Extended Abstract Prepared for Submission to WISE 2009

Search Engine Advertising: Empirical Analysis of Advertisers' Bids & Performance

Ashish Agarwal
McCombs School of Business
University of Texas, Austin
ashish.agarwal@mccombs.utexas.edu

Tridas Mukhopadhyay
Tepper School of Business
Carnegie Mellon University
tridas@cmu.edu

1. Introduction

Internet advertising spend is growing faster than any other form of advertising and is expected to surge from $16.4 billion in 2006 to $36.5 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 specific 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 ad relevance 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 are using similar auction format in other forms of online advertising such as contextual advertising.

While the phenomenon is growing rapidly, there is very little understanding of the drivers of performance in this marketplace. Search advertising differs from traditional advertising as the advertiser's can measure the performance of their advertising efforts more accurately. As a result it is important to understand how the ads perform in this medium and how the advertiser strategies are influenced by performance and other characteristics. For example, while it is known that the click performance decays with the rank of the advertiser it is not clear how it varies across the users with different search intentions, the popularity of the search terms or the relative brand popularity of the advertiser. Advertisers have to consider these differences in performance in order to make their choice of keywords, target positions, budget and the corresponding bids. Advertisers tend to differ in their budgets, advertising objectives as well as capabilities which influence their bidding strategies. So it is useful for the advertisers to know whether the choices made are appropriate to get the desired ROI in sponsored search. From a search engine's perspective, it is important to know whether the auction design is adequate to address the heterogeneity in consumer choices and advertiser decisions. Much of the work in search auctions has been theoretical in nature with the objective of evaluating the advertising bidding strategies, the profitability of search engine and the efficiency of the auction mechanism. Previous empirical work on sponsored search performance (Ghose & Yang, 2009; Rutz & Bucklin, 2007; Agarwal et al 2008) has been limited in scope as it has considered only single advertisers with a limited set of keywords.

In this paper we seek to understand how the user search characteristics, the advertiser characteristics as well as the auction outcome influence the advertiser bids. We also seek to determine how these choices along with the search characteristics and advertiser characteristics influence the performance in paid search auctions. To do this, we empirically analyze a unique panel dataset from a major search engine, which catalogs bids, rank, and click performance data for several hundred thousand keywords related to five keyword categories and sponsored by several thousand advertisers.

Our initial findings suggest that advertisers bids are determined by their past performance, budget, expenditure as well as experience. Advertisers bid differently for different types of keywords. Even though the overall performance is lower for popular keywords they tend to bid higher for such keywords. Additionally, advertisers in sponsored search are primarily interested in transactional benefits. As a consequence, they tend to bid higher on keywords where the consumer maybe more certain about the product. The click performance does depend on the advertiser and search charactertistics. We find that advertisers' with higher ad budgets may not necessarily have the highest performance.

2. Related Literature

The areas of work most closely related to this paper are consumers' online search behavior with a special emphasis on message order effects on consumer choice/action, research on advertising as a quality signal as well as recent work on search auctions.

In the literature on consumer search behavior, prior studies have shown that the depth of consumer search on the internet is very low (Johnson et al. 2004, Brynjolfsson, Dick, and Smith (2006). For example, Johnson et al. (2004) found that consumers searched less than two stores during a search session. Due to cognitive costs associated with evaluating alternatives, consumers often focus on a smaller set of results (Montgomery et al. 2004). It is thus likely that ordering and position strongly influence the attention paid to a marketing message online. Feng et al (2007) find evidence that the number of clicks for an ad decreases exponentially with its rank, and attributes this to decay in user attention as one proceeds down a list. Search behavior is also dictated by the consumer's purchase intent. Consumer search can be goal directed or exploratory (Janiszewski 1998). Online consumers include both buying consumers and information seekers (Moe 2003, Moe & Fader, 2004; Montgomery, Li, Srinivasan, & 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 purchase intent and resulting search behavior.

Studies in traditional settings show that consumers associate higher advertising expenditure with higher quality (Kirmani and Wright 1989). In case of sponsored search, as the consumers engage in sequential search, they may associate higher rank ads with higher quality. This would suggest that more prominent advertisers for a product would aim to be at the top of the list and bid accordingly. An important consideration is the interplay of search characteristics with the quality perception. A consumer in the information seeking mode has a higher degree of quality uncertainty. In such a scenario, ranking should play a more important role. Consequently, the advertisers interested in capturing the user attention should strive more for a higher rank for an uncertain user. Advertising spend has been linked to the firm performance. As a result, one can expect that large established firms will have higher ad budgets for sponsored ads and bid more aggressively. The click performance for such firms should be higher due to the consumer aware of the specific brand that the firm is trying to promote for a specific keyword. In practice, many large firms maintain several brands. Morgan & Rego (2009) have recently shown that brand portfolio can have significant impact on a firm's performance. A hypothesis for sponsored search would be that large firms would maintain larger portfolios of keywords to maintain their brand image and bid more aggressively. Consequently the click performance for such firms would be higher due to their general brand awareness among the consumers.

Edelman et al (2005) demonstrate that the generalized second price sponsored search auction, unlike the Vickrey-Clarke-Groves (VCG) mechanism, is not incentive compatible. Thus, advertisers will bid strategically in these auctions. Edelman and Ostrovsky (2007) examine data on paid search auctions and find evidence of strategic bidder behavior. Chen and He (2006) show that when advertisers are differentiated, they bid according to their product relevance. The corresponding paid placement by the search engine results in efficient search by the consumers and increases the social surplus. This would suggest that the more well known firms for a product would bid higher amounts to be in the top positions and signal. Animesh et al. (2006) empirically show that higher quality uncertainty can result in low quality firms appearing in the top position if the auction ranking mechanism is not considering the performance. This would suggest that less established firms would bid more aggressively for the keywords where users are in the information seeking mode. Ghose & Yang (2009) show that click performance is lower for longer keywords .These can be associated with more specific user queries indicating that the user is closer to the buying process. A question to evaluate is how this click performance varies with the type of advertiser and how it influences the advertisers' bids.

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

3. Model

We are interested in determining the relationship between advertiser bids and their characteristics as well as the search characteristics and past performance. We also want to evaluate the impact of these bids, as well as the advertiser and keyword characteristics on the click performance. We model these two outcomes as describe below:

Advertiser Bids

Advertisers rely on past performance and their budgets to determine the bids. We use two measures of past performance: Past average CTR and past Rank. We also control for the advertiser as well as keyword characteristics. An advertiser *a*'s bid for a keyword *k* at time *t* can be expressed as

$$Bid_{akt} = \theta_0 + \theta_1 AvgCTR_{t-1} + \theta_2 AvgRank_{t-1} + \theta_3 X_a + \theta_4 X_k + + \theta_5 X_{Int} + \alpha_{akt} \quad (1)$$

where $\boldsymbol{X_a}$ include advertiser characteristics, $\boldsymbol{X_k}$ include keyword variables and $\boldsymbol{X_{Int}}$ includes the interaction terms. $\theta = \{\theta_3, \theta_4, \theta_5\}$ are the corresponding parameter vectors.

Click Performance

Click performance is measured in terms of click through rate (CTR) which is the number of clicks conditional on ad impressions. We use a discrete choice model to capture the click performance. The user's latent utility for an advertiser *a* for keyword *k* at time *t* is expressed as follows

$$U_{akt} = \beta_0 + \beta_1 Rank_{akt} + \beta_3 Y_a + \beta_4 Y_k + \beta_5 Y_{Int} + \varepsilon_{akt} \quad (2)$$

where $\boldsymbol{Y}_a$ include advertiser characteristics, $\boldsymbol{Y}_k$ include keyword variables and $\boldsymbol{Y}_{Int}$ includes the interaction terms. $\beta = \{\beta_3, \beta_4, \beta_5\}$ are the corresponding parameter vectors. We assume that $\varepsilon_{k,t}$ are i.i.d with an extreme value distribution. Correspondingly, we use a logit model to represent the choice probability for an advertiser *a*, keyword *k* at time *t* as follows

$$\boldsymbol{P}_{akt} = \frac{1}{1+\exp(-\boldsymbol{U}_{akt})} \quad (3)$$

A similar model has been used by Misra et. Al (2006), Rutz and Bucklin (2006) & Ghose & Yang (2009) to capture clicks and conversions respectively as a function of ad attributes.
The advertiser decides on the bid for a keyword using its past performance. These bids influence the current rank of the ads and their current performance. As a consequence, equations (1) & (2) represent a system of equations with endogenous variables. In order to account for the correlation between the error terms for bid and clickthrough rate we use the following distribution

$$\begin{bmatrix} \alpha_{akt} \\ \varepsilon_{akt} \end{bmatrix} \sim \boldsymbol{N}(0,\Omega) \text{ where } \Omega = \begin{bmatrix} \Omega_{11} & \Omega_{12} \\ \Omega_{21} & \Omega_{22} \end{bmatrix} \quad (4)$$

4. Data & Results

Our dataset is provided by a major search engine. The data set consists of bids, impressions, and clicks for over a 123 day period from January 2008 to April 2008 for each rank for 607508 keywords belonging to 5 keyword categories representing common consumer durables. These keywords are associated with 16000 advertisers during the panel period. From this dataset, we select a random sample of 10023 keywords which are equally spread across the five product categories. The keywords are associated with 3216 advertisers. Bids are normalized for the entire dataset. . Summary statistics for our final sample are given in Table 1

Table 1: Summary Statistics

Variable	Mean	St. Dev.	Min	Max
Impressions	7.61	73.7	1	18868
Clicks	0.073	0.7	0	119
Rank	6.63	5.9	1	76
Bid	1519	3597	30	87670
Length	3.3	0.89	2	15
NoofAdvertisers	53	96	1	384
Advertiser budget	5.8e+07	3.5e+07	0	5.4 e+09
Advertiser Keyword Budget	1340	3e+06	0	6.1e+06
KeywordPortfolio	3319	7389	1	58070

In this draft we assume that there is no external shock which can influence the bids and the click performance simultaneously. This reduces the covariance matrix Ω to a diagonal matrix. This makes the system of equations recursive where each equation can be fully identified and estimated separately (Greene, 1999). Consequently, we estimate equation (1) using OLS and equation (2) using MLE. The estimation results are shown in Table 2.

Table 2: Parameter estimates for Bid & CTR

Bid Parameters	Estimate	Std	CTR Parameters	Estimate	Std.
(Intercept)	**6.85**	**0.00**	(Intercept)	**-3.57**	**0.01**
Past Avg CTR (pctr)	**-0.28**	**0.02**	rank	**-0.27**	**0.00**
Past Avg Rank	**-0.02**	**0.00**	Advertiser Budget (ad_B)	**-0.12**	**0.01**
Advertiser Budget (ad_B)	**0.35**	**0.00**	Advertiser Keyword Budget (adkw_B)	**1.63**	**0.03**
Advertiser Keyword Budget (adkw_B)	**4.88**	**0.02**	keyword Portfolio (kw_num)	**-0.39**	**0.01**
keyword Portfolio (kw_num)	**0.09**	**0.00**	size	**0.08**	**0.01**
size	**0.46**	**0.00**	NoofAdvertisers (ad_Num)	**-0.48**	**0.01**
NoofAdvertisers (ad_Num)	**0.51**	**0.00**	rank x size	**-0.01**	**0.00**
pctr x size	**-0.19**	**0.01**	rank x ad_Num	**0.03**	**0.00**
pctr x ad_Num	**0.50**	**0.03**	rank x ad_B	**0.00**	**0.00**
pctr x ad_B	**0.04**	**0.02**	rank x adkw_B	**-0.01**	**0.00**
pctr x adkw_B	**-0.39**	**0.06**	rank x kw_Num	**0.04**	**0.00**
pctr x kw_Num	**-0.28**	**0.02**	size x ad_B	**0.10**	**0.00**
size x ad_B	**0.03**	**0.00**	size x adkw_B	**1.02**	**0.02**
size x adkw_B	**3.34**	**0.02**	size x kw_Num	**-0.05**	**0.01**
size x kw_Num	**-0.02**	**0.00**	ad_Num x ad_B	**0.02**	**0.00**
ad_Num x ad_B	**-0.11**	**0.00**	ad_Num x adkw_B	**-0.70**	**0.01**
ad_Num x adkw_B	**-2.16**	**0.01**	ad_Num x kw_Num	**-0.14**	**0.02**
ad_Num x kw_Num	**0.03**	**0.00**			

5. Discussion

We find that advertisers bid lower for keywords with higher CTR. This accounts for the fact that the search engine uses a combination of past CTR and bid to rank the ads. As a result, firms with higher CTR can bid lower to maintain the same rank. Interestingly firms tend to lower bids in response to higher CTR at different rates. Large budget firms with smaller portfolio of keywords tend to bid higher as compared to firms with larger portfolio of keywords or large keyword specific budget.

Also as expected advertisers with higher budget and larger portfolio in general bid higher indicating that these firms are large and established brands. However, the realized performance is higher only for the firms with high keyword specific budget. Firms with large overall budget or large portfolio of keywords have lower click performance. This indicates that the creating awareness for a portfolio of brands through several keywords does not have an impact on the keyword specific performance.

Advertisers bid higher for longer keywords. This is consistent with the performance as the longer keywords have higher CTR. This contradicts the observation made by Ghose & Yang (2009) that longer keywords have lower CTR. Also firms with larger budget tend to bid higher for longer keywords and generate better click performance

We find that firms tend to bid higher for more popular search terms. This is expected as the demand for the ad slots for these keywords is high. However, this does not translate to higher performance for the advertisers. This is possible due to the fact that popular search terms are more generic. As a consequence, users are in an information seeking mode and may not click the ads at the same rate as they are generating impressions for these keywords. Rutz & Bucklin (2008) have pointed out that clicks are more important than impressions for generating brand awareness. Thus, advertisers seem to be overbidding for popular keywords. Advertisers differ in their valuation for popular keywords. Advertisers with large budgets and large keyword specific budget tend to bid lower for popular keywords.

Our results suggests that there advertisers adjust their bids in response to the search characteristics. The click performance varies with the search characteristics, the popularity of the search terms and the advertiser characteristics. We finally note that the advertiser strategies of managing large portfolio of keywords and bidding higher for popular search terms are misplaced as these do not result in better performance.

We plan to extend our study in several ways in time for any potential presentation at WISE. First, we plan to account for the correlation in the bid and click response. We will also incorporate keyword specific effects, advertiser specific effects as well as time effects and determine how these influence our results. We plan to investigate other parametric techniques to better address data sparseness and better model both the advertiser decision and the consumer response. We plan to have these results ready for potential presentation at WISE.

References

Agarwal, A., K. Hosanagar, M. Smith. 2008. Location, location and location: An analysis of profitability of position in online advertising markets. Working paper, Carnegie Mellon University, Pittsburgh.

Animesh, A., V. Ramachandran, S. Viswanathan. 2009. Quality uncertainty and the performance of online sponsored search markets: An empirical investigation. Inform. Systems Res., ePub ahead of print June 12, http://isr.journal.informs.org/cgi/content/abstract/isre.1080.0222v1.

Brynjolfsson, E., A.A. Dick, M.D. Smith. (2006). A Nearly Perfect Market? Differentiation vs. Price in Consumer Choice. Working Paper, Carnegie Mellon University, Pittsburgh, PA.

Chen, Y. & He, C. (2006). Paid Placement: Advertising and Search on the Internet, Working Paper.

Edelman, B., Ostrovsky, M. (2007). Strategic Bidder Behavior in Sponsored Search Auctions. *Decision Support Systems*, 43, 192-198.

Edelman, B., Ostrovsky, M. & Schwarz, M. (2007). Internet Advertising and the Generalized Second Price Auction: Selling Billions of Dollars Worth of Keywords, *American Economic Review*, forthcoming March 2007.

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.

Ganchev, K., Kulesza, A., Tan, J., Gabbard, R., Liu, Q., Kearns, M. (2007). Empirical Price Modeling for Sponsored Search. Working Paper

Ghose, A. and Yang, S. (2009). An Empirical Analysis of Search Engine Advertising: Sponsored Search in Electronic Markets *Forthcoming in Management Science.*

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

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

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

Montgomery, Alan L., Kartik Hosanagar, Ramayya Krishnan, and Karen B. Clay. (2004). Designing a Better Shopbot. *Management Science*, 50(2) 189–206.

Morgan, N, A. and Rego, L, L (2009). Brand Portfolio Strategy and Firm Performance. Journal of Marketing, 73(1), 59-74.
Rutz, O. & Bucklin, R (2006). A Model of individual Keyword Performance in Paid Search Advertising. Working Paper.
Rutz, O. & Bucklin, R (2007). From Generic to Branded: A Model of Spillover Dynamics in Paid Search Advertising. Working Paper.
Target Marketing (2006). Bid Fight. February, 29(2), 39-42