation A12 once using step 1-7 and keep them fixed. Then for each draw of Ω_a from the posterior distribution we may evaluate the value functions (step 8) so as to recover Ω_a as described below.

A.2.2 Phase 2: Recover Ω_a

Recall our goal is to assess the likelihood that 19 holds. Define $P_j\left(\mathbf{s}^0_{(ns)}, \mathbf{X}^0; \boldsymbol{\sigma}, \sigma'_j; \Omega_a\right)$ as the probability of the event

$$\left\{\widehat{V}_j(\mathbf{s}^0_{(ns)}, \mathbf{X}^0; \sigma_j, \boldsymbol{\sigma}_{-j}; \Omega_a) \geq \widehat{V}_j(\mathbf{s}^0_{(ns)}, \mathbf{X}^0; \sigma'_j, \boldsymbol{\sigma}_{-j}; \Omega_a)\right\}, \tag{A14}$$

where $s^0_{(ns)}$ stands for the ns-th initial state of bidder j. This event means that the estimated value function for the given initial state $s^0_{(ns)}$ with observed strategy σ_j is greater than the estimated value function with a deviation σ'_j. For observed data to be rational, we should have $P_j\left(\mathbf{s}^0_{(ns)}, \mathbf{X}^0; \boldsymbol{\sigma}, \sigma'_j; \Omega_a\right)$ converging to 1 under the true parameters, in the sense that all ND draws should result in the event of equation A14.

Note that $P_j\left(\mathbf{s}^0_{(ns)}, \mathbf{X}^0; \boldsymbol{\sigma}, \sigma'_j; \Omega_a\right)$ is not observed but it can be approximated with the sample analog from the simulated ND draws of $\widehat{V}_j(s^0_{(ns)}, X^0; \sigma'_j, \sigma_{-j}; \Omega_a)$ as the follows:

$$\begin{aligned} & \widehat{P}_j\left(\mathbf{s}^0_{(ns)}, \mathbf{X}^0; \boldsymbol{\sigma}, \sigma'_j; \Omega_a\right) \\ = \; & \frac{1}{ND}\sum_{nd=1}^{ND} I\left\{\widehat{V}_j(\mathbf{s}^0_{(ns)}, \mathbf{X}^0; \sigma_j, \boldsymbol{\sigma}_{-j}; \Omega_a) \geq \widehat{V}_j(\mathbf{s}^0_{(ns)}, \mathbf{X}^0; \sigma'_j, \boldsymbol{\sigma}_{-j}; \Omega_a)_{(nd)}\right\} \end{aligned} \tag{A15}$$

where the subscript (nd) indices the nd-th simulated $\widehat{V}_j(s^0_{(ns)}, X^0; \sigma'_j, \sigma_{-j}; \Omega_a)$.

By pooling together all $\widehat{P}_j\left(\mathbf{s}^0_{(ns)}, \mathbf{X}^0; \boldsymbol{\sigma}, \sigma'_j; \Omega_a\right)$'s across bidders and (ns), we are able to construct the likelihood function as the following

$$L = \prod_{j,(ns)} \widehat{P}_j\left(\mathbf{s}^0_{(ns)}, \mathbf{X}^0; \boldsymbol{\sigma}, \sigma'_j; \Omega_a\right) \tag{A16}$$

Denote the prior of Ω_a as $p(\Omega_a)$, the posterior can be written as

$$p(\Omega_a|data) \propto \prod_{j,(ns)} \widehat{P}_j\left(\mathbf{s}^0_{(ns)}, \mathbf{X}^0; \boldsymbol{\sigma}, \sigma'_j; \Omega_a\right) p(\Omega_a) \tag{A17}$$

B Sampling Chain

B.1 Advertiser Model

B.1.1 Priors

The advertiser model is specified as

$$\begin{aligned} b^t_j &= \begin{cases} y^{t*}_j & \text{if } y^{t*}_j \geq \chi \\ 0 & \text{otherwise} \end{cases} \\ y^{t*}_j &\sim N([\mathbf{s}^{t\prime}, \mathbf{X}^{t\prime}_j] \cdot \varphi + \varphi^{re}_j, \tau^2) \\ \varphi^{re}_j &\sim N(0, \tau^2_{re}) \end{aligned} \tag{A18}$$

We iterate the sampling chain for 20,000 and use the second half of the chain to make inference. The priors use a diffused variance of 100; examinations of the posteriors shows that the choice of the variance is the order of magnitude greater than posterior distributions, which assures a proper but diffused prior (Spiegelhalter et al. (1996), Gelman et al. (2004)).

	Priors	Selected Value
φ	$\varphi \sim N\left(\varphi_0, I_d\sigma_\varphi^2\right)$	φ_0: estimates of an classical Tobit model of bid b_j^t on $[\mathbf{s}^{t\prime}, \mathbf{X}_j^{t\prime}]$ with the truncation at 15. $\sigma_\varphi^2 = 100$ I_d is an identity matrix with the dimension of the number of covariates vector $[\mathbf{s}^{t\prime}, \mathbf{X}_j^{t\prime}]$.
τ_{re}	$\tau_{re} \sim TN_{(0,+\infty)}(\mu_{\tau 1}, \sigma_{\tau 1}^2)$[36]	$\mu_{\tau 1} = 5, \sigma_{\tau 1}^2 = 100$
τ	$\tau \sim TN_{(0,+\infty)}(\mu_{\tau 2}, \sigma_{\tau 2}^2)$	$\mu_{\tau 2} = 5, \sigma_{\tau 2}^2 = 100$

B.1.2 Conditional Posteriors

- To facilitate explication denote the vector $[\mathbf{s}^{t\prime}, \mathbf{X}_j^{t\prime}] \equiv Z_j^t$, the matrix $[Z_j^t]_{\forall j,t}$ as Z and the vector $[y_j^{t*}]_{\forall j,t} \equiv y^*$. We also denote the number of bidders as N and the number of observations for bidder j as N_j. So the total number of observations is $\sum_j N_j$; the dimension of Z_j^t is 1 by d (the dimension of $[s^{t\prime}, X_j^{t\prime}]$); the dimension of Z is $\sum_j N_j$ by d; the dimension of y^* is $\sum_j N_j$. The vector of bidder specific random effects, φ^{re}, is a column vector with a length of $\sum_j N_j$. The φ^{re} is constructed as N stacked sub-vectors, where the j-th sub-vector has bidder j's φ_j^{re} as its elements and a dimension of N_j.

- y_j^{t*}

  y_j^{t*} is determined by the following

$$\begin{aligned} y_j^{t*} &= b_j^t, \text{ if } b_j^t > 0 \\ y_j^{t*} &\sim TN_{(-\infty,15]}(Z_j^t\varphi + \varphi_j^{re}, \tau^2), \text{ if } b_j^t = 0 \end{aligned}$$

  y_j^{t*} is right truncated at 15 when $b_j^t = 0$; this is consistent with the 15¢ minimum bid requirement of the search engine.

- φ_j^{re}

[36] We do not use Gamma distribution as the prior. Natarajan and McCulloch (1998) show a diffused proper prior such as Gamma distribution can sometimes lead to inaccurate inference due to the long tail of the Gamma distribution. In our application, we first tried Gamma distribution as the prior and got poor quality mixing. There were some unrealistic large draws for τ_{re} in the order greater than 100. So we instead adopted the truncated normal prior.

φ_j^{re}'s are treated as latent variables. Bidder j has the same φ_j^{re} across all N_j observations. The conditional likelihood of latent φ_j^{re} is

$$L \propto \prod^{N_j} \exp(-\frac{[y_j^{t*} - (Z_j^t \varphi + \varphi_j^{re})]^2}{2\tau^2})$$

Hence in each iteration we draw φ_j^{re} from the following distribution

$$\begin{aligned}
\varphi_j^{re} &\sim N(\widehat{\mu}_j^{re}, \widehat{\Sigma}_j^{re}) \\
\widehat{\Sigma}_j^{re} &= (N_j \tau^{-2} + \tau_{re}^{-2})^{-1} \\
\widehat{\mu}_j^{re} &= \widehat{\Sigma}_j^{re} \tau^{-2} \sum_t (y_j^{t*} - Z_j^t \varphi)
\end{aligned}$$

- φ

$$\begin{aligned}
\text{Prior } \varphi &\sim N\left(\varphi_0, I_d \sigma_\varphi^2\right) \\
\text{Likelihood } L &\propto \prod_{j,t} \exp(-\frac{[y_j^{t*} - (Z_j^t \varphi + \varphi_j^{re})]^2}{2\tau^2}) \\
\text{Posterior } (\varphi|\cdot) &\sim N(\mu_\varphi, \Sigma_\varphi) \\
\Sigma_\varphi &= [Z'Z\tau^{-2} + \sigma_\varphi^{-2}]^{-1} \\
\mu_\varphi &= \Sigma_\varphi \cdot \{[y^* - \varphi^{re}]' Z \tau^{-2} + \varphi_0 \sigma_\varphi^{-2}\}
\end{aligned} \tag{A19}$$

- τ_{re}

$$\text{Prior } \tau_{re} \sim TN_{(0,+\infty)}(\mu_{\tau 1}, \sigma_{\tau 1}^2)$$

A random walk proposal density is used in the (r)-th iteration, $\tau_{re}^{(r)} \sim TN_{(0,+\infty)}(\tau_{re}^{(r-1)}, \sigma_{p1}^2)$, where $\tau_{re}^{(r-1)}$ is the value from the $(r-1)$-th iteration; σ_{p1}^2 is the variance which is tuned so that the acceptance rate is between $15\% - 50\%$.

The acceptance probability $pr^* = \min(1, pr)$ and

$$pr = \frac{L(\tau_{re}^{(r)}|\cdot) p(\tau_{re}^{(r)}|\mu_{\tau 1}, \sigma_{\tau 1}^2) \eta(\tau_{re}^{(r-1)}|\tau_{re}^{(r)}, \sigma_{p1}^2)}{L(\tau_{re}^{(r-1)}|\cdot) p(\tau_{re}^{(r-1)}|\mu_{\tau 1}, \sigma_{\tau 1}^2) \eta(\tau_{re}^{(r)}|\tau_{re}^{(r-1)}, \sigma_{p1}^2)}$$

where $p(\tau_{re}^{(\cdot)}|\mu_{\tau 1}, \sigma_{\tau 1}^2)$ is the density of $\tau_{re}^{(\cdot)}$ evaluated using the prior. $L(\tau_{re}^{(\cdot)}|\cdot)$ is the likelihood evaluated at $\tau_{re}^{(\cdot)}$. In particular,

$$L(\tau_{re}^{(\cdot)}|\cdot) \propto \prod_j \pi(\varphi_j^{re}; 0, \tau_{re}^{(\cdot)})$$

where $\pi(\varphi_j^{re}; 0, \tau_{re}^{(\cdot)})$ is the normal density of the random effect φ_j^{re} evaluated with mean 0 and standard deviation $\tau_{re}^{(\cdot)}$.

$\frac{\eta(\tau_{re}^{(r-1)}|\tau_{re}^{(r)}, \sigma_{p1}^2)}{\eta(\tau_{re}^{(r)}|\tau_{re}^{(r-1)}, \sigma_{p1}^2)}$ is used as a weight to correct the acceptance probability as the proposal density is truncated at 0 and, hence, asymmetric.

- τ

$$\text{Prior } \tau \sim TN_{(0,+\infty)}(\mu_{\tau 2}, \sigma_{\tau 2}^2)$$

A random walk proposal density is used in the (r)-th iteration, $\tau^{(r)} \sim TN_{(0,+\infty)}(\tau^{(r-1)}, \sigma_{p2}^2)$, where $\tau^{(r-1)}$ is the value from the $(r-1)$-th iteration; σ_{p2}^2 is the tuning variance.

The acceptance probability $pr^* = \min(1, pr)$ and

$$pr = \frac{L(\tau^{(r)}|\cdot)p(\tau^{(r)}|\mu_{\tau 2}, \sigma_{\tau 2}^2)\eta(\tau^{(r-1)}|\tau^{(r)}, \sigma_{p2}^2)}{L(\tau^{(r-1)}|\cdot)p(\tau^{(r-1)}|\mu_{\tau 2}, \sigma_{\tau 2}^2)\eta(\tau^{(r)}|\tau^{(r-1)}, \sigma_{p2}^2)}$$

where $p(\tau^{(\cdot)}|\mu_{\tau 2}, \sigma_{\tau 2}^2)$ is the density of $\tau^{(\cdot)}$ evaluated using the prior. $L(\tau^{(\cdot)}|\cdot)$ is the likelihood evaluated at $\tau^{(\cdot)}$. To be specific

$$L(\tau^{(\cdot)}|\cdot) \propto \prod_{j,t} \pi(y_j^{t*}; Z_j^t \varphi + \varphi_j^{re}, \tau^{(\cdot)})$$

where $\pi(y_j^{t*}; Z_j^t \varphi + \varphi_j^{re}, \tau^{(\cdot)})$ is the normal density of y_j^{t*} evaluated with mean $Z_j^t \varphi + \varphi_j^{re}$ and standard deviation $\tau^{(\cdot)}$.

B.2 Consumer Model

[To be completed.]

B.3 Second Step Estimation

B.3.1 Priors

	Priors	Selected Value
θ	$\theta \sim N\left(\theta_0, I_X \sigma_\theta^2\right)$	θ_0: a vector of zeros with the length of the number of product attributes. $\sigma_\theta^2 = 100$ I_X is an identity matrix with the dimension of the number of product attributes.
ψ	$\psi \sim TN_{(0,+\infty)}(\mu_\psi, I_2 \sigma_\psi^2)$, where $\psi' = \{\psi_1, \psi_2\}$	$\mu_\psi' = \{1, 1\}, \sigma_\psi^2 = 100$, I_2 is an identity matrix with the dimension of 2.

B.3.2 Conditional Posteriors

- θ

$$\text{Prior } \theta \sim N\left(\theta_0, I_X\sigma_\theta^2\right)$$

A random walk proposal density is used in the (r)-th iteration, $\theta^{(r)} \sim N(\theta^{(r-1)}, I_X\sigma_{p\theta}^2)$, where $\theta^{(r-1)}$ is the value from the $(r-1)$-th iteration; $\sigma_{p\theta}^2$ is a scalar and functions as the tuning variance.

The acceptance probability $pr^* = \min(1, pr)$ and

$$pr = \frac{L(\theta^{(r)}|\cdot)p(\theta^{(r)}|\theta_0, I_X\sigma_\theta^2)}{L(\theta^{(r-1)}|\cdot)p(\theta^{(r-1)}|\theta_0, I_X\sigma_\theta^2)}$$

where $p(\theta^{(\cdot)}|\theta_0, I_X\sigma_\theta^2)$ is the density of $\theta^{(\cdot)}$ evaluated using the prior. $L(\theta^{(\cdot)}|\cdot)$ is the likelihood evaluated at $\theta^{(\cdot)}$. The likelihood is defined in equation A16.

- ψ

$$\psi \sim TN_{(0,+\infty)}(\mu_\psi, I_2\sigma_\psi^2)$$

A random walk proposal density is used in the (r)-th iteration, $\psi^{(r)} \sim TN_{(0,+\infty)}(\psi^{(r-1)}, I_2\sigma_{p\psi}^2)$, where $\psi^{(r-1)}$ is the value from the $(r-1)$-th iteration; $\sigma_{p\psi}^2$ is a scalar and functions as the tuning variance.

The acceptance probability $pr^* = \min(1, pr)$ and

$$pr = \frac{L(\psi^{(r)}|\cdot)p(\psi^{(r)}|\mu_\psi, I_2\sigma_\psi^2)\eta(\psi^{(r-1)}|\psi^{(r)}, \sigma_{p\psi}^2)}{L(\psi^{(r-1)}|\cdot)p(\psi^{(r-1)}|\mu_\psi, I_2\sigma_\psi^2)\eta(\psi^{(r)}|\psi^{(r-1)}, \sigma_{p\psi}^2)}$$

where $p(\psi^{(\cdot)}|\mu_\psi, I_2\sigma_\psi^2)$ is the density of $\psi^{(\cdot)}$ evaluated using the prior. $L(\psi^{(\cdot)}|\cdot)$ is the likelihood evaluated at $\psi^{(\cdot)}$. The likelihood is defined in equation A16. Since the proposal density is asymmetric, the ratio $\frac{\eta(\psi^{(r-1)}|\psi^{(r)}, \sigma_{p\psi}^2)}{\eta(\psi^{(r)}|\psi^{(r-1)}, \sigma_{p\psi}^2)}$ is used to adjust the acceptance probability.

C Policy Simulation Implementation

C.1 Policy Simulation I: Incorporating Disaggregate-level Data

The first step in our analysis is to consider how the model differs in the presence of complete information on the part of the advertiser and the second step is to assess how advertiser profits change in light of complete information.

The consumer model under full information is similar to the one developed under incomplete information considered in Section 3.1. The main difference is that, in addition to download counts

$\mathbf{d}^t$, advertisers further observe every consumer i's download decision on every product j (denoted by a binary variable y_{ijt}), all their sorting/filtering choices κ_{it} and the individual level demographics, $Demo_{it}$. These are often observed across multiple periods.

Accordingly, we begin by amending the model described from Equation 1 to 11 by

$$\begin{aligned} \pi(y_{ijt}|u_{ijt}^{g\kappa}, \kappa_{it}^{g}, g_{it}) &= I\{u_{ijt}^{g\kappa} \geq 0\}y_{ijt} + I\{u_{ijt}^{g\kappa} < 0\}(1 - y_{ijt}) \quad &\text{(A20)} \\ \pi(\kappa_{it}^{g}|z_{it}^{g}, g_{it}) &= \sum_{n \in N_\kappa} I\left\{\kappa_{it}^{g} = n\right\} I\{\max\{z_{it}^{g}\} = z_{it}^{g\kappa}\} \quad &\text{(A21)} \end{aligned}$$

where N_κ indexes the number of search options (in our case four). The first equation captures the download decision and the second the search strategy decision. The new specification of disaggregate consumer model generates an augmented likelihood function of consumer i. Noting that within each segment, $\Pr(download|search\ strategy, search) * \Pr(search\ strategy|search) * \Pr(search)$, we write

$$\begin{aligned} & L_i(y_{ijt}, \kappa_{it}|\Omega_c) \quad &\text{(A22)} \\ = & \sum\nolimits_g \{\prod\nolimits_t \prod\nolimits_j [\int_{u_{ijt}^{g\kappa}} \int_{\mathbf{z}_{it}^{g}} \pi(y_{ijt}|u_{ijt}^{g\kappa}, \kappa_{it}^{g}, g_{it})\pi(\kappa_{it}^{g}|\mathbf{z}_{it}^{g}, g_{it})\pi(u_{ijt}^{g\kappa}|\cdot)\pi(\mathbf{z}_{it}^{g}|\cdot)du_{ijt}^{g\kappa}d\mathbf{z}_{it}^{g}] \\ & \times \Pr(search_{it}^{g})\}pg_{it}^{g} \end{aligned}$$

where $\Omega_c \equiv \{\alpha^g, \beta_a^g, \delta^g, \gamma^g, \lambda_0^g, \lambda_1^g\}_g$ are parameters to be estimated. Note that although we include the product operator over period t's in the likelihood function, if we observe consumer i for only once in our data, the product over periods becomes moot.

Naturally, the full posterior distribution of the model will be the product of L_i across individuals as well as the prior distributions of parameters, i.e.,

$$p(\Omega_c|data) = \prod_i L_i(y_{ijt}, \kappa_{it}|\Omega_c) \cdot p(\Omega_c) \quad \text{(A23)}$$

With the established new posterior distribution, firms can impute the consumer model using disaggregate data and adjust their download expectations accordingly. In this manner a new prediction of $d(k, X_j^t; \Omega_c)$ can be constructed similar to equation 13 using these new estimates. This prediction is then used in the advertisers bidding game to calculate new equilibrium advertiser returns.

By comparing the predicted advertiser and platform profits under complete and incomplete information, we can impute the value of that information and help to determine how the search engine should price this information.

Further, advertisers are heterogeneous in their valuations about the keyword auction. Through the policy simulation, we will also be able to observe how the information revelation impacts differently on these heterogeneous advertisers. In particular, the questions we may be able to answer include:

- How does the information revelation change the auction market structure in terms whether we still have only a few bidders dominate the auctions?
- How does the information revelation change the total returns among advertisers? The answer to this question gives us insights about how information should be valued for the whole auction market.
- How does the information revelation impact the revenue of the search engine?

C.2 Policy Simulation II: Segmentation and Targeting

Neither the search engine and advertisers actually observes the segment memberships of consumers to help with targeting. However, it is possible for the advertiser to infer the posterior probability of consumer i's segment membership conditional on its choices. These estimates can then be used to improve the accuracy and effectiveness of targeting.

More specifically, suppose the search engine observes consumer i in several periods. Let us consider consumer i's binary choices over downloading, sorting/filtering and searching in those periods. Denote these observations as $H_i(\{y_{ijt}\}_{j,t}, \{\kappa_{it}\}_t, \{search_{it}\}_t)$. The likelihood of observing $H_i(\{y_{ijt}\}_{j,t}, \{\kappa_{it}\}_t, \{search_{it}\}_t)$ is

$$\begin{aligned} & L(H_i(\{y_{ijt}\}_{j,t}, \{\kappa_{it}\}_t, \{search_{it}\}_t)) \\ = \; & \sum_g \prod_t L(H_i(\{y_{ijt}\}_j, \{\kappa_{it}\}_t, \{search_{it}\}_t)|g_{it}) \cdot pg_{it}^g \end{aligned} \tag{A24}$$

where

$$\begin{aligned} & L(H_i(\{y_{ijt}\}_{j,t}, \{\kappa_{it}\}_t, \{search_{it}\}_t)|g_{it}) \\ = \; & \prod_j \int_{u_{ijt}^{g\kappa}} \int_{\mathbf{z}_{it}^g} \pi(y_{ijt}|u_{ijt}^{g\kappa}, \kappa_{it}^g, g_{it}) \pi(\kappa_{it}^g|\mathbf{z}_{it}^g, g_{it}) du_{ijt}^{g\kappa} d\mathbf{z}_{it}^g \Pr(search_{it}^g) \end{aligned} \tag{A25}$$

Hence the posterior probability of segment membership for consumer i can be updated in a Bayesian fashion,

$$\begin{aligned} \Pr(i \;\in\; & g|H_i(\{y_{ijt}\}_{j,t}, \{\kappa_{it}\}_t, \{search_{it}\})) \\ = \; & \frac{\prod_t L(H_i(\{y_{ijt}\}_j, \{\kappa_{it}\}_t, \{search_{it}\})|g_{it}) \cdot pg_{it}^g}{\sum_{g'} \prod_t L(H_i(\{y_{ijt}\}_j, \{\kappa_{it}\}_t, \{search_{it}\})|g'_{it}) \cdot pg_{it}^{g'}} \end{aligned} \tag{A26}$$

When consumer i returns to the search engine, the engine will have a more accurate evaluation about the segment membership of that consumer.

On the other hand, suppose some consumers only visit the engine once. Before they make the product choices, the search engine cannot obtain a posterior distribution outlined in Equation A26 since their choices of products are still unavailable. Still, it is possible to establish a more informative prediction about their memberships based on their κ_{it}'s before their product choices. Similar to Equation A26, the posterior in this case is

$$\Pr(i \in g|H_i(\kappa_{it})) = \frac{L(H_i(\kappa_{it})|g_{it}) \cdot pg_{it}^{g}}{\sum_{g'} L(H_i(\kappa_{it})|g'_{it}) \cdot pg_{it}^{g'}} \tag{A27}$$

where

$$L(H_i(\kappa_{it})|g_{it}) = \int_{z_{it}^{g\kappa}} \pi(\kappa_{it}^{g}|z_{it}^{g}, g_{it}) dz_{it}^{g\kappa}$$

We can construct an analysis to consider the benefits of targeting as follows. First, we compute the return to advertisers when advertisers can only bid on key words for all segments. Second, we compute the return accruing to advertisers when they can bid for key words at the segment level. The difference between the two returns can be considered as a measure for the benefits of targeting.

References

Alba, J., J. Lynch, B. Weitz, C. Janiszewski, R. Lutz, A. Sawyer, and S. Wood (1997, July). Interactive home shopping: Consumer, retailer, and manufacturer incentives to participate in electronic marketplaces. *Journal of Marketing 61*(3), 38–53.

Ansari, A. and C. F. Mela (2003). E-customization. *Journal of Marketing Research 40*(2), 131–145.

Ansari, A., C. F. Mela, and S. A. Neslin (2008, February). Customer channel migration. *Journal of Marketing Research 45*(1), 60–76.

Bajari, P., C. L. Benkard, and J. Levin (2007). Estimating dynamic models of imperfect competition. *Econometrica 75*(5), 1331–1370.

Bajari, P. and H. Hong (2006). Semiparametric estimation of a dynamic game of incomplete information. *NBER Technical Working Paper 320*.

Ben-Akiva, M. and S. Lerman (1985). *Discrete Choice Analysis: Theory and Application to Travel Demand*. MIT Press.

Bradlow, E. T. and D. C. Schmittlein (2000). The little engines that could: Modeling the performance of world wide web search engines. *Marketing Science 19*(1), 43–.

Chen, Y. and C. He (2006). Paid placement: Advertising and search on the internet. *Working Paper*.

Diehl, K., L. J. Kornish, and J. G. Lynch (2003, June). Smart agents: When lower search costs for quality information increase price sensitivity. *Journal of Consumer Research 30*(1), 56–71.

Dubé, J.-P., G. J. Hitsch, and P. Chintagunta (2008). Tipping and concentration in markets with indirect network effects. *Working Paper*.

Edelman, B., M. Ostrovsky, and M. Schwarz (2007). Internet advertising and the generalized second price auction: Selling billions of dollars worth of keywords. *The American Economic Review 97*, 242–259.

Feng, J. (2008). Optimal mechanism for selling a set of commonly-ranked objects. *Marketing Science*, forthcoming.

Garg, D., Y. Narahari, and S. S. Reddy (2006). An optimal mechanism for sponsored search auctions and comparison with other mechanisms. *Working Paper*.

Gelman, A., J. B. Carlin, H. S. Stern, and D. B. Rubin (2004). *Bayesian Data Analysis* (2 ed.). Chapman & Hall/CRC.

Ghose, A. and S. Yang (2007). An empirical analysis of search engine advertising: Sponsored search and cross-selling in electronic markets. *Working Paper*.

Goldfarb, A. and C. Tucker (2007). Search engine advertising: Pricing ads to context. *Working Paper*.

Greene, W. H. (2003). *Econometric analysis*. Prentice Hall.

Groves, T. (1979, apr). Efficient collective choice when compensation is possible. *The Review of Economic Studies 46*(2), 227–241.

Hotz, V. J. and R. A. Miller (1993). Conditional choice probabilities and the estimation of dynamic models. *The Review of Economic Studies 60*(3), 497–529.

Häubl, G. and K. B. Murray (2003). Preference construction and persistence in digital marketplaces: The role of electronic recommendation agents. *Journal of Consumer Psychology 13*(1/2), 75–91.

Iyengar, G. and A. Kumar (2006). Characterizing optimal adword auctions. *Working Paper*.

Jofre-Bonet, M. and M. Pesendorfer (2003). Estimation of a dynamic auction game. *Econometrica 71*(5), 1443–1489.

Kamakura, W. A. and G. J. Russell (1989, November). A probabilistic choice model for market segmentation and elasticity structure. *Journal of Marketing Research 26*(4), 379–390.

Katona, Z. (2007). The race for sponsored links: A model of competition for search advertising. *Working Paper*.

Krishna, V. (2002). *Auction Theory*. Academic Press.

Lynch, J. G. and D. Ariely (2000). Wine online: Search costs affect competition on price, quality, and distribution. *Marketing Science 19*(1), 83–103.

McFadden, D. L. (1977). Modeling the choice of residential location. *Cowles Foundation Discussion Paper No. 477*.

Natarajan, R. and C. E. McCulloch (1998). Gibbs sampling with diffuse proper priors: A valid approach to data-driven inference? *Journal of Computational and Graphical Statistics 7*(3), 267–277.

Palma, A. D., S. P. Anderson, and J. F. Thisse (1992). *Discrete Choice Theory of Product Differentiation*. Massachusetts Institute of Technology Press.

Pesendorfer, M. and P. Schmidt-Dengler (2008). Asymptotic least squares estimators for dynamic games. *Review of Economic Studies 75*, 901–928.

Powell, W. B. (2007). *Approximate Dynamic Programming: Solving the Curses of Dimensionality.* Wiley, Inc. Publication.

Qiu, F., Z. Liu, and J. Cho (2005). Analysis of user web traffic with a focus on search activities. In *Proceedings of the International Workshop on the Web and Databases (WebDB)*, pp. 103–108.

Rochet, J.-C. and J. Tirole (2006). Two-sided markets: A progress report. *Rand Journal of Economics 37*(3), 645–667.

Rossi, P. E., G. M. Allenby, and R. E. McCulloch (2005). *Bayesian statistics and marketing.* John Wiley and Sons.

Russo, J. E. (1977, May). The value of unit price information. *Journal of Marketing Research 14*(2), 193–201.

Rust, J. (1994). Structural estimation of markov decision processes. In R. F. Engle and D. L. McFadden (Eds.), *Handbook of Econometrics*, Volume IV. Amsterdam: Elsevier Science.

Rutz, O. J. (2007). A model of individual keyword performance in paid search advertising. *Working Paper*.

Rutz, O. J. and R. E. Bucklin (2007). From generic to branded: A model of spillover dynamics in paid search advertising. *Working Paper*.

Ryan, S. (2006). The costs of environmental regulation in a concentrated industry. *Working Paper, Massachusetts Institute of Technology*.

Ryan, S. and C. Tucker (2007). Heterogeneity and the dynamics of technology adoption. *Working Paper, Massachusetts Institute of Technology*.

Spiegelhalter, D. J., N. G. Best, W. R. Gilks, and H. Inskip (1996). Hepatitis b: a case study in mcmc methods. In W. R. Gilks, S. Richardson, and D. J. Spiegelhalter (Eds.), *Markov Chain Monte Carlo in Practice*, Chapter 2, pp. 21–43. Chapman & Hall/CRC.

Stryszowska, M. (2005). Simultaneous vs. overlapping internet auctions. *Working Paper*.

Train, K. (2003). *Discrete Choice Methods with Simulation.* Cambridge University Press.

Tucker, C. (2005). Empirically evaluating two-sided integrated network effects: The case of electronic payments. *Working Paper, Massachusetts Institute of Technology*.

Varian, H. R. (2006). Position auction. *International Journal of Industrial Organization*, forthcoming.

Vickrey, W. (1961). Counterspeculation, auctions, and competitive sealed tenders. *Journal of Finance 16*(1), 8–37.

Yao, S. and C. F. Mela (2007). Online auction demand. *Marketing Science*, forthcoming.