

Reserve Prices in Internet Advertising Auctions: A Field Experiment*

Michael Ostrovsky[†] Michael Schwarz[‡]

November 2016

Abstract

We present the results of a large field experiment on setting reserve prices in auctions for online advertisements, guided by the theory of optimal auction design suitably adapted to the sponsored search setting. Consistent with the theory, revenues increased substantially after the new reserve prices were introduced.

*We are especially indebted to the engineers at Yahoo!, too numerous to list here, without whom this project could never have been implemented. We are also grateful to David Reiley and Preston McAfee for many helpful discussions and to Ken Hendricks, Jon Levin, and many seminar and conference participants for helpful comments and suggestions.

[†]ostrovsky@stanford.edu.

[‡]schwarz.m@gmail.com.

1 Introduction

Auctions are used to sell a wide variety of objects, ranging from flowers, paintings, and used cars to electromagnetic spectrum and Internet advertisements. One of the most natural questions about the design of an auction is revenue maximization: How should an auction be designed to generate the highest expected payoff to the seller? This question was answered by Myerson (1981) and Riley and Samuelson (1981) for the setting with one object for sale and independently distributed private bidder values. For the case with symmetric bidders, the answer is particularly elegant: the optimal mechanism can be implemented by a second-price auction with an appropriately chosen reserve price.

This theoretical work has been extended in many directions: e.g., Cremer and McLean (1988) and McAfee, McMillan, and Reny (1989) construct optimal auctions in settings with correlated and common bidder values; Maskin and Riley (1984) derive optimal mechanisms in settings with risk-averse bidders; and Maskin and Riley (1989), Armstrong (2000), and Avery and Hendershott (2000) study optimal design in settings with multiple objects.

Economists have also obtained empirical estimates of (and bounds on) optimal reserve prices for a variety of auctions (McAfee and Vincent, 1992; Paarsch, 1997; Athey, Cramton, and Ingraham, 2002; McAfee, Quan, and Vincent, 2002; Bajari and Hortaçsu, 2003; Haile and Tamer, 2003; Tang, 2011). Notably, taken together, the results of these papers present a puzzle: these papers typically find that reserve prices actually observed in real-world auctions are substantially lower than the theoretically optimal ones. This raises the possibility that reserve prices are not a particularly important part of auction design, and sellers cannot use them to substantially raise revenues. Moreover, if participation in the auction is costly for bidders, increasing the reserve price may make the auction less attractive to them and fewer will bid, leading to lower revenues. Indeed, Bulow and Klemperer (1996) find that in symmetric single-object auctions, adding just one more bidder (and setting a zero reserve price) is always preferable to setting the optimal reserve price. So perhaps reserve prices are not important in practice?

In this paper, we address this question directly, by presenting the results of a large-scale field experiment on reserve prices in a particular setting: “sponsored search” auctions conducted by Yahoo! to sell advertisements. Reserve prices in the randomly selected “treatment” group were set based on the guidance provided by the theory of optimal auctions, while in the “control” group they were left at the old level of 10 cents per click. The revenues in the treatment group have increased substantially relative to the control group, showing that reserve prices in auctions can play an important role and that theory provides a useful guide for setting them. This increase is driven by keywords with relatively high search volumes, keywords in which the theoretically optimal reserve price is relatively high, and keywords with a relatively small number of bidders.

Two prior studies have analyzed the results of controlled experiments on setting reserve prices in auctions. Reiley (2006) reports the results of a field experiment on reserve prices in a first-price online auction for trading cards for a popular collectible card game. His findings confirm several predictions of auction theory, such as the reduction in the probability of a sale when reserve prices

are present and, more subtly, the increase in bids when they are present (which is a consequence of equilibrium behavior in first-price auctions). Brown and Morgan (2009) report the results of field experiments on auctions for collectible coins conducted on Yahoo! and eBay. The primary focus of the study is the competition between platforms and market tipping, but the authors also consider the effects of reserve prices. They find that positive reserve prices, set at the level of 70% of the purchase price of the coins from the dealer, lead to significantly higher revenues and lower numbers of bidders, relative to zero reserve prices.

Our paper makes several contributions relative to these studies. First, it analyzes a much larger and economically important setting, with hundreds of thousands of keywords in a multi-billion dollar marketplace. Consequently, many of the bidders in this setting spend considerable time and resources on optimizing their advertising campaigns. Second, the reserve prices in the experiment are guided by theory, based on the estimated distributions of bidder values. To the best of our knowledge, ours is the first paper describing a successful practical application of the seminal results of Myerson (1981) and Riley and Samuelson (1981).¹ Third, unlike the previous randomized experiments, the benchmark in our analysis is not a zero reserve price, but the existing reserve price set by the company after a long period of experimentation.² Finally, our paper emphasizes the fact that the potential impact of reserve prices is much higher in multi-unit auctions than in single-unit ones (in some cases, by orders of magnitude). This observation applies not just to sponsored search auctions, but to any multi-unit, simultaneous, or sequential auctions in which substitutable goods are sold to bidders who have limited demands, ranging from timber auctions to procurement auctions.

The paper is organized as follows. Section 2 provides an overview of the sponsored search setting. Section 3 extends theoretical results on optimal auction design (Myerson, 1981; Riley and Samuelson, 1981) to the current setting and discusses simulations of revenue impact of reserve prices. Section 4 describes the design of the experiment. Section 5 presents the experimental results. Section 6 concludes.

2 Sponsored Search Auctions

We start with a brief description of the sponsored search setting; for a detailed description, see Edelman, Ostrovsky, and Schwarz (2007, subsequently EOS). When an Internet user enters a search term (“query”) into a search engine, he gets back a page with results, containing both the links most relevant to the query and the sponsored links, i.e., paid advertisements. The ads are clearly distinguishable from the actual search results, and different searches yield different sponsored links: advertisers target their ads based on search keywords. For instance, if a travel agent buys the word “Hawaii,” then each time a user performs a search on this word, a link to the travel agent

¹Walsh et al. (2008) describe the results of a small live test of an automated reserve pricing system for a reseller of returned goods over a two-month period, after which the system was turned off.

²The initial reserve price set in 1998 was 1 cent per click. The reserve price was subsequently raised to 5 cents per click in 2001 and then to 10 cents in 2003.

will appear on the search results page. When a user clicks on the sponsored link, he is sent to the advertiser’s Web page. The advertiser then pays the search engine for sending the user to its Web page. Different positions of an ad on the search results page have different desirability for advertisers: an ad shown at the top of a page is more likely to be clicked than an ad shown at the bottom.

To allocate positions to advertisers, most search engines use variations of the “generalized second-price” (GSP) auction. In the simplest GSP auction, for a specific keyword, advertisers submit bids stating their maximum willingness to pay for a click. An advertiser’s bid remains active until he changes or disables it. When a user enters a keyword, she receives search results along with sponsored links, the latter shown in decreasing order of bids. In particular, the ad with the highest bid is displayed at the top, the ad with the next highest bid is displayed in the second position, and so on. If a user clicks on an ad in position i , that advertiser is charged by the search engine the amount equal to the next highest bid, i.e., the bid of the advertiser in position $(i + 1)$.

If a search engine offered only one ad position per result page, this mechanism would be equivalent to the standard second-price auction, coinciding with the Vickrey-Clarke-Groves (VCG) mechanism (Vickrey, 1961; Clarke, 1971; Groves, 1973). With multiple ad positions, GSP generalizes the second-price auction (hence the name). Here, each advertiser pays the next highest advertiser’s bid. Aggarwal, Goel, and Motvani (2006), EOS, and Varian (2007) show that with multiple positions, the GSP auction is no longer equivalent to the VCG auction. **In particular, unlike the VCG mechanism, GSP generally does not have an equilibrium in dominant strategies, and truth-telling is not an equilibrium of GSP. Nevertheless, GSP has a natural equilibrium, with advertisers in general bidding less than their true values, in which the payoffs of advertisers and the search engine are the same as under VCG for every realization of bidder values.**

3 Theory

The ideas of Myerson (1981) can be combined with the analysis of EOS and Varian (2007) to derive the optimal mechanism for the sponsored search setting and to show how it can be implemented with minimal changes to the existing GSP auction. Below is a sketch of the derivation, using the notation of EOS.³ Suppose in an auction for a particular keyword there are K bidders and N positions on the screen. The (expected) number of clicks per period received by the advertiser whose ad was placed in position i is α_i . The value per click to advertiser k is s_k . These values are private information of the advertisers, drawn from distribution $F_k(\cdot)$ on $[0, \bar{s}_k]$. Values are independently distributed, so the distribution over vectors of values s is $F(s) = F_1(s_1) \times \cdots \times F_K(s_K)$ over $S = [0, \bar{s}_1] \times \cdots \times [0, \bar{s}_K]$. Advertisers are risk-neutral, and advertiser k ’s payoff from being in position i is equal to $\alpha_i s_k$ minus his payments to the search engine. Without loss of generality, positions are labeled in descending order ($\alpha_1 > \alpha_2 > \dots$).

³Similar derivations are contained in Iyengar and Kumar (2006), Roughgarden and Sundararajan (2007), and Edelman and Schwarz (2010).

Now consider an incentive-compatible direct revelation mechanism. Let $t_k(s_k)$ be the expected payment of bidder k with value s_k , let $x_k(s_k)$ be the expected number of clicks received by bidder k with expected value s_k , and, slightly abusing notation, let $x_k(s)$ be the expected number of clicks received by bidder k when the vector of bidder values is s . Then, using the same arguments as in the case of single-object optimal auctions (see, e.g., Krishna, 2009, for an exposition), except that the probability of receiving the object in the single-object case is replaced by the expected number of clicks in our case, we have the following equality for the expected payment of each bidder:

$$t_k(s_k) = t_k(0) + x_k(s_k)s_k - \int_0^{s_k} x_k(u_k)du_k.$$

This, following the standard argument, in turn implies that the expected payoff of the search engine is equal to

$$\sum_{1 \leq k \leq K} t_k(0) + \int_S \left(\sum_{1 \leq k \leq K} \psi_k(s_k)x_k(s) \right) f(s)ds,$$

where $\psi_k(s_k) = s_k - \frac{1-F_k(s_k)}{f_k(s_k)}$ is the virtual valuation of advertiser k with value s_k .

We now make two additional assumptions. First, assume that the virtual valuation is an increasing function.⁴ Second, assume that bidders are symmetric, i.e., have identical distributions of values. Then the revenues of the search engine are maximized when $t_k(0) = 0$ for any k and when $\left(\sum_{1 \leq k \leq K} \psi_k(s_k)x_k(s) \right)$ is maximized pointwise, for every s , which happens when (i) only bidders with positive virtual valuations are allocated clicks and (b) among them, bidders with higher virtual valuations (and thus, by assumption, with higher actual valuations) are allocated as many clicks as possible. Since each advertiser can only have one position on the screen, this simply means that the bidder with the highest value receives the top position, the bidder with the second highest value receives the second position, and so on.

Now consider an indirect mechanism: the generalized second-price auction with reserve price r^* such that $\psi(r^*) = 0$. By the argument analogous to that in EOS, in the bidder-optimal envy-free equilibrium of this auction (or, equivalently, in the unique equilibrium of the corresponding generalized English auction with reserve price r^*), bidders with values less than r^* (i.e., bidders with negative virtual valuations) will receive no clicks; among the bidders with values greater than r^* (i.e., with positive virtual valuations), the ones with higher values will receive higher positions; and finally bidders with value zero receive (and make) payments of zero. Hence, the allocations and expected payoffs in this mechanism are the same as those in the optimal direct mechanism, and thus GSP with reserve price r^* is a revenue-maximizing mechanism.

⁴As described later in the paper, reserve prices in the experiment were computed under the assumption that bidders' values are distributed lognormally. Through simulations, it was determined that for lognormal distributions with parameter values relevant for the experiment, virtual valuations are increasing. A recent paper by Ewerhart (2013) establishes sufficient conditions under which virtual valuations for lognormal distributions are monotonically increasing.

3.1 The Impact of Reserve Prices on Revenues

Remarkably, the optimal reserve price depends neither on the number of bidders nor on the number of available positions. The impact of reserve prices on revenues, however, depends critically on these parameters. In fact, in single-object auctions, reserve prices only play an important role if the number of bidders is small. To give a simple example (Table 1), suppose bidder values are distributed uniformly on $[0, 1]$, with the corresponding optimal reserve price $r^* = 0.5$. Then with just two bidders, the effect of setting the optimal reserve price rather than no reserve price is substantial: it raises the expected revenues by 25%, from 0.33 to 0.42.⁵ With six bidders, however, the effect is trivial: moving from no reserve price to the optimal reserve price changes the expected revenues by less than one third of one percent. The intuition for this decline is straightforward: reserve price r only has a positive impact when one bidder’s realized value is above r and the other bidders’ values are all below r , and the probability of this event becomes small as the number of bidders increases.

Table 1: The impact of reserve prices, uniform distribution

Bidders	$r = 0$	$r = 0.5$	Impact
<i>Single-object second-price auction</i>			
$n = 2$	0.3333	0.4166	25%
$n = 6$	0.7143	0.7165	0.31%
<i>GSP with a “decay factor” of 0.7</i>			
$n = 2$	0.1	0.475	375%
$n = 6$	0.8123	1.1764	45%

Of course, the same effect holds for multi-unit auctions, like the sponsored search ones, if the number of slots is fixed but the number of bidders increases. However, reserve prices retain their power for much higher numbers of bidders, and are in general much more important. To see this, consider a generalized second-price auction with the decline factor of 0.7 (i.e., the top position expects to receive one click, the second position expects to receive 0.7 clicks, the third position expect to receive 0.49 clicks, and so on).⁶ With two bidders and no reserve price, the expected revenue of the auctioneer is only 0.1: in essence, both bidders get 0.7 clicks “for free,” and only compete for the remaining 0.3 clicks, thus generating the revenue of $0.3 \times 0.33 = 0.1$. Note, however, that it would be feasible for the search engine to “shut down” all positions below the top one, not allocating them to anyone, and so the revenue in the optimal auction has to be at least as high as in the optimal single-object one, i.e., 0.42. As we know from the theoretical analysis, the optimal auction does not in fact involve “shutting down” any positions: the auctioneer simply sets the reserve price equal to 0.5. The resulting expected revenue turns out to be 0.475, i.e., an

⁵The expected revenue without a reserve price is equal to $E[\min\{s_1, s_2\}] = 0.33$. The expected revenue with $r^* = 0.5$ is equal to $0.25 \times E[\min\{s_1, s_2\} | s_1 > 0.5, s_2 > 0.5] + 0.5 \times 0.5 = 0.42$.

⁶The average decline factor of 0.7 is typical in the sponsored search setting, and so we use it throughout our examples and simulations.

improvement of 375% relative to the case of no reserve price.

Even with six bidders, reserve prices remain very important: the optimal reserve price improves the revenues by 45%. To see why the difference relative to the single-object case is so dramatic, consider what would have happened if the decline factor in the sponsored search auction was equal to 1 rather than 0.7 (i.e., all positions received the same number of clicks) and there were as many available positions as bidders. Without a reserve price, there would be no competition for positions and the auctioneer’s revenue would be equal to zero. With the optimal reserve price $r^* = 0.5$, revenue would be equal to the number of bidders times 0.25—an infinite improvement. Of course, with the discount factor of 0.7, the positions are no longer perfect substitutes, and the importance of reserve prices is not as dramatic, but the intuition is essentially the same. Note that in many other settings, multiple substitutable objects are also auctioned off, either simultaneously or sequentially, and if bidders in these auctions have limited demands or each is restricted to one or only a small number of objects, then for the same reason, the analysis based on an individual single-object auction may severely understate the importance of reserve prices.

In order to estimate optimal reserve prices for the experiment, it was assumed that bidders’ values are drawn from lognormal distributions. Table 2 shows the impact of various levels of reserve prices on revenues in GSP under this assumption, with the parameters of the distribution chosen in such a way that its mean is equal to 0.5 and its standard deviation is also equal to 0.5. The corresponding optimal reserve price is equal to 0.37. These parameters were chosen to give an illustration of a representative keyword; for instance, as we describe below, the optimal reserve price of 37 cents corresponds to the 75th percentile of estimated optimal reserve prices for the analyzed sample and is close to the average estimated optimal reserve price. The table presents the expected revenues for four levels of reserve prices: 0, 0.10 (corresponding to the old reserve price at Yahoo!, 10 cents), 0.235 (corresponding to the midpoint between the old reserve price and the theoretically optimal reserve price), and 0.37 (the theoretically optimal reserve price). Similar to the example with the uniform distribution of values, the impact of optimal reserve prices on revenues in the GSP auction is substantial: with six bidders, setting the reserve price at zero instead of the optimal level results in the loss of 25% of revenues; and even with ten bidders, the loss is noticeable: 9%.

Table 2: The impact of reserve prices, lognormal distribution

Bidders	$r = 0$	$r = 0.10$	$r = 0.235$	$r = 0.37$
$n = 2$	0.08 (24%)	0.22 (63%)	0.32 (93%)	0.34 (100%)
$n = 6$	0.68 (75%)	0.78 (86%)	0.87 (96%)	0.91 (100%)
$n = 10$	1.24 (91%)	1.28 (94%)	1.33 (98%)	1.36 (100%)

Table 3 presents the results of analogous impact calculations for four actual keywords taken from our sample: a “median” keywords with the mean and the standard deviation of the estimated distribution of values (0.31, 0.26), the average number of bidders ($n = 6$), and the corresponding optimal reserve price ($r^* = 0.22$) close to the median ones for the sample; two keywords with similar numbers of bidders but a substantially higher or lower means of the distribution of values: (0.78, 0.11, $n = 6$, $r^* = 0.63$) and (0.22, 0.13, $n = 5$, $r^* = 0.15$); and a keyword with a much higher mean value and substantially more bidders: (2.18, 0.99, $n = 9$, $r^* = 1.49$).

Table 3: The impact of reserve prices, sample keywords

Keyword parameters (mean, std.dev., n)	r^*	$r = 0$	$r = 0.10$	$r = \frac{0.10+r^*}{2}$	$r = r^*$
(0.31, 0.26, 6)	0.22	0.446 (76%)	0.541 (92%)	0.578 (98%)	0.591 (100%)
(0.78, 0.11, 6)	0.63	1.446 (71%)	1.547 (76%)	1.812 (89%)	2.037 (100%)
(0.22, 0.13, 5)	0.15	0.272 (68%)	0.381 (95%)	0.396 (99%)	0.401 (100%)
(2.18, 0.99, 9)	1.49	5.500 (90%)	5.551 (91%)	5.909 (96%)	6.131 (100%)

While for most of these keywords the reserve price of 0.10 produces higher revenues than the reserve price of 0, it still falls far short of the optimal revenue. In contrast, moving to the midpoint between 0.10 and the optimal reserve price allows the search engine to capture most of the upside from optimal reserve prices, likely because in that region, the derivative of expected revenues with respect to reserve prices is small (of course, at the optimal price level itself, that derivative equals zero).⁷ As discussed below, this observation plays an important role in the implementation of reserve price levels in the field experiment.

4 Experiment

In practice, sponsored search auctions have a number of complicating features that make the model of Section 3 only a stylized representation of reality. Nevertheless, the model was viewed as a useful approximation and was used as the basis for the experiment. In this section, we outline the implementation of the experiment and discuss several of the complicating features. Broadly speaking, the implementation of the experiment involved two steps: estimating the distributions of bidder values and setting reserve prices.

⁷Note that in this setting, one should be careful about various convexity and concavity statements. In fact, with two or more bidders, the function mapping reserve prices to revenues is neither convex nor concave: its derivative equals zero both at the optimal reserve price and at the reserve price of zero.

4.1 Estimating the Distributions of Bidder Values

The company picked a set of criteria for choosing keywords suitable for the experiment (for instance, keywords that had very few bidders or very few searches were not included in the sample, because the distributions of bidder values for such keywords could not be reliably estimated). The resulting sample consisted of 461,648 keywords. It also picked a time interval of several weeks during which the data for estimation were collected.

For each keyword in the sample, the following moments were computed: the average number of advertisers bidding on this keyword, the average bid, and the average standard deviation of the bids, where the average was taken over all searches (i.e., every time the keyword was searched, the three statistics were computed, and then the average over all searches for the given keyword was taken). The bid of the highest bidder in every auction was excluded from the statistics, because the theory does not allow us to pin it down (just like in a single-object second-price auction, under GSP, every bid of the highest bidder above a certain value results in the same vector of payoffs).

Next, it was assumed that bidders' values were drawn from a lognormal distribution with a mean and a standard deviation to be estimated. The number of potential bidders needed to be estimated as well: during the period when data were collected, Yahoo!'s sponsored search auctions had a uniform reserve price of 10 cents, and so bidders with per-click values of less than 10 cents were not observed in the data.

The next step was to simulate the three moments (observed number of bidders, average bid in positions 2 and below, and the standard deviation of the bids in positions 2 and below) for various true values of the number of potential bidders and the mean and the standard deviation of the lognormal distribution of values. To do that, for each combination of true values of the variables of interest, several hundred draws of the vectors of bidder values were drawn. For each draw, equilibrium bids were computed, taking into account the 10 cents reserve price and assuming that the bidders were playing the unique perfect Bayesian equilibrium of the Generalized English Auction in EOS. The moments of interest were then computed, and averaged over all draws of vectors of bidder values.

For each keyword, the number of bidders and the parameters of the distribution of bidder values were then estimated by matching the observed moments to the simulated ones. Note that the number of bidders is irrelevant for setting the optimal reserve price, but it needs to be estimated in order to get an accurate estimate of the mean and the standard deviation of the distribution of values. Finally, for each keyword, the theoretically optimal reserve price was computed using the formula in Section 3.

Figure 1 shows the histogram of the distribution of estimated optimal reserve prices for the sample, and Table 4 lists several key percentiles. The median optimal reserve price is 20 cents, the 10th percentile is 9 cents, and the 90th percentile is 72 cents. Note that just like in the previous empirical studies of reserve prices in auctions, we find that for most of the sample (almost 90%), the estimated optimal reserve price exceeds the actual reserve price used in the auction (10 cents), and for much of the sample, the difference is substantial. Unlike the previous studies, however,

in the current paper we can directly measure the importance of this difference, by conducting a controlled experiment.

Figure 1: The distribution of estimated keyword-specific optimal reserve prices (histogram)

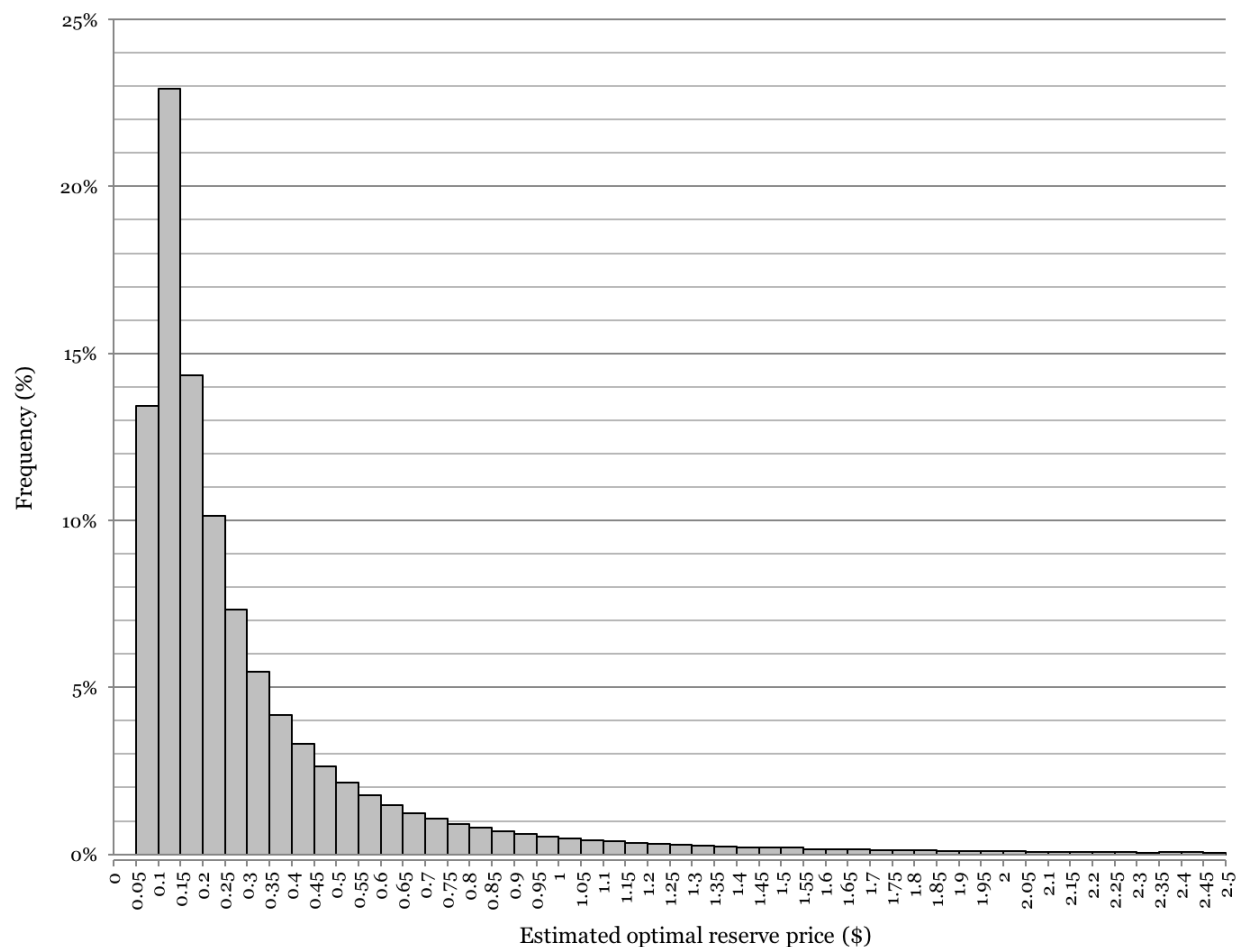


Table 4: The distribution of estimated keyword-specific optimal reserve prices (percentiles)

Percentile	10%	25%	Median	75%	90%
Estimated r^*	9¢	12¢	20¢	37¢	72¢

Several details of the estimation procedure deserve additional attention.

First, as a simplification, in the simulation procedure it was assumed that each ad’s probability of being clicked (conditional on where it is shown) is the same and the ads are ranked solely based on bids. This is an approximation: in practice, ads’ “clickabilities” may differ, and auctions rank ads based not only on the bids, but also on their quality scores.

Second, in the simulations, the same “CTR curve” was used for all keywords. The click-through rate (CTR) curve is a function that estimates the ratios of the numbers of clicks the same ad would receive in different positions on the screen. The CTR curve used in simulations was calibrated to the average estimated CTR curve for a number of auctions. Note also that this assumption implicitly rules out the possibility that the number of clicks that an ad receives, conditional on its position, is influenced by what other ads are shown on the screen (Jeziorski and Segal, 2015).

Third, it was assumed that the values that advertisers assigned to clicks did not depend on where on the screen the ads were shown or on which or how many other ads appeared on the screen. Athey and Ellison (2011) present an alternative model of sponsored search auctions that allows for this possibility by endogenizing advertiser values and discuss how the derivation of optimal reserve prices in that setting differs from the current one.

Fourth, in sponsored search auctions on the Yahoo! platform, advertisers could allow the platform to “advanced match” their ads, by showing them not only for the keyword on which the advertiser submitted a bid, but also for other closely related keywords (e.g., an ad for the keyword “car insurance” might also be shown to a user searching for “auto insurance”). For the purposes of the experiment, this possibility was ignored.

Next, the “theoretical optimality” of the computed reserve prices ignores the dynamic aspects of the real-world sponsored search environment: if bidders know that their bids will be used to set reserve prices in the future, they will change their bids. This problem can in principle be circumvented by setting each advertiser’s reserve price based only on the bids of *other* advertisers. However, the company’s view was that all advertisers for a given keyword should face the same quality-score-adjusted reserve price (more on that below). In addition, with sufficiently many bidders, this dynamic effect becomes small. Hence, it was ignored.

Finally, note that while the estimation procedure is based “in spirit” on the method of moments, we cannot make any claims about its consistency, because that would require a large number of independent observations for each keyword. Nevertheless, based on a number of simulations, this procedure was viewed as providing sufficiently accurate estimates to be used in practice.

4.2 Setting Reserve Prices

The theoretically optimal reserve prices were computed under the assumption that all bidders had the same quality scores and were ranked solely on the basis of their bids, which is a simplification. In practice, the ads on Yahoo! were ranked based on the product of their quality scores and bids, and the amount each advertiser paid was lower when his ad’s quality was higher. Thus, in order to keep the implementation of reserve prices consistent with the company’s ranking and pricing

philosophy, the theoretical reserve prices were converted into advertiser-specific reserve prices that reflected the quality scores of the ads: ads with higher quality scores faced lower per-click reserve prices, and vice versa. Note that this is a deviation from the theoretically optimal auction design with asymmetric bidders, to the extent that for keywords with the level of optimal reserve prices close to the original 10 cents, the change can in fact lead to a *reduction* in expected revenue.⁸

In addition, the company wanted to be conservative when implementing the first large-scale application of a previously untested theory and also wanted to experiment with various levels of reserve prices to see how much they impact revenues in practice. For that reason, each keyword was randomly assigned an “adjustment factor,” with values .4, .5, or .6 for most keywords, and the final reserve price was set equal to (optimal reserve price) \times (adjustment factor) + (10 cents) \times (1 – adjustment factor); a number between the old reserve price of 10 cents and the theoretically optimal reserve price. Simulations like the one presented in Tables 2 and 3 suggest that this “conservatism” need not be very costly, since most of the upside from the reserve prices is already obtained once the price is set at the midpoint between 10 cents and the optimal reserve price. Moreover, “overshooting” by the same amount may be considerably more costly, and therefore the seller facing uncertainty about the optimal reserve price will prefer to be conservative. This may, in fact, be a part of the explanation of the reserve price puzzle. The flatness of the revenue function around the optimal reserve price level also makes distinguishing between different “adjustment factors” statistically hard, and partly for that reason, the company allocated 95% of the keywords to the “treatment” group (with subgroups receiving different adjustment factors) and only 5% to the “control” group.^{9,10}

⁸For a basic example, consider a single-slot auction with two bidders whose per-click values are distributed uniformly from 0 to 20 cents. Bidder A receives 2 clicks per hour (and has a quality score of 2), while bidder B receives 1 click per hour (and has a quality score of 1). With a common reserve price of 10 cents per click, the expected revenue in the auction is 12.5 cents per hour (if bidder A’s value is above 10 cents, he wins the auction regardless of the value of bidder B and pays 20 cents for the two clicks that he receives; if bidder A’s value is below 10 cents and bidder B’s value is above 10 cents, bidder B wins the auction and pays 10 cents for the one click that he receives; $12.5 = 0.5 \times 20 + 0.25 \times 10$).

With personalized reserve prices, bidder A faces a reserve price of 7.5, while bidder B faces a reserve price of 15. (The average quality score of bidders in the auction is 1.5, and a hypothetical “average” bidder with that quality score would face the reserve price of 10 cents. The reserve prices to the actual bidders are set so that the product of the personalized reserve price and quality score is the same for all bidders.) The expected revenue in such an auction is approximately 11.15 cents per hour—a more than 10% decrease relative to the auction with a common reserve price.

For keywords with a high average per-click value, the benefits of moving to a higher reserve price outweigh the costs of giving advantage to stronger bidders. E.g., if in the same single-slot auction example we assume that the bidders’ per-click values are distributed uniformly from 0 to 100 cents, then the expected revenue in the “pre-intervention” auction (where every bidder faces the same reserve price of 10 cents) is approximately 43.4 cents, while the expected revenue in the “post-intervention” auction (where bidder A faces the reserve price of 37.5 cents and bidder B faces the reserve price of 75 cents) is approximately 55.8 cents.

⁹This concern did turn out to be valid: the experiment did not find any systematic differences between different adjustment factors. Hence, when discussing the results, we put all of the treatment keywords in the same group.

¹⁰Another reason for the small size of the control group was that both theoretical considerations and a smaller pilot experiment were strongly suggestive that new reserve prices would substantially increase revenues, and so allocating more keywords to the control group would be costly for the company.

5 Experimental Results

In this section, we report the results of the experiment. We started with comparing the absolute changes in revenues in the treatment and the control groups. We computed the actual overall dollar change in revenues in the treatment group and the actual overall dollar change in revenues in the control group. These numbers are not directly comparable, because the treatment group contains a lot more keywords than the control group, so we divided them by the numbers of keywords in the corresponding groups. When we performed this calculation, the change in per-keyword revenue in the treatment group was much higher than the change in per-keyword revenue in the control group. Revenues for the keywords in the treatment group increased relative to those in the control group by 12.85% of the average pre-intervention per-keyword revenue. However, this estimate is not robust: for instance, by excluding a single keyword from the control sample, this number can be reduced to around 8%. The reason why this estimate is not robust is that average revenues per keyword are affected not only by the bids of the advertisers, but also by the number of searches per keyword, and the number of searches per keyword turns out to be highly skewed and, for some keywords, highly volatile. To address this issue, we do two things. First, we exclude from our sample the top 0.1% of keywords by search volume.¹¹ Second, in addition to looking at the change in total revenue for each keyword, we look at the change in revenue-per-search, which is not directly affected by search volume (see Section 5.1 for details).

Summary statistics and the test of treatment-control balance are presented in Table 5. The unit of observation in the experiment is an auction market, separate for each keyword (one auction market for keyword “car insurance,” another auction market for keyword “cheap laptop,” and so on), and the sample contains 438,198 observations in the treatment group and 22,989 observations in the control group. The data in Table 5 come from a 30-day period before the introduction of reserve prices, in May and June of 2008. We normalized to 1 both the average revenue per keyword over that period and the average revenue per search.¹² The other statistics are reported without renormalization. The average keyword in our sample was searched by users 232 times over the 30-day period and had on average 5.9 advertisers active in the auction (“depth”). The average estimated optimal reserve price was approximately 35 cents. Comparing the summary statistics for the treatment group and the control group, for only one variable (depth) is the difference statistically significant, and the absolute value of that difference is small.

¹¹For completeness, in the Online Appendix (Table A.1) we also report the results for the uncensored sample.

¹²This normalization was needed to keep revenue per search confidential. The formulas for the normalization procedures are as follows. Suppose each keyword i (in the sample of n keywords) was searched s_i times over the 30-day period and generated the total revenue of R_i (and its revenue per search is therefore $RPS_i = \frac{R_i}{s_i}$). That keyword’s normalized revenue is then equal to $\hat{R}_i = R_i \frac{\sum_{j=1}^n R_j}{\sum_{j=1}^n R_j^2}$ and its normalized revenue per search is equal to $\widehat{RPS}_i = RPS_i \frac{\sum_{j=1}^n RPS_j}{\sum_{j=1}^n RPS_j^2}$.

Table 5: Summary statistics and test of treatment-control balance

Variable	All	Treatment	Control	Difference	p-value
Revenue [†]	1 (10.4636)	0.9984 (10.4614)	1.0305 (10.5060)	-0.0321 (0.0713)	0.6505
Revenue per search [†]	1 (2.0807)	0.9996 (2.0793)	1.0085 (2.1078)	-0.0089 (0.0142)	0.5262
Number of searches	231.94 (855.19)	231.93 (854.62)	232.15 (865.86)	-0.22 (5.50)	0.9697
Depth	5.9173 (1.9176)	5.9159 (1.9172)	5.9427 (1.9258)	-0.0268 (0.0129)	0.0389**
Estimated optimal reserve price	0.3515 (0.5129)	0.3514 (0.5129)	0.3525 (0.5124)	-0.0011 (0.0034)	0.7596
Sample size	461,187	438,198	22,989		

[†]Revenue and Revenue per search are renormalized, to the average value of one across the overall sample. *, **, ***—significant at 10%, 5%, and 1% levels. Numbers in parentheses give the standard deviations for the statistics in Columns 1–3 and the standard errors for the differences in Column 4.

5.1 Outcome Measures

To measure the effect of new reserve prices on various quantities of interest, we consider differences-in-differences estimates: we compare the pre-intervention to post-intervention change in the average quantity of interest in the treatment group to that change in the control group.¹³ We look at the effects of new reserve prices on three outcome variables. First, we look at the effect on “depth”—the average number of advertisers whose bids exceed the reserve price and whose ads are thus shown to the search engine users. Second, we look at the effect on the average monthly revenue per keyword. While this outcome measure is very natural, the estimates based on average revenue per keyword are unfortunately not robust. The reason for that is that the average revenue per keyword is affected not only by the bids of the advertisers, but also by the number of searches per keyword in a given period. Our experimental intervention has no effect on the number of searches, yet that variable is highly skewed and highly volatile, to the point that the results for just a handful of keywords can substantially impact the overall estimates for the 400,000+ keyword sample. To address this issue, the last outcome variable that we consider is “revenue-per-search”: for each keyword, we compute the average revenue generated by the search engine every time a user searches for this keyword, and then look at averages of these revenue-per-search estimates across keywords for various subsamples. This measure is not affected by the highly skewed and volatile number of searches, and is not disproportionately affected by outliers.¹⁴

¹³The pre-intervention data come from a 30-day period before the introduction of reserve prices, in May and June of 2008; then several weeks of data are skipped, because new reserve prices were phased in gradually and advertisers had grace periods before the new reserve prices became binding; and then the post-intervention data come from a 30-day period after all reserve prices were phased in and all grace periods ended, in August of 2008.

¹⁴For each subsample that we analyze, for both Revenue and Revenue-per-search outcome variables, we report the results in percentage terms relative to the pre-intervention averages of those variables in the subsample.

We study the effects of new reserve prices on the full sample of keywords and on several subsamples. First, we compare the effects of new reserve prices on the subsamples of frequently vs. rarely searched keywords. Rarely searched keywords are relatively less important for advertisers, who are therefore less likely to adjust their bids in response to changes in reserve prices. Since in GSP auctions, most of the impact of reserve prices on revenue is due to advertisers adjusting their bids (Edelman and Schwarz, 2010, Section II.A), we thus expect the effect of reserve prices to be higher for frequently searched keywords (which is what we find; see more detail below and in Table 6). Second, we compare the effects of new reserve prices on subsamples of keywords with high vs. low optimal reserve prices. We expect the effect to be stronger in the subsample with high optimal reserve prices, because for both subsamples, the “old” reserve price is the same (10 cents), and thus the relative change in reserve price is higher for the subsample with high optimal reserve prices. Moreover, as we discussed in Section 4.2, for keywords with optimal reserve prices close to 10 cents the theory suggests that the impact of treatment may be negative. That is what we find; see more detail below and in Table 7. Third, we compare the effects of new reserve prices on subsamples of “deep” and “shallow” keywords (those with many advertisers vs. few advertisers). Theory predicts that holding all else equal, the effect of reserve prices on revenue should be stronger for shallow keywords (which is consistent with our findings: the point estimate of the impact is greater for shallow keywords, although it is not statistically significant; see more detail below and in Table 8).

5.2 Results

The results for the overall sample are reported in the first column of Table 6. The introduction of new reserve prices has a strong negative effect on the number of advertisements shown on the page: on average, new reserve prices reduce this number by 0.91, i.e., almost one fewer ad per page is shown as a result of these reserve prices. This number is highly statistically significant. This effect should not be surprising: as Table 4 shows, most reserve prices were raised substantially, thus “pricing out” many advertisers whose ads were previously shown. Next, we consider the effect on the average revenue per keyword. The new reserve prices raised revenues by 3.8%, although the estimate is not statistically significant. Finally, looking at revenue-per-search estimates, the new reserve prices *reduced* revenue per search for the average keyword by 1.45% (although this estimate is also not statistically significant). While this contrast between the signs of the effects of new reserve prices on the average revenue and on the average revenue-per-search may at first appear puzzling, the reason for the difference becomes clear when we split the overall sample into the subsamples by search volume: the rarely searched keywords (those that are on average searched less than ten times per day) and the frequently searched ones (those that are on average searched at least ten times per day). The latter sample contains only 12.6% of keywords, but receives 66.9% of searches and generates 75.1% of revenues. The impact of new reserve prices on the average

Table 6: Results (full sample, split by search volume)

	Full sample	< 10 searches per day	≥ 10 searches per day
Δ -in- Δ Depth	-0.91***	-0.90***	-0.97***
• t -statistic	[-80.4]	[-75.5]	[-27.8]
• p -value	(< 0.0001)	(< 0.0001)	(< 0.0001)
Δ -in- Δ Revenue	3.80%	10.34%	2.06%
• t -statistic	[0.94]	[1.19]	[0.45]
• p -value	(0.347)	(0.235)	(0.653)
Δ -in- Δ Revenue per search	-1.45%	-2.53%**	3.90%**
• t -statistic	[-1.55]	[-2.36]	[2.31]
• p -value	(0.121)	(0.018)	(0.021)
N. obs. in treatment group	438,198	382,860	55,338
N. obs. in control group	22,989	20,133	2,856
Fraction of total revenue	100%	24.9%	75.1%

*, **, ***—significant at 10%, 5%, and 1% levels. Changes in Revenue and Revenue per search are reported relative to the average revenue and average revenue per search in the corresponding subsample before the experiment.

revenue per search in the former subsample is negative, reducing it by 2.5%,^{15,16} while the effect in the latter subsample is positive, at 3.9%, with both numbers being highly statistically significant. When one computes the average impact on revenue-per-search for the overall sample, the first subsample dominates (since it contains the majority of the keywords), while when one computes the average impact on revenue, the second subsample dominates (since it is responsible for the majority of revenue), resulting in the net positive impact on overall auction revenue.

Given the size of the sponsored search advertising market, each percentage point of positive impact translates into potential improvements to search engine profits and revenues on the order of hundreds of millions of dollars per year. Moreover, by identifying the segments of keywords where new reserve prices perform relatively poorly and modifying them accordingly, the impact can be further improved.

We have already presented one such example above, by splitting the sample into rarely and frequently searched keywords. In the next two sections, we consider two additional sets of subsamples: keywords with high and low theoretically optimal reserve prices, and then keywords with many and with few advertisers.¹⁷

¹⁵As discussed in Section 5.1, one likely reason why the effect of reserve prices on revenues for rarely searched keywords is negative is that advertisers simply do not spend much time optimizing bids on these keywords. They may set bids less carefully or update them less frequently, thus making the theory less applicable. Another related possibility is that low search volumes result in less accurate estimates of bidder values, which in turn leads our methodology to set less accurate reserve prices for these keywords.

¹⁶The effect on average revenue (rather than revenue per search) appears to be large and positive, at 10.3%, but this number is driven by outliers and is not robust. E.g., removing just two keywords from the treatment group reduces this number to 0.19%.

¹⁷The additional “cuts” of data are presented for the subsamples on which the new reserve prices have a statistically

5.3 Experimental Results by Reserve Price Level

One dimension along which keywords differ substantially is the estimated theoretically optimal reserve prices, varying from 9 cents for the 10th percentile of keywords to 72 cents for the 90th. Since the original reserve price for all keywords was equal to 10 cents, the shift from the old reserve prices to the new ones (set midway between the old and the theoretically optimal ones) is relatively much more important for the keywords with optimal reserve prices much larger than 10 cents than for the keywords with optimal reserve prices close to 10 cents, and thus we expect the impact on revenues in these two groups to be different as well.

Table 7 presents the results for two subsamples of keywords: those with the optimal reserve price lower than 20 cents and those with the optimal reserve price greater than or equal to 20 cents (following the analysis in the previous section, we restrict attention to the sample of frequently searched keywords). These two subsamples are of comparable sizes; however, the revenue generated by the subsample with the lower optimal reserve prices is an order of magnitude smaller than the revenue generated by the subsample with the higher reserve prices, because the keywords in the latter subsample have, on average, much higher revenues per search than those in the former. The average keyword in the second subsample receives approximately 34% more searches, and 8.2 times as much revenue per search, as the average keyword in the first subsample.

For keywords with high theoretically optimal reserve prices, the intervention is very successful: the impact of new reserve prices on revenue per search is equal to 4.9% and is highly statistically significant. For the other group, however, the intervention reduced revenues per search by a large amount: 8.7%. Both the increase in revenue in the former subsample and the decrease in revenue in the latter subsample are consistent with the theory, as we discussed in Section 4.2. Under the old regime all bidders faced reserve prices of 10 cents. Under the new regime, consistent with the rest of the company’s pricing practices, but inconsistent with optimal auction theory, bidders with higher quality scores faced lower reserve prices, leading to a decrease in revenues for keywords in which the average reserve price remained relatively close to 10 cents.

5.4 Results by the Number of Advertisers

Looking at the subsample of frequently searched keywords with the estimated optimal reserve prices of at least 20 cents, we further split the data by “depth”: the average number of bidders placing ads on the keyword (pre-intervention). Theory predicts that reserve prices should be particularly effective for relatively “shallow” keywords that have relatively few advertisers, and less effective for “deeper” keywords. Table 8 presents the results for two further subsamples of keywords: those with the average depth of less than 5.5 and those with the average depth of a least 5.5 (where 5.5 is the median depth in the full sample of keywords). The results are consistent with theory: the average impact on revenue per search in the “shallow” subsample is 7.8%, noticeably higher than

significant positive effect on revenue per search (frequently searched keywords for Table 7 and frequently searched keywords with high optimal reserve prices for Table 8). For completeness, in the Online Appendix (Table A.2), we present the results for these “cuts” of data applied to the overall sample.

Table 7: Results (keywords with at least 10 searches per day, split by the level of estimated optimal reserve price)

	Full subsample	$r^* < 20\text{¢}$	$r^* \geq 20\text{¢}$
Δ -in- Δ Depth	-0.97***	-1.00***	-0.94***
• t -statistic	[-27.8]	[-21.2]	[-19.9]
• p -value	(< 0.0001)	(< 0.0001)	(< 0.0001)
Δ -in- Δ Revenue	2.06%	-13.64%*	3.06%
• t -statistic	[0.45]	[-1.66]	[0.64]
• p -value	(0.653)	(0.097)	(0.525)
Δ -in- Δ Revenue per search	3.90%**	-8.73%**	4.88%***
• t -statistic	[2.31]	[-2.04]	[2.75]
• p -value	(0.021)	(0.042)	(0.006)
N. obs. in treatment group	55,338	21,760	33,578
N. obs. in control group	2,856	1,122	1,734
Fraction of total revenue	75.1%	4.6%	70.5%

*, **, ***—significant at 10%, 5%, and 1% levels. Changes in Revenue and Revenue per search are reported relative to the average revenue and average revenue per search in the corresponding subsample before the experiment.

Table 8: Results (keywords with at least 10 searches per day and the estimated optimal reserve price of at least 20 cents, split by the average number of advertisers)

	Full subsample	depth < 5.5	depth \geq 5.5
Δ -in- Δ Depth	-0.94***	-0.98***	-0.92***
• t -statistic	[-19.9]	[-18.2]	[-15.6]
• p -value	(< 0.0001)	(< 0.0001)	(< 0.0001)
Δ -in- Δ Revenue	3.06%	8.63%	2.48%
• t -statistic	[0.64]	[1.08]	[0.47]
• p -value	(0.525)	(0.280)	(0.639)
Δ -in- Δ Revenue per search	4.88%***	7.83%	4.51%**
• t -statistic	[2.75]	[1.22]	[2.48]
• p -value	(0.006)	(0.223)	(0.013)
N. obs. in treatment group	33,578	11,378	22,200
N. obs. in control group	1,734	590	1,144
Fraction of total revenue	70.5%	7.1%	63.4%

*, **, ***—significant at 10%, 5%, and 1% levels. Changes in Revenue and Revenue per search are reported relative to the average revenue and average revenue per search in the corresponding subsample before the experiment.

the average impact in the “deep” subsample (4.5%). However, the first of these estimates is not statistically significant, so we view this evidence as merely suggestive rather than conclusive.

6 Conclusion

The results of the experiment described in this paper show that setting appropriate reserve prices can lead to substantial increases in auction revenues. These results also show (to the best of our knowledge, for the first time) that the theory of optimal auction design is directly applicable in practice. Following the experiment, Yahoo! continued using and further fine-tuning this methodology for setting reserve prices. An executive described the overall impact of improved reserve prices on company revenues as follows:

On the [revenue per search] front I mentioned we grew 11% year-over-year in the quarter [...], so that's north of a 20% gap search growth rate in the US and that is a factor of, attributed to rolling out a number of the product upgrades we've been doing. [Market Reserve Pricing] was probably the most significant in terms of its impact in the quarter. We had a full quarter impact of that in Q3, but we still have the benefit of rolling that around the world.

Sue Decker, President, Yahoo! Inc. Q3 2008 Earnings Call.^{18,19}

Following the circulation of an early working paper version of our results, other researchers and companies have also experimented with using and extending our approach to setting reserve prices in sponsored search auctions. Sun et al. (2014) report promising simulation results using data from the largest Chinese search engine, Baidu (although the paper does not present any experimental results). Topinsky (2014) presents the results of a controlled experiment conducted at the largest Russian search engine, Yandex, using methodology similar to ours. The results of the experiment confirm our findings: the introduction of theory-based reserve prices has led to a substantial increase in auction revenues.

References

- [1] Aggarwal, Gagan, Ashish Goel, and Rajeev Motwani (2006), “Truthful Auctions for Pricing Search Keywords,” *Proceedings of the 7th ACM Conference on Electronic Commerce (EC06)*, pp. 1–7.
- [2] Armstrong, Mark (2000), “Optimal Multi-Object Auctions,” *Review of Economic Studies*, 67(3), pp. 455–481.
- [3] Athey, Susan, Peter Cramton, and Allan Ingraham (2002), “Setting the Upset Price in British Columbia Timber Auctions,” Working paper.

¹⁸<http://seekingalpha.com/article/101002-yahoo-inc-q3-2008-earnings-call-transcript>.

¹⁹The experiment described in this paper was one of several experiments comprising the Market Reserve Pricing project.

- [4] Athey, Susan, and Glenn Ellison (2011), “Position Auctions with Consumer Search,” *Quarterly Journal of Economics*, 126(3), pp. 1213–1270.
- [5] Avery, Christopher, and Terrence Hendershott (2000), “Bundling and Optimal Auctions of Multiple Products,” *Review of Economic Studies*, 67(3), pp. 483–497.
- [6] Bajari, Patrick, and Ali Hortaçsu (2003), “The Winner’s Curse, Reserve Prices, and Endogenous Entry: Empirical Insights from eBay Auctions,” *RAND Journal of Economics*, 34(2), pp. 329–355.
- [7] Brown, Jennifer, and John Morgan (2009), “How Much Is a Dollar Worth? Tipping Versus Equilibrium Coexistence on Competing Online Auction Sites,” *Journal of Political Economy*, 117(4), pp. 668–700.
- [8] Clarke, Edward H. (1971), “Multipart Pricing of Public Goods,” *Public Choice*, 11(1), pp. 17–33.
- [9] Cremer, Jacques, and Richard P. McLean (1988), “Full Extraction of the Surplus in Bayesian and Dominant Strategy Auctions,” *Econometrica*, 56(6), pp. 1247–1257.
- [10] Edelman, Benjamin and Michael Ostrovsky (2007), “Strategic Bidder Behavior in Sponsored Search Auctions,” *Decision Support Systems*, 43(1), pp. 192–198.
- [11] Edelman, Benjamin, Michael Ostrovsky, and Michael Schwarz (2007), “Internet Advertising and the Generalized Second Price Auction: Selling Billions of Dollars Worth of Keywords,” *American Economic Review*, 97(1), pp. 242–259.
- [12] Edelman, Benjamin and Michael Schwarz (2010), “Optimal Auction Design and Equilibrium Selection in Sponsored Search Auctions,” *American Economic Review Papers and Proceedings*, 100(2), pp. 597–602.
- [13] Ewerhart, Christian (2013), “Regular Type Distributions in Mechanism Design and ρ -Concavity,” *Economic Theory*, 53(3), pp. 591–603.
- [14] Groves, Theodore (1973), “Incentives in Teams,” *Econometrica*, 41(4), pp. 617–31.
- [15] Haile, Philip A., and Elie Tamer (2003), “Inference with an Incomplete Model of English Auctions,” *Journal of Political Economy*, 111(1), pp. 1–51.
- [16] Iyengar, Garud, and Anuj Kumar (2006), “Characterizing Optimal Keyword Auctions,” *Proceedings of the 2nd Workshop on Sponsored Search Auctions, ACM EC’06*.
- [17] Jeziorskiy, Przemyslaw, and Ilya Segal (2015), “What Makes Them Click: Empirical Analysis of Consumer Demand for Search Advertising,” *American Economic Journal: Microeconomics*, 7(3), pp. 24–53.
- [18] Krishna, Vijay (2009), *Auction Theory*, 2nd ed. (San Diego, CA: Academic Press).
- [19] Maskin, Eric, and John Riley (1984), “Optimal Auctions with Risk Averse Buyers,” *Econometrica*, 52(6), pp. 1473–1518.
- [20] Maskin, Eric, and John Riley (1989), “Optimal Multi-Unit Auctions,” in *The Economics of Missing Markets*, ed. by Frank Hahn (Oxford, U.K.: Oxford University Press).

- [21] McAfee, R. Preston, John McMillan, and Philip J. Reny (1989), “Extracting the Surplus in a Common Value Auction,” *Econometrica*, 57(6), pp. 1451–1459.
- [22] McAfee, R. Preston, Daniel C. Quan, and Daniel R. Vincent (2002), “How to Set Minimum Acceptable Bids, with Application to Real Estate Auctions,” *Journal of Industrial Economics*, 50(4), pp. 391–416.
- [23] McAfee, R. Preston, and Daniel R. Vincent (1992), “Updating the Reserve Price in Common Value Auctions,” *American Economic Review Papers and Proceedings*, 82(2), pp. 512–518.
- [24] Myerson, Roger B. (1981), “Optimal Auction Design,” *Mathematics of Operations Research*, 6(1), pp. 58–73.
- [25] Paarsch, Harry J. (1997), “Deriving an Estimate of the Optimal Reserve Price: An Application to British Columbian Timber Sales,” *Journal of Econometrics*, 78(2), pp. 333–357.
- [26] Reiley, David (2006), “Field Experiments on the Effects of Reserve Prices in Auctions: More Magic on the Internet.” *RAND Journal of Economics*, 37(1), pp. 195–211.
- [27] Riley, John G., and William F. Samuelson (1981), “Optimal Auctions,” *American Economic Review*, 71(3), pp. 381–392.
- [28] Roughgarden, Tim, and Mukund Sundararajan (2007), “Is Efficiency Expensive?,” *Proceedings of the 3rd Workshop on Sponsored Search Auctions, WWW2007*.
- [29] Sun, Yang, Yunhong Zhou, and Xiaotie Deng (2014), “Optimal Reserve Prices in Weighted GSP Auctions,” *Electronic Commerce Research and Applications*, 13(3), pp. 178–187.
- [30] Tang, Xun (2011), “Bounds on Revenue Distributions in Counterfactual Auctions with Reserve Prices,” *RAND Journal of Economics*, 42(1), pp. 175–203.
- [31] Topinsky, Valery (2014), *Reserve Prices in Asymmetric Auctions*, Ph.D. dissertation, Higher School of Economics, Moscow. Available at <http://www.ccas.ru/avtorefe/0023d.pdf>. [In Russian]
- [32] Varian, Hal R. (2007), “Position Auctions,” *International Journal of Industrial Organization*, 25(6), pp. 1163–1178.
- [33] Vickrey, William (1961), “Counterspeculation, Auctions, and Competitive Sealed Tenders,” *Journal of Finance*, 16(1), pp. 8–37.
- [34] Walsh, William E., David C. Parkes, Tuomas Sandholm, and Craig Boutilier (2008), “Computing Reserve Prices and Identifying the Value Distribution in Real-world Auctions with Market Disruptions,” *Proceedings of the 23rd AAAI Conference on Artificial Intelligence (AAAI’08)*, pp. 1499–1502.

Online Appendix

Table A.1: Results for the sample that includes the top 0.1% of keywords by search volume

	(A)	(B)	(C)	(D)	(E)	(F)	(G)
Δ -in- Δ Depth	-0.91***	-0.90***	-0.97***	-1.01***	-0.95***	-0.97***	-0.94***
• t -statistic	[-80.5]	[-75.5]	[-28.0]	[-21.3]	[-20.1]	[-18.1]	[-15.9]
• p -value	(< 0.0001)	(< 0.0001)	(< 0.0001)	(< 0.0001)	(< 0.0001)	(< 0.0001)	(< 0.0001)
Δ -in- Δ Revenue	12.85%*	10.34%	14.03%	-18.18%**	16.04%	49.00%	12.61%
• t -statistic	[1.69]	[1.19]	[1.51]	[-2.50]	[1.63]	[1.64]	[1.21]
• p -value	(0.092)	(0.235)	(0.132)	(0.013)	(0.104)	(0.101)	(0.227)
Δ -in- Δ Revenue per search	-1.43%	-2.53%**	3.97%**	-8.62%**	4.95%***	7.99%	4.55%***
• t -statistic	[-1.53]	[-2.36]	[2.38]	[-2.02]	[2.81]	[1.25]	[2.53]
• p -value	(0.126)	(0.018)	(0.017)	(0.043)	(0.005)	(0.212)	(0.011)
N. obs. in treatment group	438,632	382,860	55,772	21,864	33,908	11,481	22,427
N. obs. in control group	23,016	20,133	2,883	1,129	1,754	595	1,159
Fraction of total revenue	100%	19.3%	80.7%	4.8%	75.9%	7.2%	68.8%

*, **, *** — significant at 10%, 5%, and 1% levels. Changes in Revenue and Revenue per search are reported relative to the average revenue and average revenue per search in the corresponding subsample before the experiment.

Columns in the table correspond to the following subsamples:

- (A) Full sample
- (B) Keywords that receive fewer than 10 searches per day before the experiment
- (C) Keywords that receive at least 10 searches per day before the experiment
- (D) Keywords with at least 10 searches per day and the estimated optimal reserve price of less than 20 cents
- (E) Keywords with at least 10 searches per day and the estimated optimal reserve price of at least 20 cents
- (F) Keywords with at least 10 searches per day, the optimal reserve price of at least 20 cents, and depth of less than 5.5
- (G) Keywords with at least 10 searches per day, the optimal reserve price of at least 20 cents, and depth of at least 5.5

Table A.2: Results for subsamples by reserve price and depth

	(A)	(B)	(C)	(D)	(E)
Δ -in- Δ Depth					
• t -statistic	-0.91*** [-80.4]	-0.86*** [-60.3]	-0.97*** [-55.0]	-0.93*** [-77.7]	-0.91*** [-52.8]
• p -value	(< 0.0001)	(< 0.0001)	(< 0.0001)	(< 0.0001)	(< 0.0001)
Δ -in- Δ Revenue					
• t -statistic	3.80% [0.94]	-7.85% [-1.60]	4.65% [1.07]	4.17% [0.86]	3.41% [0.74]
• p -value	(0.347)	(0.110)	(0.283)	(0.392)	(0.461)
Δ -in- Δ Revenue per search					
• t -statistic	-1.45% [-1.55]	-2.89%** [-2.41]	-1.25% [-1.19]	0.22% [0.11]	-2.06%* [-1.92]
• p -value	(0.121)	(0.016)	(0.235)	(0.909)	(0.055)
N. obs. in treatment group	438,198	222,145	216,053	216,937	221,261
N. obs. in control group	22,989	11,608	11,381	11,195	11,794
Fraction of total revenue	100%	7.6%	92.4%	15.1%	84.9%

*, **, ***—significant at 10%, 5%, and 1% levels. Changes in Revenue and Revenue per search are reported relative to the average revenue and average revenue per search in the corresponding subsample before the experiment.

Columns in the table correspond to the following subsamples:

- (A) Full sample
- (D) Keywords with the estimated optimal reserve price of less than 20 cents
- (E) Keywords with the estimated optimal reserve price of at least 20 cents
- (F) Keywords with depth of less than 5.5
- (G) Keywords with depth of at least 5.5