2008

An Empirical Examination of the Relation between Bids and Positions of Ads in Sponsored Search

Tanja Stepanchuk

School of Business and Economics, University of Frankfurt, stepanchuk@wiwi.uni-frankfurt.de

Follow this and additional works at: http://aisel.aisnet.org/bled2008

An Empirical Examination of the Relation between Bids and Positions of Ads in Sponsored Search

Tanja Stepanchuk

University of Frankfurt, School of Business and Economics,
Department of Marketing, Mertonstr. 17-25, 60554 Frankfurt, Germany
stepanchuk@wiwi.uni-frankfurt.de

Abstract

The rapid growth of e-business in the last years made sponsored search a multi-billion dollar industry, which will continue to grow in the upcoming years. Approximately 50% of the total online advertising spending today is used for sponsored search, where search engine providers use sponsored search auctions for pricing the clicks and ranking the ads based on the bids advertisers submit for a search term. The bid determines not only the price per click, but also the position of the ad in the sponsored search results, consequently costs, revenue and finally profitability of sponsored search. As the advertisers do normally not know the relationship between bids and positions of the ads in the sponsored search results and can thus not calculate the optimal bid, it is particularly challenging for advertisers to know: which response model allows for the robust prediction of a position of the ad by a certain bid and is easy to use. Using real Yahoo! Search Marketing data for bids and resulting positions from diverse e-business sectors we conduct an empirical examination of the relation between bids and positions of the ads in sponsored search results by calibrating different response models. Our findings reveal that the semi-logarithmic model i) is the most robust function for predicting a position of the ad, and ii) provides clear assistance for advertisers in terms of decision making about the bid for a search term, which is necessary to gain a certain position of the ad in the sponsored search results.

Keywords: E-Commerce, Sponsored Search

Introduction

A continuously growing number of Internet users make use of search engines to find and buy products online. In August 2007, 750 million users, about 95% of the worldwide Internet audience, generated in total 61 billion searches (Burns, 2007). The total U.S. online e-commerce sales in the B2C sector in 2006 amounted to $170.8 billion (Lipsman, 2007). In the B2B sector more than 50% of the purchases happen online, where 42.6% of them occur via sponsored search or search engine marketing (Hotchkiss, 2007). Nowadays sponsored search is a one of the most effective marketing instruments. In 2007 sponsored search accounted for 43.4% of total U.S. online ad spending and is expected to increase from $11.76 billion up to $26.79 billion in 2011 (Kedrosky, 2008). Why do advertisers spend so much money on sponsored search?

In contrast to traditional forms of advertising, e.g. TV advertising, sponsored search allows advertisers to make their advertisement more targeted. The ads are displayed in the sponsored

search results depending on the term the user is searching for. Hence, the advertisers get highly qualified visitors on their Web sites, which are already interested in buying the product or a service or at least in getting information about. Sponsored search works as follows: A consumer types a search term, e.g. "Hotels in Bled", into a search engine, here Yahoo! (Figure 1).
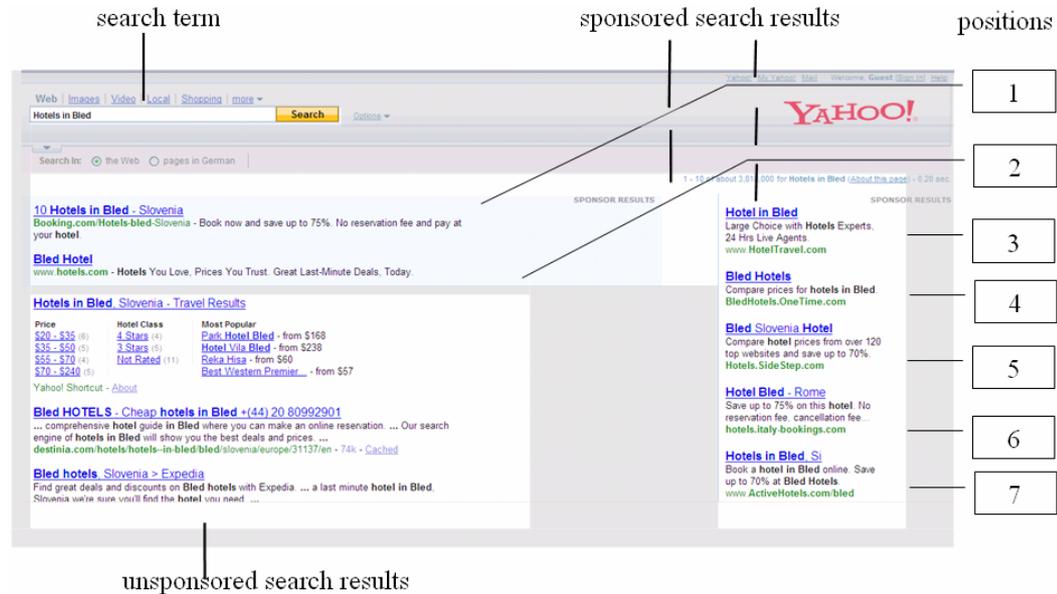

Figure 1: *Yahoo! search results*

He receives two types of results. The left side of the screen displays unsponsored search results sorted by relevance and on the right side and above it shows the ads that are displayed depending on the search term "Hotels in Bled". These ads are also called sponsored search results. The two ads on the top indicate position 1 to 2 of the ads in the sponsored search results, whereas the ads on the right side of the screen represent position 3 to 7. By clicking on one of those ads, the consumer is directed to the advertiser's landing page, which provides further information about the search term and an opportunity to act, e.g. to book a hotel in Bled. At the same time, the advertiser pays the search engine provider, here Yahoo!, for the click. The cost per click is a result of a sponsored search auction, where i) a search engine provider sells his advertising space and ii) the advertiser submits a bid for a search term, here "Hotels in Bled", the so-called "maximum cost per click" that they are willing to pay for each click on the displayed ad in the sponsored search results. (Edelman, Ostrovsky, Schwarz, 2007).

Meanwhile the costs for sponsored search in some e-business sectors are very high. In Germany's travelling sector the average costs per click on position 1 are $2.00, in insurance sector the click on the ad displayed on position 1 costs about $6.00 on average (Figure 2), whereas in the U.S. a cost per click on position 1 for a term "online car insurance" is about $36.32 ("Google Adwords Keyword Tool").
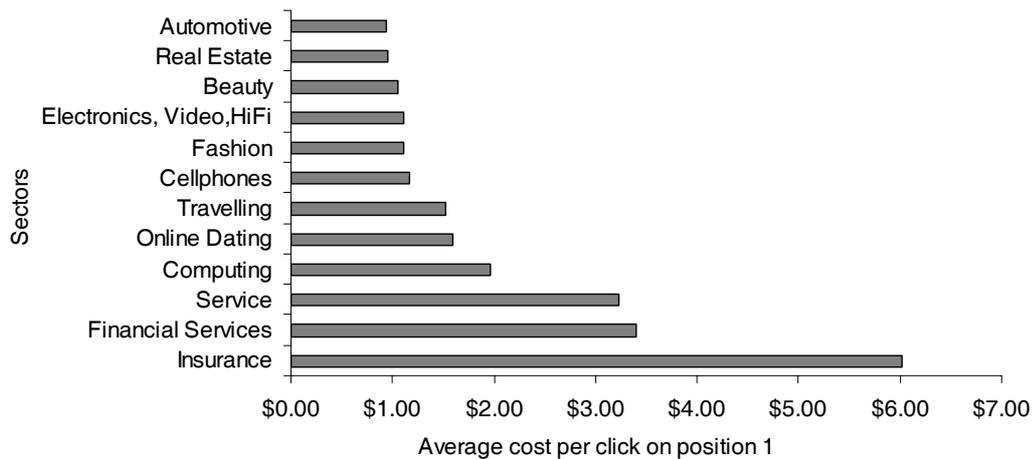
Figure 2: *Average costs per click on position 1 in different e-business sectors in Germany in December 2007*

Do advertisers know the relation between bids and the resulting positions of the ad, the resulting costs, revenues and profits in sponsored search? This paper answers those questions and provides insights into i) the related literature, ii) the measurement of the success in sponsored search and ii) the empirical examination of the relation between bids and positions of the ads in the sponsored search results by calibrating different response models.

Related Works

Although sponsored search is nowadays the most popular online advertising instrument, it has received very little attention as research topic in the academic literature. Most research comes from the economics and information systems literature. These papers focus on i) the design of the sponsored search auctions (Aggarwal, Goel, Motwani, 2006; Feng, Bhargava, Pennock, 2007; Lahaie, 2006; Varian, 2006) and ii) bidding behavior on sponsored search auctions (Animesh, Ramachandran, Viswanathan, 2006; Edelman, Ostrovsky, Schwarz, 2007; Kleinberg, 2005). Some researchers examine sponsored search as a revenue maximization problem (Borgs et al., 2006; Chakrabarty, Zhou, Lukose, 2007; Iyengar, Kumar, 2006; Kitts, Leblanc, 2004; Mehta, Saberi, Vazirani, Vazirani, 2005; Szymanski, Lee, 2006). Research on sponsored search in marketing is scarce. (Gerstmeier, Skiera, Stepanchuk, 2007; Rutz, Bucklin, 2007). Rutz, Bucklin (2007) as well as Misra, Pinker, Kauffman (2006) use a logit model for modeling conversion and clicks as a function of ad attributes. Finally, only two of the described papers mention which response models they apply to map the relationship between the bid and resulting position of the ad in the sponsored search results. Kitts, Leblanc (2004) use an exponential model, whereas Gerstmeier, Skiera, Stepanchuk (2007) use a semi-logarithmic response model. But in both cases, these authors neither calibrate their functional forms nor compare other response models. Therefore the unique contribution of this paper is the empirical examination of the relation between bids and positions of the ads in the sponsored search by calibrating different response models.

How to Measure Success in Sponsored Search

Figure 3 outlines how advertisers can measure their profitability of sponsored search. The submitted bids on sponsored search auctions do not only determine the average cost per click but simultaneously the position of the ads in the sponsored search results (KittsLeblanc 2004; Edelman, Ostrovsky et al. 2005; Varian 2006). The position of the ad then determines i) the click through rate (CTR), which measures the percentage of Internet users who clicked on the displayed ad as well as ii) the conversion rate (CR), which indicates the percentage of the Internet users who

200

clicked on the ad and finally bought the product. The multiplication of the CTR with the number of Internet users searching for a term, e.g. "Hotels in Bled" results in the number of clicks. The number of customers, acquired via sponsored search can then be calculated by multiplying the number of clicks and the CR. The multiplication of the number of clicks with the average cost per click defines the costs for sponsored search. Finally, the advertiser can calculate his profit in sponsored search as the profit per customer times the number of customers, acquired via sponsored search, minus costs for sponsored search.
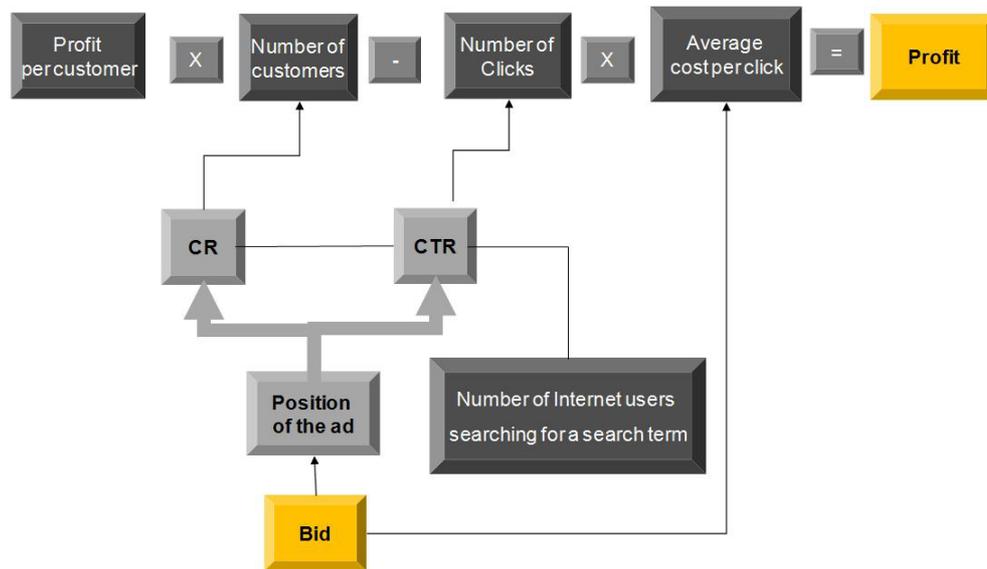

Figure 3: The measurement of profitability in sponsored search

The Google Eye tracking study revealed that top positions (1 to 4) in the sponsored search results are usually more attractive for the advertiser because they lead to more awareness, and consequently more clicks and thus very likely more customers (Hotchkiss, Alston, Edwards, 2005). However the costs per click at those positions are also much higher than at the lower positions. These effects require advertisers to trade-off between the number of acquired customers and the cost per acquired customer. To make it successfully the advertisers have to know how bids influence the positions of the ad in the sponsored search results in order to find the optimal bid for a search term, that determines resulting costs, revenues and finally their profit in sponsored search. Therefore it is vitally important for the advertisers to examine empirically which response model provides the most robust prediction for the position under a certain bid and for the bid to gain a certain position of the ad in sponsored search results.

Empirical Examination

In this section, we first describe the Yahoo! data for search terms from different e-business sectors we use for our examination and analyze the properties of the relationship between a bid and the resulting position of the ad. Second, based on the related works on sponsored search auctions and marketing modeling literature, we make up a set of the appropriate response models for the calibration. Finally, using the predictive validity and the goodness of fit criteria we compare the performance of the response models i) on the aggregate level, for each e-business sector and ii) on the individual level, for each search term separately.

Data

As Yahoo! Search Marketing offered the generalized first price open bid auction, where the submitted bid is visible, until 2007, we use its database (Overture, 2006) and obtained a dataset containing 326 widely used generic search terms for the following e-business sectors: "Financial Services & Insurance", "Travel & Wellness" and "Computing & Electronics". For each search term s the data contains the bid values in Euro $bid_s^1, bid_s^2, ..., bid_s^{p_s}$ which we define as prices that each bidder pays for a position $p_s = 1, ..., 20$. The bid values in our dataset decrease within the positions: the highest bid is assigned to bid_s^1 at the position on the top, $p_s = 1$, and the lowest bid is assigned to $bid_s^{p_s}$ at the lowest position, $p_s = 20$.

Some of search terms in our dataset contain only five instead of twenty data points, e.g. the search term "stock investment". Thus, we eliminate then those search terms from the consideration set. Table 1 reports the number of search terms, the number of observations (N), and the descriptive statistic for bids for each e-business sector.

Table 1: Descriptive statistic for bids from Yahoo! Search Marketing

e-business sectors	No. search terms	Example	N	Min	Max	Mean	Std. Dev.
Financial Services & Insurance	64	Online banking, car insurance online	947	0.07	5.98	0.93	0.84
Travel & Wellness	35	Hotel make-up, aloe Vera	487	.15	2.00	.45	.0.25
Computing & Electronics	90	Notebook, printer MP3-player, fax,	1021	0.05	4.00	0.47	0.50
Total	**189**		**2455**	**0.05**	**5.98**	**0.64**	**0.67**

Having a closer look at the data, we reveal that the bids are very heterogeneous even within a singular e-business sector. Figure 4 shows the substantial differences in bids for "Financial Services & Insurances" and even the "winning" bids on position 1 are quite heterogeneous.
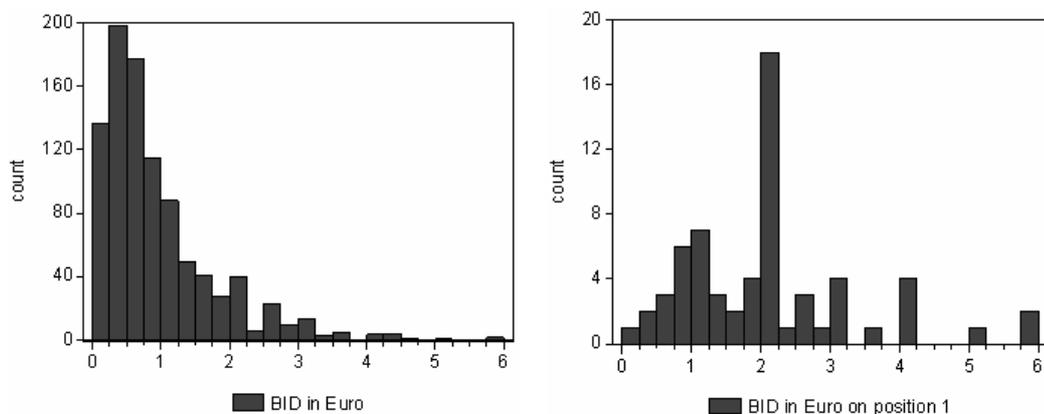

Figure 4: Bid values for e-business sector "Financial Services & Insurance"

To compensate the heterogeneity in bids, we scale bid values by the bid on position 1, getting a standardized bid that equals 1 on position 1:

$$\Delta\mathrm{bid}_s^{p_s} = \begin{cases} \dfrac{\mathrm{bid}_s^{p_s}}{\mathrm{bid}_s^1}, & 0 \le \mathrm{bid}_s^{p_s} < \mathrm{bid}_s^1, \\ 1, & \mathrm{bid}_s^{p_s} \ge \mathrm{bid}_s^1, \ p_s = \overline{1,20}. \end{cases} \tag{1}$$

In the following, we analyze the functional properties of the relation between standardized bid and the position of the ad, and compose the set of appropriate response models that we would like to calibrate.

Examined Response Models

The analysis of the data indicates that the examined relation has the following properties: is a nonlinear, generally convex function with increasing returns (Leeflang, Wittink, 2000a; Lilien, Kotler, Moorthy, 1992) and a "long tail" at the lower positions of the ad in the sponsored search results. This means that i) the position of the ad increases with the decrease in standardized bids and ii) the standardized bids provide almost no differences on lower positions. Therefore, looking thoroughly at the related works and the model building literature in marketing (Hanssens, Parsons, Schultz, 2001; Leeflang, Wittink, 2000b; Lilien, Kotler, Moorthy, 1992), in the following we set up the frequently used response models that meet all the properties, described above. We do not consider the linear, power series (quadratic, cubic) and S-shape functions like logistic, logarithmic reciprocal, Gompertz and ADBUDG models (Leeflang, Wittink, Wedel, Naert, 2000).

Whereas Kitts, Leblanc (2004) calibrate the position of the ad using absolute bid values, Gerstmeier, Skiera, Stepanchuk (2007) propose to model the relation between the standardized bid (1) and the position of the ad in the sponsored search results as a function of only one parameter β_s:

$$p_s(\Delta\mathrm{bid}_s^{p_s}) = \begin{cases} 1 - \dfrac{\ln(\Delta\mathrm{bid}_s^{p_s})}{\ln(\beta_s)}, & 0 \le \Delta\mathrm{bid}_s^{p_s} < 1, \\ 1, & \Delta\mathrm{bid}_s^{p_s} \ge 1, \end{cases} \tag{2}$$

where s is an index of search terms. Parameter β_s reflects the decay rate in bids across the positions of the ad and is calculated as a portion of the bid on position x to the bid on position (x+1). This model actually represents a classical semi-logarithmic model (Leeflang, Wittink, 2000a; Lilien, Kotler, Moorthy, 1992) and is denoted as *semi-log* in our examination[1].

Following the idea of Kitts, Leblanc (2004), we consider the estimation of the examined relation using an exponential model or *exp* afterwards (Leeflang, Wittink, Wedel, and Naert, 2000):

$$p_s(\Delta\mathrm{bid}_s^{p_s}) = \begin{cases} \alpha_s \cdot e^{\beta_s \cdot \Delta\mathrm{bid}_s^{p_s}}, & 0 \le \Delta\mathrm{bid}_s^{p_s} < 1, \\ \alpha_s \cdot e^{\beta_s}, & \Delta\mathrm{bid}_s^{p_s} \ge 1, \end{cases} \tag{3}$$

where α_s and β_s are exponential shape parameters.

---

[1] Find the details of estimation in the Appendix.

Based on the model building literature in marketing, the next model that meets the properties of the examined relationship is the multiplicative model (*power*):

$$p_s(\Delta bid_s^{p_s}) = \begin{cases} \alpha_s \cdot (\Delta bid_s^{p_s})^{\beta_s}, & 0 \leq \Delta bid_s^{p_s} < 1, \\ \alpha_s, & \Delta bid_s^{p_s} \geq 1, \end{cases} \tag{4}$$

where α_s and β_s are shape parameters of the multiplicative model.

Another response model that meets requirements is the reciprocal or *inverse* model in our analyses:

$$p_s(\Delta bid_s^{p_s}) = \begin{cases} \alpha_s + \dfrac{\beta_s}{\Delta bid_s^{p_s}}, & 0 \leq \Delta bid_s^{p_s} < 1, \\ \alpha_s + \beta_s, & \Delta bid_s^{p_s} \geq 1, \end{cases} \tag{5}$$

The constraint is that the sum of both parameters has to be equal 1: $\alpha_s + \beta_s = 1$.

Table A.1 in the Appendix provides the notation and characteristics of the examined response models.

Results and Implications

In order to analyze which of the introduced response models mirrors the relationship between standardized bids and positions of the ads in the sponsored search results best, we calibrate each response model in two steps: i) on the aggregate level, for every e-business sector, and ii) on the individual level, for each randomly chosen search term.

In a first step on the aggregate level, we i) create a holdout sample of 30% of search terms for each e-business sector, and ii) estimate linear forms of the specified response models on the predictive sample of 70% of search terms minimizing the mean squared error (MSE) using the Newton search method with a forward stepwise selection (Himmelblau, 1972).

Afterwards, we iii) compare the models' performance on the aggregate level using a) the predictive validity, computing the deviation of the MSE of the holdout sample from the MSE of the prediction sample $MSE\,dev.$ for each response model and e-business sector, and b) the goodness of fit criteria: adjusted R squared (R_a^2), MSE, Akaike Information Criterion (AIC) (Akaike, 1974), and Bayesian Information Criterion (BIC) (Schwarz, 1978).

Thus, in a first step of calibration, we specify 20 search terms with 268 observations for the e-business sector "Financial Services & Insurance" as a holdout sample and, correspondingly, 44 search terms with 679 observations as prediction sample. The results of calibration show that for "Financial Services & Insurance", the semi-logarithmic model (2) provides the lowest deviation of 0.10% from the MSE of the predictive sample (see Table 2). The MSE of the exponential response model (3) of the holdout sample deviates from the predictive MSE only on -2.89%. The largest deviation of -8.43% from the predictive MSE gives the multiplicative model (4). Additionally, we consider three main measures of goodness of fit to diagnose the performance of the response models: i) the lowest MSE, ii) the lowest values of AIC, and iii) the lowest value of BIC. Thus we reveal that for the e-business sector "Financial Services & Insurance" the semi-logarithmic response model (2) has the lowest values of MSE, AIC and BIC on the prediction and on the holdout samples.

Table 2: Results for "Financial Services & Insurance"

Financial Services & Insurance							
Response model	R_a^2	α_s	β_s	MSE	AIC	BIC	MSE dev. in %
Prediction sample: 44 search terms, N=679							
Semi-log	0.58		1.18	17.65	1526.39	1530.05	
Exp	0.57	20.34	-3.01	18.73	1544.05	1551.38	
Power	0.48	1.00	-1.02	43.92	1795.27	1798.93	
inverse	0.51	-0.25	1.25	42.15	1783.12	1786.78	
Holdout sample: 20 search terms, N=268							
Semi-log	0.60		1.18	**17.63**	**602.87**	**605.73**	**0.10**
Exp	0.53	20.34	-3.01	19.27	613.36	619.07	-2.89
Power	0.42	1.00	-1.02	47.63	718.53	721.39	-8.43
inverse	0.44	-0.25	1.25	43.73	708.61	711.46	-3.75

Table 3 summarizes the results of calibration of the relationship between the standardized bid and the position of the ad for "Travel & Wellness" which contains 10 search terms with 128 observations as a holdout sample and 25 search terms with 359 observations for the estimation of the response models.

Table 3: Results for "Travel & Wellness"

Travel & Wellness							
Response model	R_a^2	α_s	β_s	MSE	AIC	BIC	MSE dev. in %
Prediction sample: 25 search terms, N=359							
Semi-log	0.61		1.14	26.91	872.32	875.43	
Exp	0.58	23.25	-3.15	31.78	899.27	905.49	
Power	0.58	1.00	-1.37	46.13	957.36	963.59	
Inverse	0.59	-1.62	2.62	39.86	934.60	940.82	
Holdout sample: 10 search terms, N=128							
Semi-log	0.26		1.14	**6.61**	**233.02**	**235.24**	75.42
Exp	0.28	23.25	-3.15	11.27	263.66	268.09	64.53
Power	0.38	1.00	-1.37	28.12	314.48	318.91	**39.04**
Inverse	0.38	-1.62	2.62	18.73	291.89	296.31	53.01

In this case, on the holdout sample, the multiplicative response model (4) provides the lowest deviations of 39.04% from the MSE of the prediction sample. But the semi-logarithmic model has a better performance on the holdout sample: its MSE here is better on 75.42% than on the prediction sample. Analyzing the goodness of fit measures, the lowest values of MSE, AIC and BIC at both samples confirm the superior performance of the semi-logarithmic response model (2).

Finally, Table 4 reports the results of calibration of the response models for the e-business sector "Computing & Electronics". We reveal that the MSE values of all the response models are worse on the holdout than on the prediction sample. According to the goodness of fit criteria, we find that for the e-business sector "Computing & Electronics", the semi-logarithmic model (2) performs better on the prediction sample, but the exponential model performs better on the holdout sample.

Table 4: *Results for "Computing & Electronics"*

Computing & Electronics							
Response model	R_a^2	α_s	β_s	MSE	AIC	BIC	MSE dev. In %
Prediction sample: 63 search terms, N=508							
Semi-log	0.45		1.29	8.42	997.37	1000.80	
Exp	0.44	13.05	-2.57	8.62	1003.55	1010.40	
Power	0.42	1.00	-0.89	15.11	1129.81	1136.67	
Inverse	0.44	0.25	0.75	16.33	1147.33	1154.19	
Holdout sample: 27 search terms, N=503							
Semi-log	0.65		1.29	**21.96**	**1177.86**	**1181.26**	-160.7
Exp	0.57	13.05	-2.57	21.70	1175.25	1178.65	-151.8
Power	0.44	1.00	-0.89	37.20	1294.00	1300.81	-146.2
Inverse	0.49	0.25	0.75	38.35	1300.63	1307.44	-134.8

In a second step of calibration on the individual level, i) we calibrate the response models minimizing the MSE, but now for each of the randomly chosen 28 search terms separately, ii) we compare the models' performance using the goodness of fit criteria, and finally iii) we conduct a t-test to examine whether the differences in the models' MSE on the individual level are significant. Thus, we calibrate the response models for each of 28 search terms, randomly chosen from the e-business sector "Financial Services & Insurance", separately. We find that for 42.86% of search terms, the semi-logarithmic response model (2) provides the minimum MSE, the power (4) and the inverse (5) response models perform for 21.43% of search terms better, and the exponential function shows the lowest performance on the individual level (Figure 5).

According to the criteria of minimum BIC and minimum AIC, the semi-logarithmic response function achieves better performance than the exponential model, but the t-test shows that the differences in the MSE values of both models on the individual level are insignificant (t =-0.986). Whereas the differences in MSE values of i) the semi-logarithmic and power model (t= -3.056), and ii) the semi-logarithmic and inverse model (t= -2.286) are significant on 0.05- level.

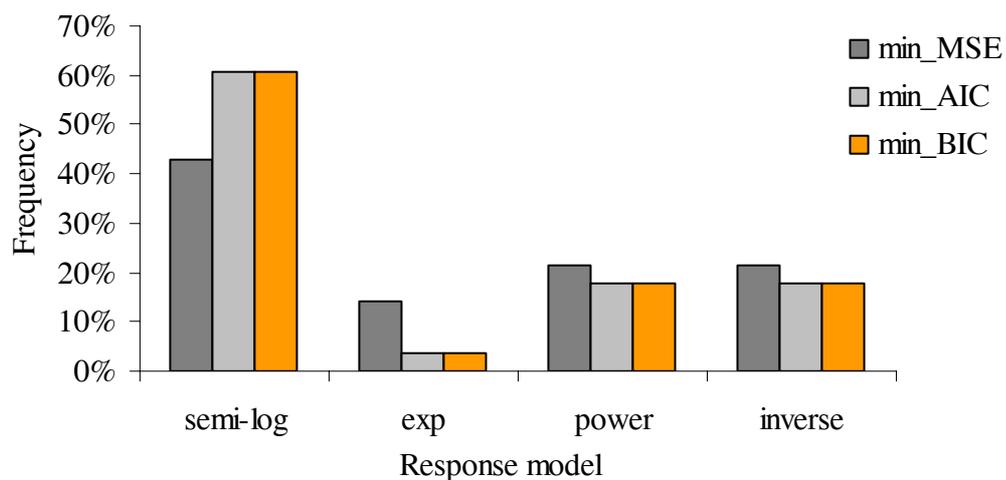**Figure 5**: *Comparison of the response models' performance on the individual level*

Summarizing our results, we reveal that the semi-logarithmic model (2) better reflects the relation between the standardized bids and positions of the ads in the sponsored search results providing the best goodness of fit and robustness in most cases. Moreover this response model can be intuitively interpreted. The inversion of this model results in the compound price function (Gerstmeier, Skiera, Stepanchuk, 2007):

$$\text{Price}\,(p_s) = \begin{cases} \dfrac{\text{bid}_s^1}{\beta_s^{(p_s-1)}}, & p_s > 1 \Leftrightarrow 0 \le \text{bid}_s^{p_s} < \text{bid}_s^1 \\[2mm] \text{bid}_s^1, & p_s = 1 \Leftrightarrow \text{bid}_s^{p_s} \ge \text{bid}_s^1, \end{cases} \tag{6}$$

which allows the advertiser i) to approximate the vector of bid values for a search term across the positions and ii) to predict thus his profitability in sponsored search.

Below, we demonstrate the importance of the price function for practice. For example, 5000 users searching for a search term "Hotels in Bled" with an average profit per customer of 100€. Supposed that the average price per click on position 1 for "Hotel in Bled" is $1€$, the estimated decay rate in prices within positions β_s is 1.29, using (6) the advertiser can easily calculate the cost per click that he would pay on position 3: $0.60€{=}1€/1.29^2$. Bidding on position 1 generates 400 clicks (CTR of 8%), where only 1.5% of the clicks convert into 6 new customers. Then the advertiser's costs for sponsored search are $400€{=}1€\cdot400$. Finally, according to Figure 3, the advertiser's profit of bidding on position 1 is $200€{=}100€\cdot6 - 400€$. In contrast, bidding on position 3 generates only 292 clicks (CTR of 5.84%), consequently only 4 customers are acquired via sponsored search. Correspondingly, the advertiser's costs on position 3 are lower than on position 1 (175.47€), but the advertiser's profit on position 3 is higher, and equals $224.53€{=}100€\cdot4 - 175.47€$. Thus, we can recommend advertisers to bid on top positions in sponsored search only if: i) the profit per customer is high; ii) the average price per click on position 1 for a search term is rather low, and iii) the number of clicks on the ad strongly decreases on lower positions.

Conclusions

In this paper we have outlined the functionality and measurement of the sponsored search. We have exposed the bid as a decision variable that determines not only the price per click, but the position of the ad in the sponsored search results. The position of the ad influences the number of clicks and the number of customers, acquired via sponsored search. Thus the bid determines costs, revenue and finally profitability. This paper contributes thereby to a major problem in sponsored search: How to predict a position achieved under a certain bid and a bid to gain a certain position of the ad in the sponsored search results? To find the response model that allows for the most precise prediction we have conducted an empirical examination i) using the real Yahoo! Search Marketing bid and the resulting position data for the e-business sectors "Financial Service & Insurance", "Travel & Wellness" and "Computing & Electronics", and ii) calibrating the different response models on aggregate and individual levels.

Our findings reveal on the aggregate level viz. for every e-business sector that the widely used exponential response model works well but does not provide the best performance in the majority of the cases. The semi-logarithmic model with only one shape parameter shows the best performance. The calibration on the individual level viz. separately for each search terms shows that a) the exponential model performs on the aggregate level better than on the individual level, b) the semi-logarithmic response function shows again the best performance, but c) there are no significant differences in the mean squared errors of exponential and semi-logarithmic model on the individual level.

Since the semi-logarithmic response model i) shows its superior performance in terms of goodness of fit and robustness, ii) is easy to use and iii) clearly assists advertisers in terms of decision making about the bid for a search term, we would recommend using it to predict the position of the ad in the sponsored search results. Additionally we would highly recommend advertisers:

- to choose search terms and texts of your ads carefully, be more specific and creative than your competitors;
- to consider your success as the profit after costs for sponsored search;
- to compute optimal bids for each search term separately, trading off between the number of acquired customers and the costs per customer, acquired via sponsored search;
- and to monitor the performance of your campaign constantly based on the daily data for each search term.

Appendix

Gerstmeier, Skiera, Stepanchuk (2007) propose the following response model:

$$p_s\left(\Delta bid_s^{p_s}\right) = \begin{cases} 1 - \dfrac{\ln\left(\Delta bid_s^{p_s}\right)}{\ln\left(\beta_s\right)}, & 0 \le \Delta bid_s^{p_s} < 1, \\ 1, & \Delta bid_s^{p_s} \ge 1, \end{cases} \tag{A.1}$$

which represents a classical semi-logarithmic model (Leeflang, Wittink, 2000a):

$$\begin{aligned} & p_s\left(\Delta bid_s^{p_s}\right) = 1 + \beta_{s_1} \cdot \ln\left(\Delta bid_s^{p_s}\right) + \varepsilon_s, \\ & \text{s.t. } \beta_{s_1} < 0, \end{aligned} \tag{A.2}$$

where the shape parameter β_{s_1} equals $-1/\ln(\beta_s)$. Consequently the decay rate in bids across the positions of the ad β_s can be computed as $1/e^{1/\beta_{s_1}}$.

Kitts, Leblanc (2004) use an exponential model:

$$p_s\left(\Delta bid_s^{p_s}\right) = \begin{cases} \alpha_s \cdot e^{\beta_s \cdot \Delta bid_s^{p_s}}, & 0 \le \Delta bid_s^{p_s} < 1, \\ \alpha_s \cdot e^{\beta_s}, & \Delta bid_s^{p_s} \ge 1, \end{cases} \tag{A.3}$$

where α_s and β_s are exponential shape parameters. Taking logarithms of both parts of the expression we estimate this model as follows:

$$\begin{aligned} & \ln\left(p_s\right) = \alpha_{s_1} + \beta_s \cdot \Delta bid_s^{p_s} + \varepsilon_s, \\ & \text{s.t. } \alpha_{s_1} + \beta_s = 0, \ \beta_s < 0, \end{aligned} \tag{A.4}$$

where parameter $\alpha_s = e^{\alpha_{s_1}}$. Here parameter β_s has to be negative to capture the effect that the position of the ad increases with a decrease in standardized bids.

Multiplicative or power model has the following form

$$p_s\left(\Delta bid_s^{p_s}\right) = \begin{cases} \alpha_s \cdot \left(\Delta bid_s^{p_s}\right)^{\beta_s}, & 0 \le \Delta bid_s^{p_s} < 1, \\ \alpha_s, & \Delta bid_s^{p_s} \ge 1, \end{cases} \tag{A.5}$$

where α_s and β_s are shape parameters of the multiplicative model. The increase in the position of the ad with the decrease in standardized bids guaranteed when parameter β_s is negative (Leeflang, Wittink, Wedel, Naert, 2000). Taking logarithms, we estimate it as follows:

$$\begin{aligned} \ln(p_s) &= \alpha_{s_1} + \beta_s \cdot \ln\left(\Delta bid_s^{p_s}\right) + \varepsilon_s, \\ \text{s.t. } &-1 < \beta_s < 0 \text{ or } \beta_s < -1, \alpha_{s_1} = 0. \end{aligned} \tag{A.6}$$

Table A.1 provides the notation and characteristics of the examined response models:

Table A.1: *Examined response models (RM) for calibrating the relationship between the standardized bid and the position of the ad*

Notation of RM	Form for position $p_s(\Delta bid_s^{p_s})$	Shape: convex	Elasticity	Marginal response
semi-log	$1 - \dfrac{\ln(\Delta bid_s^{p_s})}{\ln(\beta_s)}$	$\beta_s > 1$	$\dfrac{\beta_{s_1}}{1 + \beta_{s_1} \cdot \ln(\Delta bid_s^{p_s})}$ $\beta_{s_1} = -1/\ln(\beta_s)$	$\dfrac{1}{\beta_{s_1}}$, when $\Delta bid_s^{p_s} = e^{-1/\beta_{s_1}}$
Exp	$\alpha_s \cdot e^{\beta_s \cdot \Delta bid_s^{p_s}}$ $\alpha_s \cdot e^{\beta_s} = 1$	$\beta_s < 0$	$\beta_s \cdot \Delta bid_s^{p_s}$	$\alpha_s \cdot \beta_s \cdot e^{\beta_s \cdot \Delta bid_s^{p_s}}$
Power	$\alpha_s \cdot (\Delta bid_s^{p_s})^{\beta_s}$ $\alpha_s = 1$	$-1 < \beta_s < 0$ $\beta_s < -1$	β_s	$\alpha_s \cdot \beta_s \cdot (\Delta bid_s^{p_s})^{\beta_s - 1}$
Inverse	$\alpha_s + \dfrac{\beta_s}{\Delta bid_s^{p_s}}$ $\alpha_s + \beta_s = 1$		$\dfrac{-\beta_s \cdot (\Delta bid_s^{p_s})^{-2}}{\left(\alpha_s \cdot \Delta bid_s^{p_s} + \beta_s\right)}$	$\dfrac{-\beta_s}{(\Delta bid_s^{p_s})^2}$

References

Aggarwal, G., Goel, A., Motwani, R., (2006): 7th ACM Conference on Electronic Commerce, "Truthful Auctions for Pricing Search Keywords", Ann Arbor, Michigan, USA.

Akaike, H., (1974): A New Look at the Statistical Model Identification, IEEE Transactions on Automation Control, Vol. AC, No. 19, pp. 716-723.

Animesh, A., Ramachandran, V., Viswanathan, S., (2006): "Quality Uncertainty and the Performance of Online Sponsored Search Markets: An Empirical Investigation", Working Paper, SSRN eLibrary.

Borgs, C., Chayes, J., Etesami, O., Immorlica, N., Jain, K., Mahdian, M., (2006): 2nd Workshop on Sponsored Search Auctions, "Bid Optimization in Online Advertisement Auctions", Ann Arbor, Michigan, USA.

Burns, E., (2007): Worldwide Internet: Now Serving 61 Billion Searches per Month, http://searchenginewatch.com/showPage.html?page=3627304.

Chakrabarty, D., Zhou, Y., Lukose, R., (2007): Third Workshop on Sponsored Search Auctions, "Budget Constrained Bidding in Keyword Auctions and Online Knapsack Problems", Banff, Canada.

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, Vol. 97, No. 1, pp. 242-259.

Feng, J., Bhargava, H.K., Pennock, D.M., (2007): Implementing Sponsored Search in Web Search Engines: Computational Evaluation of Alternative Mechanisms, INFORMS Journal on Computing, Vol. 19, No. 1, pp. 137-148.

Gerstmeier, E., Skiera, B., Stepanchuk, T., (2007): "Influence of Bidding Heuristics on Profitability in Paid Search", Working Paper, University of Frankfurt, Frankfurt.

Google Adwords Keyword Tool, (2008).

Hanssens, D.M., Parsons, L.J., Schultz, R.L., (2001): "Market Response Models: Econometric and Time Series Analysis", Kluwer Academic, Boston et al.

Himmelblau, D.M., (1972): "Applied Nonlinear Programming", McGraw-Hill, New York.

Hotchkiss, G., (2007): Business to Business Survey 2007, Enquiro Search Solutions, Inc.

Hotchkiss, G., Alston, S., Edwards, G., (2005): Google Eye Tracking Report.

Iyengar, G., Kumar, A., (2006): 2nd Workshop on Sponsored Search Auctions, "Optimal Keyword Auctions", Ann Arbor, Michigan, USA.

Kedrosky, P., (2008): Forecast for Accelerating Online Ad Spending in 2008, http://seekingalpha.com/article/58908-forecast-for-accelerating-online-ad-spending-in-2008.

Kitts, B., Leblanc, B., (2004): Optimal Bidding on Keyword Auctions, Electronic Markets, Vol. 14, No. 3, pp. 186-201.

Kleinberg, R.D., (2005): 16th ACM-SIAM Symposium on Discrete Algorithms, "A Multiple-Choice Secretary Algorithm with Apllications to Online Auctions." Vancouver, British Columbia, Canada.

Lahaie, S., (2006): 7th ACM Conference on Electronic Commerce, "An Analysis of Alternative Slot Auction Designs for Sponsored Search", Ann Arbor, Michigan, USA.

Leeflang, P.S.H., Wittink, D.R., (2000a): Building Models for Marketing Decisions: Past, Present and Future, International Journal of Research in Marketing, Vol. 17, No. 2-3, pp. 105-126.

Leeflang, P.S.H., Wittink, D.R., (2000b): Models for Marketing Decisions: Postscriptum, International Journal of Research in Marketing, Vol. 17, No. 2-3, pp. 237-253.

Leeflang, P.S.H., Wittink, D.R., Wedel, M., Naert, P.A., (2000): "Building Models for Marketing Decisions", Kluwer, Dordrecht.

Lilien, G.L., Kotler, P., Moorthy, K.S., (1992): "Marketing Models", Prentice Hall, Englewood Cliffs (New Jersey).

Lipsman, A., (2007): comScore Networks Reports Total Non-Travel E-Commerce Spending Reaches $102 Billion in 2006, http://www.comscore.com/press/release.asp?press=1166.

Mehta, A., Saberi, A., Vazirani, U.V., Vazirani, V.V., (2005): 48th Annual IEEE Symposium on Foundations of Computer Science, "Adwords and generalized on-line matching", Providence, RI.

Misra, P., Pinker, E., Kauffman, A.R., (2006): WISE, "An Empirical Study of Search Engine Advertising Effectiveness", Evanston, Illinois.

Overture, (2006): Ranking Check.

Rutz, O.J., Bucklin, R.E., (2007): "A Model of Individual Keyword Performance in Paid Search Advertising", Working Paper, SSRN-Library.

Schwarz, G., (1978): Estimating the Dimension of a Model, Annals of Statistic, Vol. 6, No. 2, pp. 461-464.

Szymanski, B.K., Lee, J.S., (2006): 2nd Workshop on Sponsored Search Auctions, "Impact of ROI on Bidding and Revenue in Sponsored Search Advertisement Auctions", Ann Arbor, Michigan, USA.

Varian, H.R., (2006): Position Auctions, International Journal of Industrial Organization, Vol. 25, No. 6, pp. 1163 -1178.