Sponsored Search: Search Characteristics, Advertiser Quality & Click Performance

Ashish Agarwal
Assistant Professor
McCombs School of Business
University of Texas, Austin
Austin, TX 15213
Phone: (512) 471 5814
Email: ashish.agarwal@mccombs.utexas.edu

Tridas Mukhopadhyay*
Deloitte Consulting Professor of e-Business
Tepper School of Business
Carnegie Mellon University
4800 Forbes Avenue
Pittsburgh, PA 15213
Phone: (412) 268 5978
Email: tridas@cmu.edu
* Please Do Not Circulate Without Permission*

ABSTRACT

Sponsored search ads appear as ordered lists in search engine results pages. While it is established that the click performance decays with the position of the ad, little is known about the impact of search characteristics and advertiser characteristics on this performance.

We use a dataset for the ad campaigns of two competing advertisers for several hundred keywords provided by a search engine to evaluate the click performance. We use a hierarchical bayesian model to account for the click decisions made by consumers. We find that the click performance decreases with position at a higher rate for high quality advertiser. We also find that while the high quality advertiser gets a higher click performance for more specific keywords, the rate of decrease in click performance with position further increases for the high quality advertiser. This suggests that the quality perception of the advertiser maybe influenced by the position of their advertisement.

Our results inform the advertising strategies of firms participating in sponsored search auctions and provide insight into consumer behavior in these environments. Our results also show that the search engine may be overcharging some advertisers and promoting low quality advertisers.

Introduction

Internet advertising spend is growing faster than any other form of advertising and is expected to surge from $19.4 billion in 2006 to $35.5 billion in 2011 (eMarketer). 40% of this ad spend is on sponsored search where advertisers pay to appear alongside the algorithmic search results of a search engine. Most search engines including Google, Yahoo, and MSN use auctions to sell their inventory of ad space. In these auctions, advertisers submit bids on keywords based on their willingness to pay for every click from a consumer searching on that (or a closely related) keyword. Search engines use a combination of the submitted bids and advertiser quality for every keyword to rank the ads. Sponsored search is widely regarded as one of the most effective forms of advertising because it occurs close to a user's purchase decision and is matched based on the user's stated information need. The search engines use similar auction format in other forms of online advertising such as contextual advertising.

Even though the phenomenon is growing rapidly, there is very little understanding of the drivers of performance in this marketplace. While it is known that the click performance decays with the position of the ads it is not clear how it varies across advertisers. This is vital information as popular search ranking mechanisms use the click propensity of an advertiser as a measure of advertiser quality.[1] A misrepresentation of the advertiser quality based on click propensity would result in a loss of revenue for the search engine or higher costs for advertisers. A common assumption in the literature is that the impact of the type of advertiser is separate from the position effect on the click performance (Lahaie and Pennock, 2007; Liu, Chen, and Whinston, 2010; Gerath et al. 2011; Athey and Ellison, 2011). In other words, the assumption is

1 http://adwords.google.com, https://**adcenter**.microsoft.com/,
http://help.yahoo.com/l/sg/yahoo/ysm/sps/start/overview_newadrank.html

that the decay in click performance of ads is independent of the quality of advertiser. However, there is no empirical validation of this assumption. While the actual quality measure used by search engines is proprietary information, the quality information revealed to advertisers by search engines is independent of the actual rank placement.[2] Advertisers use this quality information to formulate their bidding strategies. Many empirical studies in sponsored search (Ghose and Yang 2009, Agarwal, Hosanagar and Smith, 2011; Animesh, Viswanathan and Agarwal, 2011) are also single advertiser studies and do not consider the effect of advertiser quality while evaluating the click performance. Jeziorski and Segal (2009) model the consumer click behavior and show that relative ordering of ads is important to determine the click performance of ads. They attribute this to the differences in the quality of ads to satisfy information needs of consumers. However, they do not systematically determine how the click performance varies across advertisers of different quality.

It is also not clear how consumers with different search characteristics respond to different advertisers. Specific keywords are less popular than common keywords as illustrated in figure 1 which shows the search volumes for different keywords obtained from Google trends. Additionally, there is lower competition for such keywords which would make these a very attractive target for advertisers. Consumers using specific keywords are expected to be more informed and expected to search more (Moorthy et al, 1997; White and Morris, 2007; White et al. 2009). Previous work on consumer behavior in sponsored search (Ghose and Yang 2009, Agarwal et al. 2011) has shown contradictory evidence on the impact of specific keywords on

---

2 http://adwords.google.com/support/aw/bin/answer.py?hl=en&answer=6111,
https://adwords.google.com/support/aw/bin/answer.py?answer=100305,
http://adcenterhelp.microsoft.com/help.aspx?project=adcenter_live_std&market=en-us&querytype=topic&query=welcome.htm,
http://help.yahoo.com/l/us/yahoo/ysm/sps/faqs_all/faqs.html#position_qi

the performance. Moreover, the effect of the keyword characteristics on the click performance across advertisers of different quality is not known.

= = Insert Figure 1 about here = =

In this paper we seek to understand how the click performance varies across advertisers in sponsored search and how keyword characteristics influence this performance across advertisers. To do this, we empirically analyze a unique panel dataset from a major search engine, which catalogs bids, position, and click performance data for several hundred keywords related to two competing advertisers. We label these advertisers as high quality and low quality based on their click propensity for individual keywords using criteria suggested by popular search engines as well as existing literature. For our analysis we consider a time period of two weeks during which none of the advertisers (including the ones not considered in our analysis) change bids for any of these keywords.[3] This allows us to control for impact of advertiser decision on the ad position. We use a hierarchical bayesian model to capture the click decisions made by consumers. We also account for the endogeneity of the ad position due to the decisions made by the search engine as well as the missing variables commonly observed by consumers but not available to us. We find that the click performance depends on both the advertiser quality and the ad position of the advertiser. While the higher quality advertiser gets a higher click performance as expected, it also experiences a higher decay in the click performance with the position. This suggests that consumers evaluating lower positions are less likely to consider the quality difference between advertisers. We also find that the decay in click performance of high quality advertiser is higher for more specific keywords even though it gets a higher overall click performance for these

3 It is very common for advertisers to update their bids on a weekly or even a monthly basis. We have verified this with our search engine as well as search engine marketing firm managing sponsored search ads for several firms.

keywords. This suggests that more informed consumers who use more specific keywords are even less likely to consider the quality difference between advertisers at lower positions. We verify our results across other competing advertiser pairs. We quantify the effect of evaluating click propensity of advertisers as a function of ad position instead of a simple average click propensity. We find that our high quality advertiser is over charged for top positions and under charged for lower positions if a simple click propensity measure is used. The cost of clicks can deviate by 50% by using a position based measure of quality.

We make several contributions in this paper. First, we provide insight into how different consumers respond to different advertisers in the sponsored search environment. Existing search cost literature attributes the decay in click performance with position to sequential search as well as heterogeneity in search costs (Arbatskaya 2007). However, it is not known how this would influence the consumer decision to evaluate firms with different quality perception. We show that in the sponsored search context, where consumers are expected to search sequentially, this can result in differential response to advertiser quality at different positions. Consumers, who search more, tend to place lower emphasis on perceived advertiser quality at lower position. This can be due to the fact the consumers are satiated with information as suggested by Jeziorski and Segal (2009) and the expected utility gain from any advertiser is lower at lower position irrespective of their perceived quality. Consumers using specific keywords are expected to search more (Moorthy et al, 1997; White and Morris, 2007; White et al. 2009) and are less likely to be satiated at any given position. Additionally, they are more likely to click the high quality advertiser as they are more informed. In that case the perceived quality difference should reduce at a lower rate with position for these keywords as compared to the less specific keywords. However, we find that consumers using more specific keywords are even more likely to discount perceived quality at lower positions. This suggests that the quality perception itself maybe a function of position.

Our results inform advertisers on their current advertising strategies. Our results suggest that advertisers should consider both position and their relative quality at different positions while determining their bidding strategy. Gerath et al. (2011) suggest that a high quality advertiser can get higher click performance at lower position as compared to a low quality advertiser at a higher position. Our results show that while this may be true, the high quality advertiser may not get the expected click performance at lower positions due to lower quality perception. However, the bidding efficiency is better for these advertisers at lower positions due to potential undercharging by the search engine using an average measure of advertiser quality. Previous work (Ghose and Yang 2009, Agarwal et al. 2011) shows that bidding efficiency can be better at lower positions because of lower costs. Our study shows that apart from the better bidding efficiency, advertisers with lower quality may benefit from being at lower positions as the decrease in click rate with position is lower for these advertisers.

Finally, our results inform the search engines on the current gap in the auction rules. Our results show that search engines maybe overcharging high quality advertisers and undercharging low quality advertisers for top positions. This can reduce advertiser participation. Additionally, if the quality represented by the overall click propensity represents the true quality, then this can lead to adverse selection as more low quality advertisers are likely to be at the top. This can lead to lower search engine revenues in the long run due to lower customer participation. Athey and Ellison (2011) suggest that the ranking mechanism can be inefficient if it uses only the click propensity to evaluate advertisers instead of their true quality as the click propensity maybe driven by familiarity with the advertiser. As a result, consumers waste clicks before discovering the true quality of advertisers. Our results show less difference in click propensity across advertisers at lower positions which indicate that consumers evaluating ads at lower positions are less affected by the click propensity as a measure of quality of advertisers. Thus, if there is indeed a mismatch between the true quality and the click propensity, the auction mechanism may not be that inefficient at lower positions from a consumer welfare point of view.

Previous Work

Our work is related to the consumer behavior in the online context with a specific focus on sequential search, search characteristics and quality perception. We also explain the related work in sponsored search.

Sequential Search

Prior work in traditional media has demonstrated that message ordering influences ad persuasion (Rhodes et al. 1979, Brunel and Nelson, 2003). Similar results have been shown in online environments. In fact, Hoque and Lohse (1999) find that consumers are more likely to choose advertisements near the beginning of an online directory than they are when using paper directories. Ansari and Mela (2003) have found that the higher position of links in email campaign can lead to higher probability of clicking. Prior studies that have shown that the depth of consumer search on the Internet is low suggest that message ordering plays a key role on search engines as well. For example, Johnson et al. (2004) found that consumers searched fewer than two stores during a typical search session. Similarly, Brynjolfsson, Dick, and Smith (2009) find that only 9% of users of a shopbot select offers beyond the first page. Due to cognitive costs associated with evaluating alternatives, consumers often focus on a small set of results (Montgomery et al. 2004). In case of search results, there is evidence from eye tracking studies

(Granka et al. 2004) to suggest that consumers do indeed focus on the top of the list in case of organic search results which are displayed along with the sponsored search results. Feng et al. (2007) find evidence that the number of clicks for an ad decreases exponentially with its position, and attribute this to decay in user attention as one proceeds down a list. However, they do not distinguish keywords as well as advertisers and the relative influence of their characteristics on the search performance. Similarly, single advertiser studies (Ghose and Yang, 2009; Rutz and Trusov, 2011; Agarwal, Hosanagar, and Smith2011) have established that click performance decays with position of the ad. An important consideration is how the perception of advertiser quality influences the click performance.

Advertiser Quality

In a study of the effect of competition on advertising memory recall, Kent and Allen (1994) show that consumers are more likely to recall familiar brands. Dodds et al. (1991) show that consumer perceive familiar brands with higher quality. In the sponsored search context, this would suggest that consumers would perceive familiar sites as of higher quality and this would influence how they select the sponsored search ad. Animesh, Viswanathan and Agarwal (2011) show that in sponsored search, ad's unique selling proposition such as quality or price information drives its performance along with ad's position for the same advertiser. Similarly, Park and Park (2010) find that the consumer click decision depends on the advertiser's selling proposition as well as unknown advertiser characteristics. This confirms that consumers do pay attention to factors other than ad position and may demonstrate propensity to click ads based on their perception of quality.

As consumers are expected to evaluate ads sequentially, starting from the top it is important to know how the perception of quality weighs against the ad position. Erdem et al. (2008) show that price is a signal of quality. Similarly in the sponsored search context, one can expect that the position of the ad signals its quality. Previous work (Chen, Feng, and Whinston 2009; Animesh et al. 2010; Gerath et al. 2011) has suggested that low quality advertisers prefer top position as it leads to higher clicks. Gerath et al. (2011) also suggest that a high quality advertiser can get better performance at lower position as compared to a low quality advertiser who appears at a higher position. Thus, it is possible that advertisers may not be ordered in the decreasing order of quality due to higher bids submitted by lower quality advertisers. As the consumers engage in sequential search, they may associate higher position ads with higher quality even if the quality is lower. Animesh, Viswanathan and Agarwal (2011) show that competitive intensity has a positive effect at lower positions. This suggests that consumers clicking at lower positions behave differently. However, it is not clear how the performance will differ across advertisers at lower position. Jeziorski and Segal (2009) show that advertisers appearing higher up in the search results impose a negative externality on the lower ads and attribute it to information satiation. However, Athey and Ellison (2010) suggest that advertisers appearing higher up can exert positive externality as consumers learn about the relevance of ads.

Keyword Characteristics

Search behavior is also dictated by the consumer’s purchase intent. Online consumers include both buying consumers and information seekers (Moe 2003, Moe and Fader, 2004; Montgomery, Li, Srinivasan, and Lietchy, 2004). Consumers with high purchase intent tend to be very focused in their search, targeting a few products and categories versus consumers with low purchase intent, who have broad search patterns targeting a higher variety of products (Moe 2003). A similar pattern can be expected in sponsored search i.e. consumers may be heterogeneous in terms of their search behavior. Consumers use different types of queries to search for information which may indicate their purchase intent. Using path analysis, Montgomery, Li, Srinivasan, and Lietchy (2004) show that consumers with directed search have higher probability of purchase. Specific keywords used by consumers are an indication of a more directed search and can potentially reflect their underlying purchase intent. Thus, the search volume for specific keywords may reflect a higher proportion of buyers as compared to less specific keywords. Urbany et al. (1989) show that consumers with higher uncertainty about the information for the alternatives are likely to search less while consumers with uncertainty about choice are likely to search more. Brucks (1985) shows that product class knowledge increases search. Srinivasan and Ratchford (1991) also show that prior knowledge leads to an increase in search. Moorthy et al. (1997) show that consumers with low expertise are likely to search less. This would suggest that consumers using common keywords may search less because of their uncertainty about the information. On the other hand consumers using specific keywords are less satiated and may search more. White and Morris (2007) and White et al. (2009) find that advanced users submit longer queries and click more pages in regular search. In the sponsored search context, Ghose and Yang (2009) show that the click performance is actually lower for longer keywords. However, Rutz and Trusov (2011) and Agarwal, Hosanagar, and Smith (2011) do not find any association between the click through rate and the specificity of a keyword. These studies suggest that there can be heterogeneity in among advertisers in terms of the effect of keyword specificity on the click performance.

Another important consideration is the interplay of search characteristics with the quality perception. A consumer in the information seeking mode and using less specific queries has a higher degree of quality uncertainty. Consumers using specific keywords are advanced in their information search. One possibility is that these consumers might have started of their information search using less specific keywords and have become informed through multiple keyword searches. In that case, we can expect a spillover effect (Rutz and Bucklin, 2011) which will make them more aware of relevant advertisers for such keywords. Thus, if they have to choose between a high quality advertiser and a low quality advertiser, they should be more likely to select the high quality advertiser. Another possibility is that specific keywords reflect consumer's personal preference. For example, a consumer searching for 'dress shirts' may be just revealing personal preference for the kind of 'shirts' they are interested in. However, personal preference is driven by the past consumption of the product (Dube, Hitsch and Rossi 2010). Past consumption would suggest that such consumers should be more familiar with firms selling products as compared to those who start with a non specific search. Consumers using specific keywords can be expected to search more. However, after accounting for their search intensity, these consumers can be expected to click high quality advertisers at a higher rate. Using an experimental setup, Dou et al. (2010) find evidence which suggest that an unknown brand is more likely to be selected when it appears before the known brand in sponsored search. They also find that this would be the case when the users are less experienced. This suggests that the perception of quality depends on the experience of the user. However, it is not known how this perception of quality will be influenced by the position of the advertiser. One possibility is that at lower positions consumers are satiated as suggested by Jeziorski and Segal (2009) and the quality difference between advertisers is not important. However, consumers using specific keywords are expected to be less satiated and search more. Additionally, these consumers are expected to have lower quality uncertainty as they are more informed. In that case, for consumers using specific keywords, quality difference should be higher at lower positions as compared to that for consumers using less specific keywords.

The prior literature reveals that there can be heterogeneity in the user response to the ads in sponsored search, depending on their search stage and advertiser quality. Thus, the net impact of user and advertiser characteristics and their interaction on the click performance in sponsored search is an open and managerially significant research question.

Data Description

Our dataset is provided by Yahoo[4] and consists of bids, impressions, and clicks for different positions over 123 days from January 2008 to April 2008 for a large number of keywords and the corresponding advertisers. During the time period of our data panel, Yahoo's sponsored search design was similar to that of Google in terms of the auction mechanism and search results.[5]

We consider two advertisers who bid on several hundred keywords. We consider only those keywords which are common to both advertisers. Together these advertisers account for 42% clicks for these keywords. We also consider only those keywords where all advertiser bids do not change for a consecutive two week period. This takes away the endogeneity of the ad positions due to strategic bidding by advertisers. Additionally, advertisers are unlikely to make ad and website changes as they keep their bids constant. This is a valid assumption as the time period of our panel is very short for advertisers to make these strategic changes. As we are interested in measuring the click performance of these advertisers for different positions, we consider only those positions where there is some click data available. This allows us to get estimates for the coefficient of the ad position without the influence of very low positions where none of the advertisers get any clicks. Our final sample consists of 348 sample keywords which generate 524029 impressions and 6562 clicks for the two advertisers during our panel data

[4] Yahoo held about 20% search market share in 2008 http://searchengineland.com/search-market-share-2008-google-grew-yahoo-microsoft-dropped-stabilized-16310

[5] http://www.reuters.com/article/2007/06/05/us-yahoo-advertising-idUSN0420377020070605

period. Note that we do not have detailed click stream information which will reveal exact arrangement of ads seen by customers. However, we verify that the selected keywords have performance data for both advertisers on a daily basis. Summary statistics for our sample are shown in Table 1. Next we describe how we determine the advertiser quality as well as keyword specificity.

Advertiser Quality: We label our advertisers as high quality and low quality based on their historical click rate for each of the keywords. This is the standard measure used by search engines and advertisers to evaluate the quality of advertisers.[6] The ability to generate higher click performance reflects the potential of an advertiser to drive traffic to its website. Web traffic has been used as a measure of perceived quality in previous studies (Brynjolfsson and Smith 2000, Palmer 2002, Animesh et al 2010). Website traffic has been found to have strongly positive relationship with the market value of firms (Trueman et al. 2000, Demers and Lev, 2001).

The advertiser quality ordering is the same for all keywords in our sample. For any keyword, the ad performance can also be influenced purely by its position due to sequential search. Additionally, the difference in performance of advertisers can be different at different positions. We have to adjust for these factors in order to derive the true advertiser click propensity from the past click performance.[7] We mean center the click through rate for each keyword -advertiser pair for each period and each position using the mean performance of all advertisers for that keyword at that position in that period and the standard deviation of the advertisers' performance for that keyword at that position in that period. We then use the average value of this normalized

---

6 http://adwords.google.com, https://**adcenter**.microsoft.com/, *http://help.yahoo.com/l/sg/yahoo/ysm/sps/start/overview_newadrank.html*

[7]http://adwords.google.com/support/aw/bin/answer.py?hl=en&answer=115967, http://help.yahoo.com/l/us/yahoo/ysm/sps/faqs_all/faqs.html#position_qi

advertiser CTR performance for each keyword across different positions in a period to come up with a measure of the advertiser click propensity for a keyword in that period. We then take an average of the click propensity of an advertiser over past 10 weeks to determine the click propensity of the advertiser for that keyword. Appendix 1 shows how this click propensity is calculated. Search engines are also known to evaluate the relative quality of the advertiser using past click through performance which is adjusted for the position. Also, the relative quality is evaluated for each ad which is specific to a keyword.[8] Note that this click propensity is just used to label the advertiser as high quality or low quality and is not used to estimate the parameters. We have verified the robustness of our labeling of advertisers using other measures such as the click propensity for most recent weeks in the past, overall click propensity across multiple keywords.

Other factors such as promotions can influence the past click performance of an advertiser for a keyword which can artificially improve the click performance for a short period. As we evaluate the click propensity of an advertiser for each keyword based on different lag periods the effect of these short term changes should be minimized. It is possible that all past performance for an advertiser was a result of promotions and advertising. However, such long term spending is more likely from a high quality advertiser ashigh quality advertiser is expected to spend higher amount for advertising (Kirmani and Wright 1989; Moorthy and Zhao, 2000). The high quality advertiser in our sample also has a larger portfolio of keywords and gets more clicks than the low quality advertiser. This again shows that our labeling confirms to the expected

---

8https://adwords.google.com/support/aw/bin/answer.py?answer=100305,
https://adwords.google.com/support/aw/bin/answer.py?hl=en&answer=115967,
http://help.yahoo.com/l/sg/yahoo/ysm/sps/start/overview_qualityindex.html

advertising outcomes for advertisers of different quality. It is important to note that the measure of advertiser quality in our study is relevant to the context.

Click performance can be influenced by the regular (organic) search results. Yang and Ghose (2010) show that organic and sponsored search are complements. As organic results are purely driven by the advertiser quality, higher click performance on sponsored search ads due to organic results also indicates higher quality. [9] Thus, we don't have to adjust the click performance for the influence of organic results in order to label our advertisers as high quality and low quality.

Keyword Specificity: Actual keywords are not available to us.[10] Instead, each word is represented by a unique id. For example, if the words 'red' and 'dress' have been assigned ids 1 and 2 respectively, then the keywords 'red', 'dress' and 'red dress' would be represented as '1', '2' and '12' in our dataset. In order to determine keyword specificity we determine parent child linkages between the keywords using matching words. For example, dress → red dress, blue dress→red cocktail dress, blue cocktail dress. We can determine this easily in our setup because of the number ids assigned to each word in a keyword. Using this approach we generate a hierarchy and determine the number of keywords above each keyword. We call this number as the specificity of the keyword. In the above example, keyword 'dress' will have specificity 1, keywords 'red dress' and 'blue dress' will have specificity 2, and keywords 'red cocktail dress' and 'blue cocktail dress' will have specificity 3. Similar measure of specificity has been used by Agarwal, Hosanagar, and Smith (2011). Note that an alternate measure of keyword specificity is the number of keywords (Ghose and Yang, 2009; Yang and Ghose 2010). For most of our keywords, the length of keyphrase is equivalent to the specificity level.

---

9 http://www.google.com/support/webmasters/bin/answer.py?answer=35291

10 Search engine has masked the data to prevent revealing competitive information.

Model

We are interested in modeling the click performance of keywords for different advertisers. We do not have detailed clickstream data to represent the exact clicking behavior of consumers. Instead, we have the daily aggregate data for each advertiser keyword combination. Such aggregate data is very common in the sponsored search context. Current single advertiser studies in sponsored search (Ghose and Yang 2009; Yang and Ghose 2010; Rutz and Bucklin, 2011; Rutz and Trusov, 2011; Agarwal, Hosanagar, and Smith, 2011) use this kind of aggregate data to model the click and conversion performance. We extend this framework to account for the multiple advertisers in our setup.

Our unit of analysis is a keyword because the search engine auction is keyword specific. Keyword characteristics are an indication of the underlying search behavior, which varies across consumers. For example, the keyword 'shirt' is less specific and indicates an initial stage of information search, while more specific keywords like 'white shirt', 'formal blue shirt' indicate a more advanced and directed stage of information search. To account for these differences across keywords, we capture how specific a keyword is using the measure of keyword specificity described earlier. The specificity of a keyword is based on the nearness of its landing page to a product. Advertisers organize their websites in a hierarchical fashion to accommodate the search intent of users and to reduce their search cost. Various levels in the hierarchy represent product categories, sub-categories and products. When consumers are routed through a search engine, the landing page coincides with a level in the website hierarchy that is chosen based on the search intent of the consumer as reflected in the keyword. Besides specificity, there can be additional variables that capture keyword characteristics. For example, presence of 'retailer' or 'brand' information captures preference for the retailer or brand. However, in our datasets none of the keywords carry brand or retailer information.[11] Another variable that has been used in the past is the 'size' of the keyword which indicates the number of words in the keyword. This can capture additional aspects of consumer preference. However, this variable is redundant as the information is already captured in 'specificity' variable. There are other unobservable characteristics associated with a keyword that can influence consumer choice. For example, the regular search results are different for different keywords.

We use a hierarchical model to capture the effect of keyword characteristics. This provides a flexible random component specification, allowing us to incorporate observable and unobservable keyword-specific heterogeneity given the limited observations for each keyword. Hierarchical models are commonly used to draw inferences on individual level characteristics (Rossi and Allenby, 2003). Hierarchical Bayesian (HB) models have also been used to study sponsored search data with keyword as a unit of analysis (Ghose and Yang 2009, Yang and Ghose 2010, Rutz and Trusov, 2011; Agarwal, Hosanagar, and Smith 2011).

Consumer search behavior suggests that the clicking choice of an advertisement in the sponsored search results depends on the expected utility gain of clicking on one additional advertisement and the associated search cost (Jeziorski and Segal, 2009; Park and Park, 2010). As a result consumers could be selecting more than one ad if the expected gain is more than the search cost. In our setup we do not have the exact arrangement of ads for each consumer visit to represent the sequential click behavior. What we have is the aggregate position information for each advertiser on a daily basis. It is also possible that during the day some consumers may not see all competing advertisers in the search results as some advertisers may have dropped out due

11 As the actual keywords are not available to us, we have separately verified this with the search engine

to budget constraints. This is reflected in the position variations on a daily basis for each advertiser even though the bids do not change. Also, competing advertisers do not get the same number of impressions for each keyword. This prevents us from using a conditional approach (Park and Park, 2010) where we make the click choice of the advertiser with a lower average daily position dependent on the click choice of the advertiser with higher average daily position.

To represent this scenario, we assume that the consumer decision of clicking each advertiser is a binary choice where the sequential effect is purely captured by the position of the advertiser's ad. We assume an i.i.d. extreme value distribution of the error term for individual choices and use a logit model to represent the click probability for a keyword k at time t. This probability for advertisers of high quality H and low quality L can be represented as

$$(1) \qquad \Lambda^{a}_{k,t} = \frac{\exp(U^{a}_{kt})}{1+\exp(U^{a}_{kt})} \text{ where a = \{H,L\}}$$

where U^{a}_{kt} is the latent utility of clicking.

Apart from the position, consumer choice is also influenced by the competitive intensity which can vary across keywords (Animesh, Viswanathan and Agarwal, 2011). For example, more specific keywords have less competition. We represent this using the number of other advertisers bidding on the keyword and showing up in the search results. In order to represent other characteristics not observed by us, we use a random component specification. For a keyword k at time t, the latent utility of clicking can be expressed as

$$(2) \qquad U^{a}_{kt} = \beta^{ka}_{0} + \beta^{ka}_{1}\mathrm{Pos}_{kat} + \beta^{ka}_{2}\mathrm{CI}_{kat} + \beta^{a}_{\mathrm{Time}}\mathrm{Time}_{kat} + \varepsilon^{\beta_a}_{kt}$$

$$\beta^{ka} = \Delta^{\beta_a} z_k + u^{\beta_a}_{k} \qquad u^{\beta_a}_{k} \sim N(0, V^{\beta_a}) \text{ where } \beta^{ka} = [\beta^{ka}_{0}, \beta^{ka}_{1}, \beta^{ka}_{2}]$$

where Pos represents the position of the ad in sponsored search results, and CI is the competitive intensity, z_k represents keyword specificity, Δ^{β_a} is a matrix capturing the relationship between the keyword characteristics and the mean values of coefficients, $u_k^{\beta_a}$ represents the unobservable heterogeneity for the random coefficients, which we assume is normally distributed with a mean 0 and covariance matrix V^{β_a}, and $\varepsilon_{kt}^{\beta_a}$ represents the time varying unobserved keyword attributes which are common for all consumers. We also control for the time dynamics of the auction using a time variable Time_{kat}. As consumer choice of clicking for the competing advertisers is correlated, we have

(3) $$\begin{bmatrix} \varepsilon_{kt}^{\beta_H} \\ \varepsilon_{kt}^{\beta_L} \end{bmatrix} \sim N\left[\begin{pmatrix} 0 \\ 0 \end{pmatrix}, \begin{pmatrix} \Omega_{HH}^{cl} & \Omega_{HL}^{cl} \\ \Omega_{HL}^{cl} & \Omega_{LL}^{cl} \end{pmatrix}\right]$$

Endogeneous Ad Position

The ad position is determined by the search engine based on bids submitted by advertisers and their expected click propensity. In our setup, bids for none of the sample keywords change during the panel period. So the influence of advertisers bid on the ad position is addressed. Additionally, advertisers are not expected to make any ad related changes as well as website changes in the short panel period considered. In that case, we do not expect any changes in the click propensity of advertisers due to ad specific changes. However, the search engine can still vary the position of advertisers for a keyword. Advertisers place different limits on the spending budget for clicks (Hosanagar and Cherapanov, 2008). When the budget limit is reached advertiser is not considered for the auction. For example, an advertiser may bid very high but have limited budget. In that case the advertiser will appear in the top positions for a small number of searches and then won't be considered for additional auctions which will improve the position of the other advertisers with lower bid and higher budgets. Thus an advertiser can be in different positions for a keyword during the day. Additionally, the search engine can manipulate the results to maximize the clicks. For example, search engines can let an advertiser exceed their daily budget on heavy traffic days and compensate for it on other days.[12] This leads to changes in the daily average position of advertisers. This also leads to changes in the composition of the sponsored search results in terms of the cumulative ad content across all displayed advertisers. This is observed by consumers but not known to us and the advertiser position is correlated with these changes in the composition. We also do not observe regular search results. Yang and Ghose (2010) show complementarity between the click performance of regular search results and organic search results. Any changes in the regular search results would influence the clicks which in turn would influence the budget consumption of ads. Thus, the position of ad can be correlated with regular search results. As these variables are not known to us and are included in the error term, it leads to a biased estimate of the position coefficient.

We do not have the advertiser budget information to model the search engine positioning decisions based on advertiser budget constraints. In order to correct for the endogeneity bias we use the IV approach. We define the IV equation as follows

$$\ln\left(\mathrm{Pos}_{kt}^{a}\right) = \alpha_0^{ka} + \alpha_1^{ka} x_{kt} + \alpha_{\mathrm{Time}}^{a}\mathrm{Time}_{kt} + \varepsilon_{kt}^{\alpha_a} \quad (4)$$

with $\alpha^k = \Delta^{\alpha_a} z_k + u_k^{\alpha_a}$ and $u_k^{\alpha_a} \sim N(0, V^{\alpha_a})$

where x_{kt} are keyword specific instrument variables. One of the drivers of relative arrangement of ads for keywords is the search traffic for the keyword. A variation in this search traffic leads search engine to display more ads by allowing advertisers to exceed their daily budget to maximize the click performance. So we use the search traffic for keywords as an instrument. In

12 http://advertising.microsoft.com/small-business/search-advertising/budgeting-bidding,
http://help.yahoo.com/l/sg/yahoo/ysm/sps/screenref/16720.html#3,
http://adwords.google.com/support/aw/bin/answer.py?hl=en&answer=6312

demand models with endogenous prices, input prices are often used as instruments (Kuksov and Villas Boas, 2008). Lagged prices, lagged shares, cost, and prices in other markets are also used as instruments for endogenous prices (Yang et al. 2003). We also use lagged position as an instrument. In order to account for the effect of competition on the ranking decision we also control for the maximum competitive bid (c_bid). Time variable allows us to control for any time dynamics that can influence the ranking decision. We use a log specification as the position of the ad is the daily average position, and is a continuous variable.

Finally, as the position of the ad depends on the search engine's decision and is endogenous, the unobservable time varying keyword attributes for equations representing consumer decisions to click will be correlated with the error term for position equations representing the search engine decision. As such, we use the following distribution to account for correlation between the error terms for clickthrough rate, and position equations:

(5) $$\begin{bmatrix} \varepsilon_{kt}^{\beta_H} \\ \varepsilon_{kt}^{\beta_L} \\ \varepsilon_{kt}^{\alpha_H} \\ \varepsilon_{kt}^{\alpha_L} \end{bmatrix} \sim N\left[\begin{pmatrix} 0 \\ 0 \\ 0 \\ 0 \end{pmatrix}, \begin{pmatrix} \Omega_{HH}^{cl} & \Omega_{HL}^{cl} & \Omega_{HH}^{cl,pos} & \Omega_{HL}^{cl,pos} \\ \Omega_{HL}^{cl} & \Omega_{LL}^{cl} & \Omega_{LH}^{cl,pos} & \Omega_{LL}^{cl,pos} \\ \Omega_{HH}^{cl,pos} & \Omega_{LH}^{cl,pos} & \Omega_{HH}^{pos} & \Omega_{HL}^{pos} \\ \Omega_{HL}^{cl,pos} & \Omega_{LL}^{cl,pos} & \Omega_{HL}^{pos} & \Omega_{LL}^{pos} \end{pmatrix}\right]$$

Identification

The above set of simultaneous equations represents a triangular system with and has been addressed by authors in classical econometrics (Lahiri and Schmidt 1978, Hausman 1975, Greene 1999) and bayesian econometrics (Zellner 1962). It can be represented as follows

$U_{kt}^{CTR} = f(\text{Position}, X1, \varepsilon_{kt}^{\beta})$

$\text{Position} = f(X2, \varepsilon_{kt}^{\alpha})$

In this construction, position is endogenous while variables X1-X2 are exogenous. Identification comes from the fact that position is completely determined by the exogenous variables traffic and lagged position. Position in turn influences click performance. Thus, the rank and order conditions are satisfied for identification purposes (Greene, 1999).

Lahiri and Schmidt (1978) have shown that the parameter estimates for a triangular system can be fully identified using GLS. Hausman (1975) shows that the likelihood function for a triangular system is the same as for seemingly unrelated regressions. Zellner (1962) has addressed triangular systems from a Bayesian point of view, and shows that the posterior probability distribution function is the same as in a seemingly unrelated regressions setting. Triangular systems have been estimated using the classical approach (Elberse and Eliashberg 2003; Godes and Mayzlin 2004) and more recently in sponsored search using the Bayesian approach (Ghose and Yang 2009, Yang and Ghose 2010, Agarwal, Hosanagar, and Smith, 2011). We estimate the model using a Bayesian approach, applying Markov chain Monte Carlo sampling due to the non-linear characteristics of our model (Rossi and Allenby 2005). The priors and conditional posteriors of this model are discussed in Technical Appendix. For the HB Models we run the MCMC simulation for 80,000 draws, discarding the first 40,000 as burn-in. To ensure that our parameter estimates are accurate we simulated the clicks, bids and positions using our estimates. By repeating the estimation with this simulated dataset we were able to recover our parameter estimates. This indicates that our parameters are fully identified.

Results

Click through rate (CTR)

Table 2 provides the mean values for the posterior distribution of the Δ^{β_a} matrix and the estimates for V^{β_a} from equation 2 for both high quality and low quality advertiser. The coefficient for constant term for the high quality firm is higher as compared to that for the low quality firm and the mean difference test suggests that the difference is significant ($p < 0.01$). This suggests that click performance represents the expected click propensity for firms of different quality. The coefficient for *pos* is negative and significant indicating that click performance decays with position. This confirms the observation made in the previous literature that click performance decreases with position. The coefficient for *pos* is more negative for the high quality firm as compared to the low quality firm and the difference is significant ($p < 0.01$). This suggests that the high quality firm has a higher decay in click performance as compared to the low quality firm. This can be attributed to the search behavior of consumers visiting lower positions. As consumers are expected to search sequentially, one possibility is that the marginal value of an ad at lower position may be lower as consumers maybe satiated with information from ads in the higher position. This reduces the gap in the click probability of advertisers of different quality. Another possibility is that lower position of the advertiser actually lowers the quality perception of the advertiser.

= = Insert Table 2 about here = =

The coefficient for the intercept representing *specificity* is positive and more significant and higher for the high quality advertiser as compared to that for the low quality advertiser. This suggests that more customers are likely to click the high quality advertiser's ad for more specific keywords. According to industry reports, these specific keywords are long tail keywords and are expected to generate higher click through rate as consumers are closer to their purchase decision.[13] Our results suggest that there can be heterogeneity in performance across advertisers depending on their quality. This also explains contradictory results observed by previous single advertiser studies (Ghose and Yang, 2009; Agarwal et al. 2011). Consumers using specific keywords are supposed to be more informed and have less quality uncertainty. The fact that high quality advertiser gets a higher click performance for more specific keywords suggests that the click propensity does reflect the true quality to some extent. Note that we control for the effect of competitive intensity which can lead to higher click performance for more specific keywords irrespective of the quality of the advertiser.

The coefficient for the interaction term between *pos* and *specificity* is negative and significant only for the high quality advertiser. Figure 2 shows the difference in CTR performance of high quality advertiser and low quality advertiser for sample keywords using the posterior distribution of coefficients in equation 2. Thus, the gap in the CTR performance between advertisers at lower positions reduces for more specific keywords. Consumers using such keywords are expected to be less satiated at any given position as they are expected to search more. In that case, the marginal utility of evaluating ads at any position should be higher for these consumers as compared to the ones using less specific keywords. As such, the difference in the quality perception of advertisers should be higher for more specific keywords as compared to that for less specific keywords. However, our results point in the other direction.

13 http://www.calculatemarketing.com/blog/techniques/benefits-of-long-tail-keywords, http://ezinearticles.com/?Targeted-Keywords&id=2997302, http://www.thestreet.com/story/10485554/1/grab-customers-with-long-tail-keywords.html?cm_ven=GOOGLEFI

This suggests that the lower position of the advertiser may be leading to lower quality perception.

= = Insert Figure 2 about here = =

The coefficient for competitive intensity *CI* is negative and significant for both advertisers. This is reasonable as higher competitive intensity should result in lower click through rate. This is consistent with the previous research results on the impact of competition by Animesh, Viswanathan, and Agarwal (2011) and Rutz and Trusov (2011). The covariance matrix shows the significance of heterogeneity in this setup.

Position

Table 3 provides the mean values for the posterior distribution of the Δ^{α_a} matrix and the estimates for $V^{\alpha_a}$from equation 4 for both high quality and low quality advertiser. The coefficient for traffic is positive and significant for both advertisers. This is reasonable as higher traffic would result in search engine providing more choices to consumers. Resulting higher competition would lower the average position of the advertiser. The coefficient for lagged position is also positive and significant for both advertisers. This suggests that factors driving the position in the past period also influence the ad position in the current period. The coefficient for competitive bid is positive and significant. Additionally, the value is higher for the low quality advertiser. This is reasonable as the position of low quality advertiser is more likely to be sensitive to maximum competitive bid due to its lower click propensity. Higher maximum competitive bid should result in lower position.

= = Insert Table 3 about here = =

Finally, Table 4 shows covariance between unobservables for CTR, and ad positions from equations 2 and 4. Covariance between the unobservables for CTR for two advertisers is significant. This indicates that the unknown factors influencing consumer clicks for one advertiser also influence the consumer clicks for the other advertiser. The covariance between the unobservables for position equations for the two advertisers is statistically significant. Similarly, covariance between the unobservables for CTR and position is statistically significant. This correlation between the error terms for CTR with the error term for ad position shows that position is endogenous and the proposed simultaneous equation model helps to capture the endogeneity effect.

= = Insert Table 4 about here = =

Cost per Click (CPC)

We use the estimates from our analysis to determine the CTR performance for both advertisers as a function of position. We want to isolate the position effect due to sequential search from the advertiser effect on the click performance. So we calculate the position normalized CTR for each advertiser for different positions by dividing the CTR for each advertiser in each position by the average CTR for that position. This position normalized CTR for each advertiser for each position represents its position dependent quality. We use the average of position normalized CTR for each advertiser as a measure of its average quality. Cost per click paid by an advertiser at position k in sponsored search auctions is the bid for advertiser in position k+1 weighted by the expected quality of the advertiser in position k (Lahaie and Pennock 2007).

$$CPC_k \propto \frac{Quality_{k+1}}{Quality_k} Bid_{k+1}$$

If advertiser in position k+1 is the last advertiser then it can pay a minimum bid (Lahaie and Pennock 2007). given as

$CPC_{k+1} \propto \frac{1}{Quality_{k+1}} r$ where r is the reservation price

We calculate the cost per click for each advertiser using either the average quality or position based quality. In our calculation, we assume that the other advertiser is in the next position and is going to be charged the minimum bid. Figure 3 shows the CPC change by introducing position based quality for each advertiser instead of using average quality. We can see that the high quality advertiser is being overcharged and low quality advertiser is being undercharged at higher positions using the average quality as a measure of performance. The situation is reversed for the lower position. The results show that cost per click can deviate by as much as 50%

= = Insert Figure 3 about here = =

Robustness of Results

In this section we outline several steps we have taken to evaluate the robustness of our results.

Holdout Sample Analysis: We have attempted to verify the prediction accuracy of our results using a holdout sample. To do this, we consider data for the first week as the estimation sample and data for the remaining next week as the holdout sample. We use mean absolute percentage error (MAPE) for daily values at the aggregate level and the keyword level. Error values are reported in Table 5 and indicate that the model prediction accuracy is similar for both the estimation and holdout samples, suggesting that our model estimates are robust.

= = Insert Table 5 about here = =

Other Advertiser Keyword Combinations: In order to verify that our results apply to other advertisers as well, we have conducted a similar analysis for additional advertiser pairs competing for certain keywords. We randomly selected a sample consisting of 33 advertiser pairs which compete for 123 keywords. These keywords represent 5 different categories of durable goods. We verify that the same keyword is not represented by more than one advertiser pair. However, the same advertiser pair could be competing on multiple keywords. We use the same criteria to select this sample as our main sample. We label advertisers in each advertiser pair as high quality and low quality based on their past click propensity. We consider a period of few weeks where bids do not change for the corresponding keywords. We use equations 1-5 to model the click performance as well as account for the endogeneity of positions for both high quality and low quality advertiser. The effect of any unobservable advertiser characteristics is addressed by accounting for keyword heterogeneity as there is only one advertiser pair for each keyword and we have separate equations for each advertiser. We find that the results hold for this data sample. Tables 6-8 show the results for this sample. The coefficient for position is negative and significant for both advertisers (Table 6). However, the coefficient for high quality advertiser is more negative and the difference is significant ($p < 0.01$). Additionally, the coefficient of the interaction between position and specificity is significant and negative for the high quality advertiser.

= = Insert Tables 6-8 about here = =

Discussion and Conclusion

Previous work in sponsored search auctions has established that the click performance decays with the position of the ad. However, the variation of click performance across advertisers is not very well understood. This is important to verify the efficacy of the auction design which serves advertisers of different perceived quality and for advertisers to come up with appropriate advertising strategies. We investigate the effect of advertiser quality by empirically analyzing the click performance of two competing advertisers of different quality over several hundred keywords. We find that the click performance decays at a higher rate for the higher quality advertiser. We also find that the decay rate increases further for the higher quality advertiser for more specific keywords. These results suggest that the advertiser quality is being evaluated differently by consumers visiting the top positions and the lower positions. Additionally, consumers using more specific keywords maybe further de-emphasizing quality at lower position.

Our study has several implications. We provide insight into the consumer clicking behavior in sponsored search and its impact on the performance of advertisers with different perceived quality levels. We show that that perceived quality of advertisers depends on the extent of search and the perceived quality difference reduces at lower position. This maybe attributed to information satiation after consumers have evaluated ads sequentially from the top. However, we also find that consumers, who are likely to search more and more certain of the quality difference, are more likely to de-emphasize quality difference at lower positions. This suggests that the quality perception itself maybe a function of position. Our results suggest that advertisers should consider their position based performance and not rely on a single quality measure. High quality advertisers may want to remain in the top position as the perceived quality decreases at a higher rate for them as compared to the low quality advertisers. Previous theory work has suggested that low quality advertisers bid higher to be in the top position. This assumes that the quality perception is independent of position. However, our results show that it maybe beneficial for low quality advertisers to be in lower positions as the brand recognition is low at these positions.

Our results also point to the inefficiency in the current auction mechanism where advertisers' position based performance is not considered for the evaluation of quality. We show that this can result in overcharging of high quality advertisers for top position and undercharging low quality advertisers for these positions. This may lead to an increase in participation by the low quality advertisers. However, if the low quality perception also represents the true quality of advertiser then this can result in adverse selection. This can reduce consumer participation in the long run. As a result, it is important for the search engine to determine if the perceived quality represented by the click propensity is an adequate measure to determine the true quality of advertiser. One way to do this is to get conversion information from the advertisers. Additionally, the search engine ranking mechanism should consider keyword characteristics for determining the ranking of advertisers. Currently, all keywords are treated in the same way when it comes to ranking advertisers.[14] For example, the search engine could use different weights for bid and performance for different keywords to discourage lower quality advertisers from bidding high and appearing at the top position. This would reduce adverse selection for less specific keywords. Also, this will ensure that the low quality advertisers appear in lower positions and pay according to their expected performance. Additionally, search engines can inform advertisers of the consequences of bidding too high to compensate for their low quality in order to get a higher position. This will ensure advertiser participation in the long run.

Our study has several limitations. We use a binary classification for advertisers into high and low quality instead of using a continuous measure of quality. While this allows us to evaluate the effect of position on click performance of these advertisers, it does not allow us to comment on the relative importance of the magnitude of quality and position. We also do not have information about the advertiser attributes that contribute to this quality. These attributes include existing brand recognition as well as appearance in organic results. Future studies should investigate the impact of each of these advertiser characteristics on the click performance for

---

14 https://adwords.google.com/support/aw/bin/answer.py?hl=en&answer=115967http://help.yahoo.com/l/sg/yahoo/ysm/sps/start/overview_qualityindex.html

different types of keywords. We do not have information about the unique selling proposition of the ads and how it contributes to the perception of quality. Animesh, Viswanathan, and Agarwal (2011) find that an ad with price as the unique selling proposition is less affected by the competition. In our setup, we do not expect the unique selling proposition of ads to change over time. Our model also accounts for heterogeneity across keywords which controls for the effect of the unique selling proposition of ad corresponding to each keyword. However, future studies should explicitly evaluate the interplay of advertiser's unique selling proposition with its quality perception. We do not consider the dynamic effect of advertising which can lead to general awareness for the advertiser and in turn lead to higher quality perception. We also do not explicitly consider the correlation of performance across keywords. Future study should look at explicitly modeling the dynamics of advertising as well as consider the impact of correlation across keywords on the click performance of individual keywords. Finally, we do not have information about the conversions at the advertiser website after consumers click on ads. This is a limitation of any such dataset obtained from search engines that cannot get conversion information from advertisers. This should not affect our results as higher click performance leads to higher conversions. However, future research should explicitly model the effect of search characteristics and advertiser quality on conversion performance.

References

Agarwal, A., K. Hosanagar, M. D. Smith (2011), "Location, location and location: An Analysis of Profitability of Position in Online Advertising Markets," *Journal of Marketing Research*, Forthcoming.

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

Animesh, A., V. Ramachandran, and S. Viswanathan (2010), "Quality Uncertainty and Adverse Selection in Online Sponsored Search Markets," *Information Systems Research*, 21, 190-201.

Animesh, A., S. Viswanathan, and R. Agarwal (2011), "Competing Creatively in Sponsored Search Markets," *Information Systems Research*, 22, 153-169.

Arbatskaya, M. (2007), "Ordered Search," *Rand Journal of Economics*, 38(1), 119-126.

Athey, S., G. Ellison (2011), "Position auctions with consumer search," *Quarterly Journal of Economics*, Forthcoming.

Brunel, F. F. and M. R. Nelson (2003), "Message Order Effects and Gender Differ-ences in Advertising Persuasion," *Journal of Advertising Research*, 43(3), 330-341.

Brucks, M. (1985), "The Effects of Product Class Knowledge on Information Search Behavior," *Journal of Consumer Research*, 12, 1–16.

Brynjolfsson, E., and M.D. Smith (2000), "Frictionless Commerce? A Comparison of Internet and Conventional Retailers," *Management Science,* 46(4) pp 563-585.

Brynjolfsson, E., A. A. Dick, and M.D. Smith. (2009), "A Nearly Perfect Market? Differentiation vs. Price in Consumer Choice," *Quantitative Marketing and Economics*, 8(1), 1-33.

Chen, J., J. Feng, and A. B. Whinston (2009), "Keyword Auctions, Unit-Price Contracts, and the Role of Commitment," *Production and Operations Management* 19(3), pp. 305–321.

Demers, E., and B. Lev(2001), "A Rude Awakening: Internet Shakeout in 2000," *Review of Accounting Studies* (6:2/3), pp. 331-359.

Dodds. W. B., K. B. Monroe, and D. Grewal (1991), "Effects of Price, Brand, and Store Information on Buyers' Product Evaluations," *Journal of Marketing Research*, 28(9), 307-19.

Dou, W.; K. H. Lim, C. Su, N. Zhou, and N. Cui (2010), "Brand Positioning Strategy Using Search Engine Marketing," *MIS Quarterly*, (34: 2) pp.261-279.

Dube JP; GH Hitsch, and P.E. Rossi (2010), "State Dependence and Alternate Explanations for Consumer Inertia," *Rand Journal of Economics*, 41(3), 417-445.

Elberse, A., J. Eliashberg (2003) Demand and Supply Dynamics for Sequentially Released Products in International Markets: The Case of Motion Pictures. *Marketing Science*, 22(3), 329-354.

EMarketer (2009). US Online Ad Spend Turns the Corner. *www.emarketer.com*, December.

Erdem, T., M. Keane and B. Sun (2008), "A Dynamic Model of Brand Choice When Price and Advertising Signal Product Quality", *Marketing Science*, 27 (6), 1111-1125.

Feng, F., Bhargava, H., and Pennock, D. (2007). Implementing Sponsored Search in Web Search Engines: Computational Evaluation of Alternative Mechanisms. *Informs Journal on Computing*, 19(1),137-148.

Gerath, K., L. Ma, YH. Park, and K. Srinivasan (2011), "A Position Paradox in Sponsored Search Auctions," *Marketing Science*, Forthcoming

Ghose, A. and S. Yang (2009), "An Empirical Analysis of Search Engine Advertising: Sponsored Search in Electronic Markets," *Management Science*, 55(10), 1605-1622.

Godes, D., and D. Mayzlin (2004), "Using Online Conversations to Study Word of Mouth Communication," *Marketing Science*, 23 (4), 545-560.

Granka, L. A., T. Joachims, and G. Gay (2004), "Eye-Tracking Analysis of User Behavior in WWW Search," *Proceedings of the 27th Annual International ACM SIGIR Conference on Research and Development in Information Retrieval*, July 25-29, Sheffield, UK.

Greene, W. (1999), "Econometric Analysis," 4th ed. Prentice Hall, Upper Saddle River, NJ.

Hausman, J. (1975), "An instrumental variable approach to full information estimators for linear and certain nonlinear econometric models," *Econometrica*, 43, 727–738.

Hoque, A. Y. and G. L. Lohse (1999), "An Information Search Cost Perspective for De-signing Interfaces for Electronic Commerce," *Journal of Marketing Research*, 36(3), 387-394.

Hosanagar, K. and V. Cherapanov (2008), "Optimal Bidding in Stochastic Budget Con-strained Slot Auctions," *Proceedings of ACM Conference on Electronic Commerce*, Chicago, USA.

Jeziorski, P., and R. I. Segal (2009), "What Makes them Click: Empirical Analysis of Consumer Demand for Search Advertising", *Working Paper*, Stanford University

Johnson, E. J., W. W. Moe, P. S. Fader, S. Bellman, G. L. Lohse (2004), "On the depth and dynamics of online search behavior," *Management Science*, 50(3), 299-308.

Lahaie, S. and D. M. Pennock (2007), "Revenue Analysis of a Family of Ranking Rules for Keyword Auctions", *Proc. of the 8th ACM Conference on Electronic Commerce*, San Diego, CA

Lahiri, K. and P. Schmidt (1978), "On the estimation of triangular structural systems," E*conometrica,* 46(5), 1217–1221.

Liu, D., J. Chen, and A. B. Whinston (2010), "Ex Ante Information and the Design of Keyword Auctions", *Information Systems Research*, 21(1), 133-153.

Kent, R. J. and C. T, Allen (1994), "Competitive Interference Effects in Consumer Memory tor Advertising: The Role of Brand Familiarity," *Journal of Marketing*, 58 (7), 97-105.

Kirmani, A., and Wright, P. (1989) Money Talks: Perceived Advertising Expense and Expected Product Quality. *Journal of Consumer Research*,(16), , pp. 344-353.

Kuksov, D. and J. M. Villas-Boas (2008), "Endogeneity and Individual Consumer Choice," *Journal of Marketing Research,* 10(6), 702-714.

Moe,W.W. (2003), "Buying, searching, or browsing: Differentiating between online shoppers using in-store navigational clickstream," *Journal of Consumer Psychology*, 13(1, 2) 29–40.

Moe, W. W. and P. S. Fader(2004), "Dynamic Conversion Behavior at e-Commerce Sites," *Management Science*, 50 (3), 326-335.

Montgomery, A. L., K. Hosanagar, R. Krishnan, and K. B. Clay (2004), "Designing a Better Shopbot," *Management Science*, 50(2) 189–206.

Montgomery, A. L., S. Li, K. Srinivasan, and J. C. Liechty (2004), "Modeling Online Browsing and Path Analysis Using Clickstream Data," *Marketing Science*, 23(4), 579-595.

Moorthy, S., B. T. Ratchford, and D. Talukdar (1997), "Consumer information search revisited: Theory and empirical analysis," *Journal of Consumer Research*, 23, 263–277.

Moorthy, S. and H. Zhao (2000), "Advertising Spending and Perceived Quality," *Marketing Letters*, 11 (3), 221-33.

Palmer, J.W. (2002), "Web site usability, design, and performance metrics," *Information Systems Research,*13(2), pp 150-167.

Park, C.H. and Y-H Park (2010), "Modeling Navigational Behavior in Sponsored Search Advertising," http://papers.ssrn.com/sol3/papers.cfm?abstract_id=1565489

Rhodes, E. W., N. B. Teferman, E. Cook, and D. Schwartz (1979), "T-Scope tests of Yellow Pages Advertising," *Journal of Advertising Research*, 19, 49-52.

Rossi, P. E., G. M. Allenby (2003), "Bayesian Statistics and Marketing," *Marketing Science*, 22(3), 304-28.

___ and G. M. Allenby (2005), *Bayesian Statistics and Marketing*, West Sussex, England, John Wiley and Sons.

Rutz, O. and R. Bucklin (2011), "From Generic to Branded: A Model of Spillover s in Paid Search Advertising," *Journal of Marketing Research*, 48(1), 87-102.

Rutz, O. and M. Trusov (2011), "Zooming In on Paid Search Ads - A Consumer-level Model Calibrated on Aggregated Data," *Marketing Science*, forthcoming.

Srinivasan, N., and B. T. Ratchford (1991), "An empirical test of a model of external search for automobiles," *Journal of Consumer Research*, 18, 233–242.

Trueman, B., M.H.F. Wong, and X. J. Zhang (2001), "Back to Basics: Forecasting the Revenues of Internet Firms," *Review of Accounting Studies*, Springer Science & Business Media B.V., pp. 305-329.

Urbany. J. E., P. R. Dicksoti, and W. L. Wilkie (1989), "Buyer Uncertainty and Information Search," *Journal of Consumer Research*. 16 (September), 208- 215.

White, R.W. and D. Morris (2007), "Investigating the querying and browsing behavior of advanced search engine users," Proc. SIGIR, 255-262.

White, R.W., S. T. Dumais, and J. Teevan (2009), "Characterizing the Influence of Domain Expertise on Web Search Behavior," *ACM International Conference on Web Search and Data Mining*, February 9-12, Barcelona, Spain.

Yang , S., Y. Chen, G. M. Allenby (2003), "Bayesian analysis of simultaneous demand and supply," *Quantitative Marketing Economics*, 1(3), 251–275.

Yang, S. and A. Ghose (2010), "Analyzing the Relationship Between Organic and Sponsored Search Advertising: Positive, Negative, or Zero Interdependence? *Marketing Science*, 29(4), 602-623.

Zellner, A. (1962). An Efficient Method of Estimating Seemingly Unrelated Regressions and Tests for Aggregation Bias. *Journal of the American Statistical Association*, 57, 348-368.

Table 1: Summary Statistics

Variable	Description	Mean	Std Dev	Min	Max
High Quality Advertiser					
Impressions	Impressions received by the advertiser for a keyword on a daily basis	37	106	1	2289
Clicks	Clicks received by the advertiser for a keyword on a daily basis	0.75	1.69	0	28
Pos	Average daily position of advertiser for a keyword	3.43	1.95	1	21
CI	Number of competing advertisers for a keyword	6	2.9	1	26
Log(Traffic)	Overall impressions received by a keyword	3.18	1.35	0.69	8.66
Log(C_Bid)	Highest competing bid for the advertiser for a keyword	1.09	0.53	0	3.04
Low Quality Advertiser					
Impressions		65	215	1	3391
Clicks		0.53	2.1	0	110
Pos		5.57	2.3	1	21
CI		6	2.9	1	26
Log(Traffic)		3.18	1.35	0.69	8.66
Log(C_Bid)		1.63	0.43	0	3.02
Keyword					
Specificity	Measure of how specific is a keyword	2.2	1.1	1	5

Table 2: Parameter Estimates for CTR

Variables	**High Quality Advertiser**		**Low Quality Advertiser**	
	Intercept	**Specificity**	**Intercept**	**Specificity**
Const	-2.35 (0.41)***	0.3 (0.1)***	-3.57 (0.59)***	0.15 (0.08)*
Pos	-0.25 (0.04)***	-0.06 (0.02)***	-0.15 (0.03)***	0.01 (0.01)
CI	-0.06 (0.02)***	0.0 (0.02)	-0.055 (0.02)***	-0.02 (0.02)
Time	-0.01 (0.001)***		-0.002 (0.00)***	

V^{β_H}	**Const**	**Pos**	**CI**
Const	0.91 (0.1)***	-0.04 (0.01)***	-0.05 (0.01)***
Pos		0.05 (0.01)***	-0.01 (0.0)***
CI			0.04 (0.01)***

V^{β_L}	**Const**	**Pos**	**CI**
Const	1.08 (0.07)***	-0.109 (0.008)***	-0.076 (0.007)***
Pos		0.063 (0.007)***	-0.005 (0.002)**
CI			0.043 (0.003)***

Table 3: Parameter Estimates for Position

Variables	High Quality Advertiser		Low Quality Advertiser	
	Intercept	Specificity	Intercept	Specificity
Const	0.76 (0.12)***	-0.02 (0.04)	1.21 (0.09)***	0.03 (0.04)
Traffic	0.1 (0.02)***	0.0 (0.01)	0.06 (0.03)**	0.0 (0.01)
Lagged Position	0.09 (0.02)***	-0.01 (0.02)	0.14 (0.03)***	0.02 (0.02)
Comp_Bid	0.02 (0.013)*		0.05 (0.01)***	
Time	-0.003 (0.001)***		-0.002 (0.00)***	

V^{α_H}	Const	Traffic	LaggedPosition
Const	0.23 (0.03)***	-0.04 (0.01)***	-0.03 (0.01)***
Traffic		0.03 (0.0)***	-0.01 (0.0)***
LaggedPosition			0.08 (0.01)***

V^{α_L}	Const	Traffic	LaggedPosition
Const	0.23 (0.03)***	-0.03 (0.01)***	-0.04 (0.01)***
Traffic		0.03 (0.0)***	-0.01 (0.0)***
LaggedPosition			0.06 (0.01)***

Table 4: Parameter Estimates for Omega

	High Quality CTR	Low Quality CTR	High Quality Position	Low Quality Position
High Quality CTR	0.24 (0.01)***	0.08 (0.01)***	0.02 (0.01)**	0.02 (0.01)**
Low Quality CTR		0.16 (0.01)***	0.02 (0.007)***	0.02 (0.01)**
High Quality Position			0.12 (0.0)***	0.01 (0.0)***
Low Quality Position				0.06 (0.0)***

Table 5: Prediction Accuracy For Estimation And Holdout Samples

	CTR Fit (MAPE)	
	Aggregate	Keyword
Estimation Sample	.43	.41
Holdout Sample	.44	.43

Note: Aggregate MAPE is the average MAPE across all datapoints. Keyword MAPE is the average of the average MAPE for different keywords

Table 6: Parameter Estimates for CTR for other Advertiser Keyword Pairs

Variables	High Quality Advertiser		Low Quality Advertiser	
	Intercept	Specificity	Intercept	Specificity
Const	-1.6 (0.24)***	0.3 (0.11)**	-2.07 (0.28)***	0.22 (0.13)*
Pos	-0.47 (0.06)***	-0.11 (0.05)**	-0.19 (0.05)***	0.01 (0.04)
CI	-0.01 (0.03)	-0.02 (0.03)***	0.01 (0.03)***	-0.05 (0.03)*
Time	-0.05 (0.6)		0.3 (0.7)	
V^{β_H}	**Const**	**Pos**	**CI**	
Const	0.58 (0.12)***	-0.08 (0.03)**	-0.05 (0.02)**	
Pos		0.15 (0.02)***	-0.02 (0.01)**	
CI			0.07 (0.01)***	
V^{β_L}	**Const**	**Pos**	**CI**	
Const	1.04 (0.15)***	-0.06 (0.02)***	-0.1 (0.02)***	
Pos		0.12 (0.02)***	-0.02 (0.01)**	
CI			0.08 (0.01)***	

Table 7: Parameter Estimates for Position for other Advertiser Keyword Pairs

Variables	High Quality Advertiser		Low Quality Advertiser	
	Intercept	Specificity	Intercept	Specificity
Const	-1.75(1.44)	-2.55 (13.89)	-0.23 (0.21)***	0.01 (0.07)
Traffic	0.06 (0.03)**	-0.02 (0.03)	0.1 (0.03)***	-0.01 (0.03)
Lagged Position	0.14 (0.05)***	0.03 (0.05)	0.13 (0.04)***	-0.04 (0.04)
Comp_Bid	0.1 (0.03)***		0.05 (0.01)***	
Time	0.001 (0.001)		-0.001 (0.001)	
V^{α_H}	**Const**	**Traffic**	**LaggedPosition**	
Const	2.4 (0.53)***	-2.3 (4.62)	4.2 (6.79)	
Traffic		0.07 (0.01)***	0.0 (0.01)	
LaggedPosition			0.12 (0.02)***	
V^{α_L}	**Const**	**Traffic**	**LaggedPosition**	
Const	0.3 (0.05)***	-0.04 (0.02)***	-0.05 (0.02)***	
Traffic		0.07 (0.01)***	-0.01 (0.01)	
LaggedPosition			0.12 (0.02)***	

Table 8: Parameter Estimates for Omega for other Advertiser Keyword Pairs

	High Quality CTR	Low Quality CTR	High Quality Position	Low Quality Position
High Quality CTR	0.173 (0.017)***	0.019 (0.007)	0.015 (0.01)	0.018 (0.005)***
Low Quality CTR		0.132 (0.007)***	0.008 (0.004)**	0.017 (0.008)**
High Quality Position			0.11 (0.004)***	0.002 (0.003)
Low Quality Position				0.10 (0.003)***

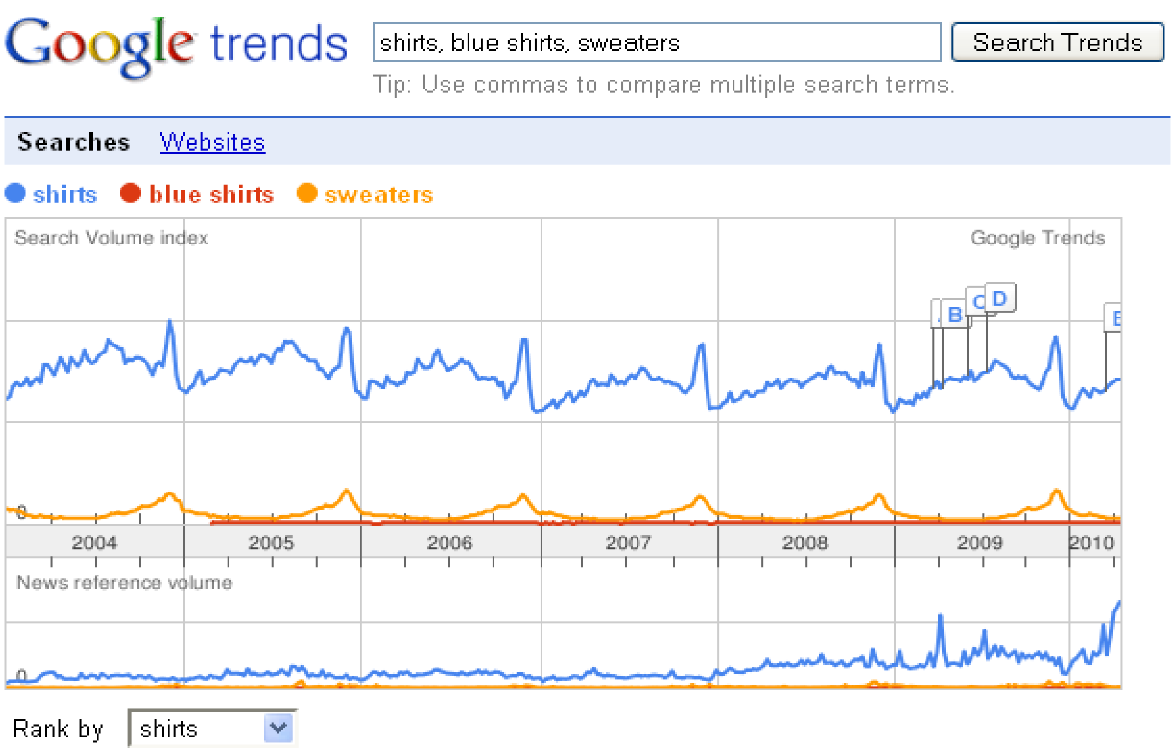

Figure 1: Search Trend for Sample Keywords

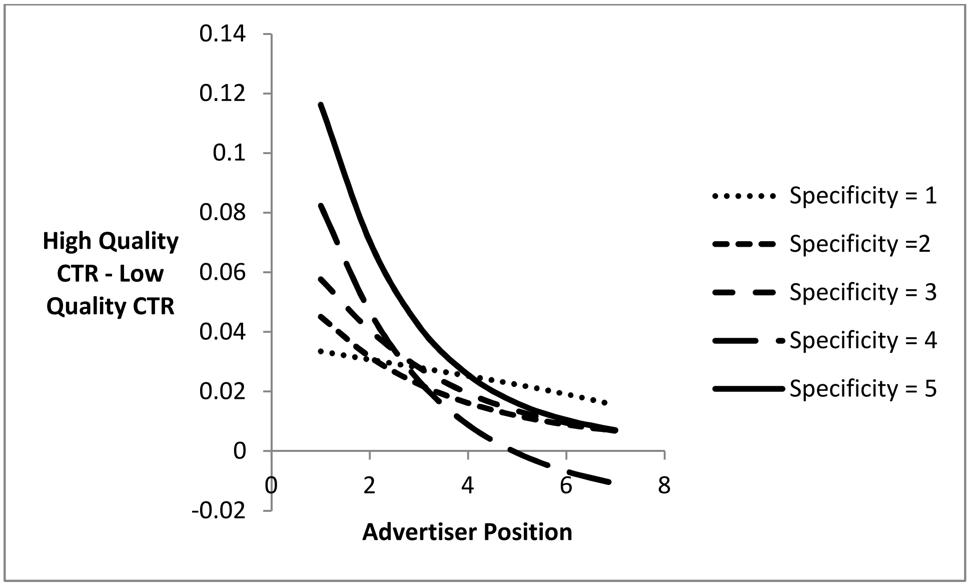

Figure 2: Click Performance difference between High Quality Advertiser and Low Quality Advertiser for Sample Keywords

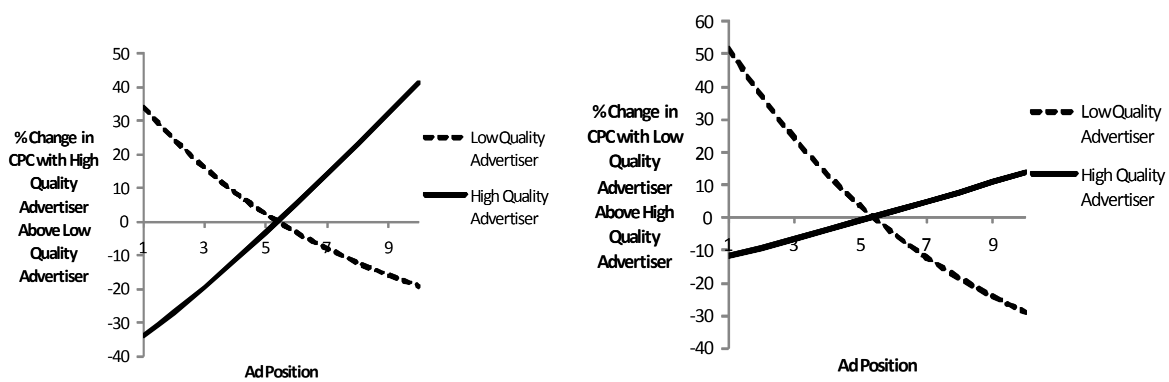

Figure 3: Change in CPC due to Position based measure of quality instead of an average measure of Quality

ONLINE APPENDIX

Advertiser Quality

The clickthrough rate for an advertiser ***a*** for a particular keyword ***k*** at a particular position ***r*** at time ***t*** can be expressed as

$CTR_{kart} = c_k q_{kat} \theta^r$ where θ^r is the base click through rate as a function of position ***r***, $c_k$represent the keyword specific effect, and q_{kat} captures the advertiser specific effect for that keyword in period ***t***.

Then, the normalized CTR can be expressed as

$NCTR_{kart} = \frac{c_k q_{kat} \theta^r - c_k q_{kt} \theta^r}{\sqrt{\frac{1}{N_{kt}} \sum_{a \in A_{kt}} (c_k q_{kat} \theta^r - c_k q_{kt} \theta^r)^2}} = \frac{q_{kat} - q_{kt}}{\sqrt{\frac{1}{N_{kt}} \sum_{a \in A_{kt}} (q_{kat} - q_{kt})^2}}$ where $q_{kt} = \frac{1}{N_{kt}} \sum_{a \in A_{kt}} q_{kat}$

and N_{kt} is the number of advertisers for keyword ***k*** in period ***t*** and A_{kt} represents all advertisers belonging to keyword ***k*** in period ***t***

Note that the effect of position as well as keyword disappears with this approach. Additionally, normalization with variance takes away the effect of differences in the distribution of advertisers' performance across positions for each keyword and also across keywords. The click propensity for an advertiser ***a*** for a keyword ***k*** can be expressed as

$AdQuality_{ka} = \frac{1}{T} \sum_t \frac{1}{N_{kat}} \sum_r NCTR_{kart}$ where N_{kat} is the number of different positionss attained by advertiser ***a*** for keyword ***k*** in period **t** and T is the total number of periods.

MCMC Algorithm

The model can be written in the hierarchical form:

$U_{kt}^{H} | \beta^{kH}, \beta^{H}, X_{kt}^{\beta^{kH}}, X_{kt}^{\beta^{H}}, \Omega$

$U^{L}_{kt}|\beta^{kL}, \beta^{L}, X^{\beta^{kL}}_{kt}, X^{\beta^{L}}_{kt}, \Omega$

$b_k|\{U^{H}_{kt}\}, \{U^{L}_{kt}\}, X^{b^k}_{kt}, X^{b}_{kt}, b, z, \Delta^b, V^b, \Omega$

$b|\{U^{H}_{kt}\}, \{U^{L}_{kt}\}, \{b_k\}, X^{b^k}_{kt}, X^{b}_{kt}, V^b, \Omega, \bar{b}$

$\Omega|\{U^{H}_{kt}\}, \{U^{L}_{kt}\}, \{b^k\}, \{b\}, X^{b^k}_{kt}, X^{b}_{kt}, \nu_\Omega, S_\Omega$

$V^b|\{b_k\}, z, \Delta^b, \nu, S$

$\Delta^b|\{b_k\}, z, A_0, \overline{\Delta^b}$

where

$b_k = [\beta^{kH}\ \beta^{kL}\ \alpha^{kH}\ \alpha^{kL}]$, $b = [\beta^{H}\beta^{L}\ \alpha^{H}\ \alpha^{L}]$, $V^b = [V^{\beta_H}\ V^{\beta_L}\ V^{\alpha_H}\ V^{\alpha_L}]$, $\Delta^b = [\Delta^{\beta_H}\ \Delta^{\beta_L}\ \Delta^{\alpha_H}\ \Delta^{\alpha_L}]$, $\bar{b} = [\overline{\beta_H}\ \overline{\beta_L}\ \overline{\alpha_H}\ \overline{\alpha_L}]$, $\overline{\Delta^b} = [\overline{\Delta^{\beta_H}}\ \overline{\Delta^{\beta_L}}\ \overline{\Delta^{\alpha_H}}\ \overline{\Delta^{\alpha_L}}]$ and where

$X^{b^k}_{kt}$ are independent variables with keyword specific coefficients and X^{b}_{kt} are independent variables with common coefficients in equations 2 and 4.

We have used 0 as the initial value for elements of b_k,b, Δ^b and an identity matrix as an initial value for elements of V^b.

The MCMC algorithm is described below.

Step I: DrawU^{L}_{kt} & U^{H}_{kt}

We use a data augmentation approach and random walk Metropolis-Hastings algorithm for sampling

$U_{kt} = (U_{kt}^{L}, U_{kt}^{H})$ (Rossi & Allenby, 2005)

${U_{kt}^{H}}^{new} = {U_{kt}^{H}}^{old} + \delta^{H}$ where $\delta^{H} \sim N(0, .02I)$

${U_{kt}^{L}}^{new} = {U_{kt}^{L}}^{old} + \delta^{L}$ where $\delta^{L} \sim N(0, .02I)$

The draws are accepted with a probability α where

$$\alpha = \min\left[\frac{\exp\left[-\frac{1}{2}\left(U_{kt}^{new} - X_{kt} - E_{kt}\right)' A\left(U_{kt}^{new} - X_{kt} - E_{kt}\right)\right] l(U_{kt}^{new})}{\exp\left[-\frac{1}{2}\left(U_{kt}^{old} - X_{kt} - E_{kt}\right)' A\left(U_{kt}^{old} - X_{kt} - E_{kt}\right)\right] l(U_{kt}^{old})}, 1\right]$$

and where $l(U_{kt})$ is the likelihood of clicks for both advertisers

$$l(U_{kt}) = \prod_{k=1}^{K} \prod_{t=1}^{T} (\Lambda_{kt}^{H})^{Clicks_{kt}^{H}} \left(1 - \Lambda_{kt}^{H}\right)^{Impressions_{kt}^{H} - Clicks_{kt}^{H}} (\Lambda_{kt}^{L})^{Clicks_{kt}^{L}} \left(1 - \Lambda_{kt}^{L}\right)^{Impressions_{kt}^{L} - Clicks_{kt}^{L}}$$

and $X_{kt} = \begin{bmatrix} \beta^{kH} X_{kt}^{\beta^{kH}} + \beta^{H} X_{kt}^{\beta^{H}} \\ \beta^{kL} X_{kt}^{\beta^{kL}} + \beta^{L} X_{kt}^{\beta^{L}} \end{bmatrix}$

$e_{kt} = \begin{bmatrix} e_{kt}^{1} \\ e_{kt}^{2} \end{bmatrix}$ where $e_{kt}^{1} = \ln\left(pos_{kt}^{H}\right) - \alpha^{kH} X_{kt}^{\alpha^{kH}} - \alpha^{H} X_{kt}^{\alpha^{H}}$ & $e_{kt}^{2} = \ln\left(pos_{kt}^{L}\right) - \alpha^{kL} X_{kt}^{\alpha^{kL}} - \alpha^{L} X_{kt}^{\alpha^{L}}$

$$E_{kt} = W_{12} W_{22}^{-1} e_{kt}$$

$$A^{-1} = W_{11} - W_{12} W_{22}^{-1} W_{21}$$

$W_{11} = \begin{bmatrix} \Omega_{HH}^{cl} & \Omega_{HL}^{cl} \\ \Omega_{HL}^{cl} & \Omega_{LL}^{cl} \end{bmatrix}$, $W_{22} = \begin{bmatrix} \Omega_{HH}^{pos} & \Omega_{HL}^{pos} \\ \Omega_{HL}^{pos} & \Omega_{LL}^{pos} \end{bmatrix}$, $W_{12} = W_{21} = \begin{bmatrix} \Omega_{HH}^{cl,pos} & \Omega_{HL}^{cl,pos} \\ \Omega_{LH}^{cl,pos} & \Omega_{LL}^{cl,pos} \end{bmatrix}$

Step II: Draw $b_k = [\beta^{kH}\ \beta^{kL}\ \alpha^{kH}\ \alpha^{kL}]$

We define

$$x_k = \begin{bmatrix} X_k^{\beta^{kH}\prime} & 0 & 0 & 0 \\ 0 & X_k^{\beta^{kL}\prime} & 0 & 0 \\ 0 & 0 & X_k^{\alpha^{kH}\prime} & 0 \\ 0 & 0 & 0 & X_k^{\alpha^{kL}\prime} \end{bmatrix} \qquad y_k = \begin{bmatrix} U_{kt}^{H} - \beta^{H} X_{kt}^{\beta^{H}} \\ U_{kt}^{L} - \beta^{L} X_{kt}^{\beta^{L}} \\ \ln\left(pos_{kt}^{H}\right) - \alpha^{H} X_{kt}^{\alpha^{H}} \\ \ln\left(pos_{kt}^{L}\right) - \alpha^{L} X_{kt}^{\alpha^{L}} \end{bmatrix}, \qquad V = \begin{bmatrix} V^{\beta_H} & 0 & 0 & 0 \\ 0 & V^{\beta_L} & 0 & 0 \\ 0 & 0 & V^{\alpha_H} & 0 \\ 0 & 0 & 0 & V^{\alpha_L} \end{bmatrix}, \overline{b_k} = \begin{bmatrix} \Delta^{\beta_H} z_k \\ \Delta^{\beta_L} z_k \\ \Delta^{\alpha_H} z_k \\ \Delta^{\alpha_L} z_k \end{bmatrix}$$

$Q_k = [(x_k'\Omega x_k)^{-1} + V^{-1}]^{-1}$ & $\widetilde{b_k} = Q_k\left[x_k'\Omega^{-1} y_k + V^{-1}\overline{b_k}\right]$

Then $b_k \sim N(\widetilde{b_k}, Q_k)$

Step III: Draw $b = [\beta^H \beta^L\ \alpha^H\ \alpha^L]$

We define

$$x = \begin{bmatrix} X_k^{\beta^{H}\prime} & 0 & 0 & 0 \\ 0 & X_k^{\beta^{L}\prime} & 0 & 0 \\ 0 & 0 & X_k^{\alpha^{H}\prime} & 0 \\ 0 & 0 & 0 & X_k^{\alpha^{L}\prime} \end{bmatrix} \qquad y = \begin{bmatrix} U_{kt}^{H} - \beta^{kH} X_{kt}^{\beta^{kH}} \\ U_{kt}^{L} - \beta^{kL} X_{kt}^{\beta^{kL}} \\ \ln\left(pos_{kt}^{H}\right) - \alpha^{kH} X_{kt}^{\alpha^{kH}} \\ \ln\left(pos_{kt}^{L}\right) - \alpha^{kL} X_{kt}^{\alpha^{kL}} \end{bmatrix}, \ \overline{V} = 100I, \overline{b} = \begin{bmatrix} 0 \\ 0 \\ 0 \\ 0 \end{bmatrix}$$

$Q = [(x'\Omega x)^{-1} + \overline{V}^{-1}]^{-1}$ & $\tilde{b} = Q\left[x'\Omega^{-1} y + \overline{V}^{-1}\overline{b}\right]$

Then $b \sim N(\tilde{b}, Q)$

Step IV: Draw Ω

$\Omega \sim IW(\nu_\Omega + N, \sum_{k=1}^{K}\sum_{t=1}^{T} Y_{kt}' Y_{kt} + S_\Omega)$ where $Y_{kt} = \begin{bmatrix} U_{kt}^{H} - \beta^{kH} X_{kt}^{\beta^{kH}} - \beta^{H} X_{kt}^{\beta^{H}} \\ U_{kt}^{L} - \beta^{kL} X_{kt}^{\beta^{kL}} - \beta^{L} X_{kt}^{\beta^{L}} \\ \ln\left(pos_{kt}^{H}\right) - \alpha^{kH} X_{kt}^{\alpha^{kH}} - \alpha^{H} X_{kt}^{\alpha^{H}} \\ \ln\left(pos_{kt}^{L}\right) - \alpha^{kL} X_{kt}^{\alpha^{kL}} - \alpha^{L} X_{kt}^{\alpha^{L}} \end{bmatrix}$, N = No of observations, $\nu_\Omega = 10$, $S_\Omega = 10I$

Step V: Draw $V^b = \left[V^{\beta_H}\ V^{\beta_L}\ V^{\alpha_H}\ V^{\alpha_L}\right]$

$V^{\beta_H} \sim IW(\nu + N, \sum_{k=1}^{K} \left(\beta^{kH} - \Delta^{\beta_H} z_k\right)' \left(\beta^{kH} - \Delta^{\beta_H} z_k\right) + S)$

where N = No of keywords, $\nu = 10, S = 10I$

$V^{\beta_L} \sim IW(\nu + N, \sum_{k=1}^{K} \left(\beta^{kL} - \Delta^{\beta_L} z_k\right)' \left(\beta^{kL} - \Delta^{\beta_L} z_k\right) + S)$

where N = No of keywords, $\nu = 10, S = 10I$

$V^{\alpha_H} \sim IW(\nu + N, \sum_{k=1}^{K} \left(\alpha^{kH} - \Delta^{\alpha_H} z_k\right)' \left(\alpha^{kH} - \Delta^{\alpha_H} z_k\right) + S)$

where N = No of keywords, $\nu = 10, S = 10I$

$V^{\alpha_L} \sim IW(\nu + N, \sum_{k=1}^{K} \left(\alpha^{kL} - \Delta^{\alpha_L} z_k\right)' \left(\alpha^{kL} - \Delta^{\alpha_L} z_k\right) + S)$

where N = No of keywords, $\nu = 10, S = 10I$

Step VI: Draw $\Delta^{\beta_H}\ \Delta^{\beta_L}\ \Delta^{\alpha_H}\ \Delta^{\alpha_L}$

Then

$\Delta^{\beta_H} \sim N(\widetilde{\Delta^{\beta_H}}, q_{\beta_H})$ where $q_{\beta_H} = [(z_k' z_k)^{-1} + A_0]^{-1}$ & $\widetilde{\Delta^{\beta_H}} = q_{\beta_H}\left[z_k' \beta^{kH} + A_0 \overline{\Delta^{\beta_H}}\right]$

$\overline{\Delta^{\beta_H}} = 0, A_0 = .01I$

$\Delta^{\beta_L} \sim N(\widetilde{\Delta^{\beta_L}}, q_{\beta_L})$ where $q_{\beta_L} = [(z_k' z_k)^{-1} + A_0]^{-1}$ & $\widetilde{\Delta^{\beta_L}} = q_{\beta_L}\left[z_k' \beta^{kL} + A_0 \overline{\Delta^{\beta_L}}\right]$

$\overline{\Delta^{\beta_L}} = 0, A_0 = .01I$

$\Delta^{\alpha_H} \sim N(\widetilde{\Delta^{\alpha_H}}, q_{\alpha_H})$ where $q_{\alpha_H} = [(z_k' z_k)^{-1} + A_0]^{-1}$ & $\widetilde{\Delta^{\alpha_H}} = q_{\alpha_H}\left[z_k' \alpha^{kH} + A_0 \overline{\Delta^{\alpha_H}}\right]$

$\overline{\Delta^{\alpha_L}} = 0, A_0 = .01I$

$\Delta^{\alpha_L} \sim N(\widetilde{\Delta^{\alpha_L}}, q_{\alpha_L})$ where $q_{\alpha_L} = [(z_k' z_k)^{-1} + A_0]^{-1}$ & $\widetilde{\Delta^{\alpha_L}} = q_{\alpha_L}\left[z_k' \alpha^{kL} + A_0 \overline{\Delta^{\alpha_L}}\right]$

$\overline{\Delta^{\alpha_L}} = 0, A_0 = .01$