Comparing Predicted Prices in Auctions for Online Advertising

Eric Bax* and Julian Romero†

December 2, 2009

Abstract

In auctions where advertisers compete to place ads in online content, some types of bids are proportional to estimates of probabilities of clicks or conversions. The highest bidder wins the auction. As a result, the auction winner is more likely to have a bid based on an over-estimate than on an under-estimate of a probability, even if the probability estimates are unbiased. This paper explores the impact of this effect on auction revenue and fairness, and it outlines some methods to improve revenue and fairness.

Keywords: Reversion, Validation, Bias, Auction, Prediction

1 Introduction

This paper focuses on auctions to place ads in online content. Each opportunity to show an ad is called an *impression*. Advertisers buy impressions; publishers sell impressions. An exchange manages the auctions. The exchange stores bids and selection criteria for impressions from advertisers. The exchange also stores selection criteria for ads from publishers. The exchange hosts an auction for each impression offered by a publisher on the exchange. In the auction for each impression, the exchange uses selection criteria from advertisers and publishers to determine which ads are eligible for the ad slot, rewards the ad slot to an eligible ad that maximizes expected revenue for the publisher, and sets the price for the impression. The winning advertiser pays the price, which is usually partitioned between the publisher and the exchange provider. For background on auction theory and designing auctions, refer to (Milgrom 2004, Krishna 2002.)

Some advertisers, called *brand advertisers*, show ads to build brand awareness. Brand advertisers generally prefer paying by the impression, known as CPM pricing. The abbreviation CPM stands for cost per mille, which is a thousand impressions. For

*Yahoo! Marketplace Design, PO Box 60543, Pasadena, CA 91116-6543. ebax@yahoo-inc.com

†Division of the Humanities and Social Sciences, Mail Code 228-77, California Institute of Technology, Pasadena, CA 91125. jnr@hss.caltech.edu

simplicity, we ignore the fact that prices for CPM are for a thousand impressions and instead use CPM to mean cost per impression.

Other advertisers, called *performance advertisers*, show ads to encourage users to complete some *action* - for example, clicking on the ad to see a landing page designated by the advertiser, offering the advertiser contact information by filling out a form on a landing page, or making a purchase. Performance advertisers generally prefer paying by the action, called CPA (cost per action) pricing. In CPA pricing, the advertiser pays for an impression only if showing their ad on the impression leads to an action. CPC (cost per click) pricing is a common special case of CPA pricing.

Brand advertisers and performance advertisers can compete for the same impressions based on the expectations of the values of their offers. The expected value of an offer is called its *eCPM*, meaning expected CPM. For CPM pricing, eCPM is simply equal to the advertiser's bid for the impression. For CPA pricing, eCPM is the advertiser's bid for an action times the probability that the action will result from awarding the impression to the advertiser, called the *probability of action*.

The probability of action is not known exactly, so the exchange must estimate the probability of action, and hence estimate the eCPM for performance advertisers. In contrast, the eCPM for brand advertisers is known exactly. So when the exchange selects a winner to maximize eCPM, it compares exact eCPMs for brand advertisers to estimated eCPMs for performance advertisers. If an auction has multiple ads from performance advertisers and multiple ads from brand advertisers, the auction can be seen as first selecting a maximum estimated eCPM among performance ads and then comparing it to a maximum exact eCPM among brand ads.

Even if the estimates of eCPM are unbiased, taking the maximum among estimates introduces bias, making the expected maximum among estimated eCPMs greater than the expected maximum among actual eCPMs for the performance advertisers. (Think of flipping seven fair coins ten times each to select a coin with a maximum rate of heads as the "winner." The winner's experimental rate of heads is likely to be greater than its true probability of heads.) This bias is a hallmark of uniform validation (Vapnik and Chervonenkis 1971, Vapnik 1998, Devroye et. al. 1996.) The fact that the winner's actual eCPM is likely to be less than the estimated eCPM that caused it to be selected as a winner can also be viewed as a manifestation of reversion to the mean or regression to the mean (Galton 1886, Samuels 1991.)

Methods to correct for uncertainty in estimating statistics of payoffs have been developed for portfolio allocations in financial markets (Jorion 1986, Jobson et. al. 1979, Lintner 1965.) The methods explored in this paper are similar to these efforts. Display advertising auctions differ from portfolio allocations in financial markets in that the auctions have single winners. So the goal in this paper is to correct uncertainty in a way that focuses on improving the selection of a single winner rather than correcting uncertainty for statistics over the whole set of offers. However, it is possible to treat inventory allocation in online display advertising as a portfolio allocation problem rather than a series of auctions (Bax and Gopalakrishnan 2009.) For search advertising auctions, there are multiple winners. Selecting those winners is similar to selecting weights for multiple investments in a portfolio. Methods for financial markets use the same correction methods and parameters across investments. In contrast, this paper explores using different corrections to select winners for different slots in search auctions.

The bias introduced by maximizing over estimates impacts fairness and revenue. In general, advertisers with less accurate eCPM estimates gain an unfair advantage over advertisers with more accurate eCPM estimates. Specifically, performance advertisers obtain an unfair advantage over brand advertisers. Since performance advertisers pay based on real actions rather than on estimates, the publishers and exchange suffer a loss of revenue. In more detail, when the bias in eCPM causes a performance advertiser to win an auction over a brand advertiser, the expected loss of revenue is the difference between the CPM bid for the brand advertiser and the actual (lower) eCPM for the performance advertiser.

The revenue loss to the seller due to selecting an offer with the highest estimated payoff is similar to the well-known phenomenon called the winner's curse (Thaler 1988, Krishna 2002.) The winner's curse occurs when multiple bidders estimate the value of an item and submit bids based on those estimates. The auction, by selecting the highest bid, tends to select a bid based on an overestimate of value. As a result, the winner tends to realize less value than their bid. Both the winner's curse and the revenue loss studied in this paper are the result of the difference between actual values and first order statistics (David and Nagaraja 2003) of estimates of values. The revenue loss studied in this paper is borne by the seller or market-maker, because the seller or market-maker must estimate the values of bids and bears the exposure from misestimation. When the bidders, rather than the market-maker, bear the risk, there is a known method to correct for bias (Wilson 1969.)

The rest of the paper is organized as follows. Section 2 introduces definitions and notations for online advertising auctions. Section 3 analyzes the impact of estimating probabilities of actions on auction revenue and fairness. Sections 4 and 5 outline methods to improve revenue and fairness. The method in Section 4 can operate without information about the accuracy of probability estimates. The more sophisticated method in Section 5 uses information about how probabilities are estimated to adjust the estimates. Section 6 proposes methods to use with search auctions. Section 7 concludes with advice on applying the techniques from this paper in practice and suggestions for further work.

2 Online Advertising Auctions - Definitions and Notation

This section introduces an auction for impressions. In the initial sections of this paper, we focus on auctions for display advertising, which feature competition among offers with different payment types (CPM, CPC, and CPA) and produce a single winning offer. In contrast, auctions for text advertising (which includes search advertising) usually feature only CPC offers and produce a ranked list of winning offers. We focus on search auctions in Section 6.

The auction selects as the winner an offer that maximizes expected cost per impression. The auction is second-price; a runner-up offer determines the price charged to the winner. The term *offer* refers to an advertiser bidding to show an advertisement on an impression.

Let b_i be the bid for offer i:

- For a CPM offer, b_i is the price the advertiser is willing to pay to show its ad.
- For a CPC offer, b_i is the price the advertiser is willing to pay for a click on its ad.
- For a CPA offer, b_i is the price the advertiser is willing to pay for an action/conversion.

Let p_i be the predicted probability of a payout for offer i:

- For a CPM offer, p_i is 1.
- For a CPC offer, p_i is the predicted probability of a click.
- For a CPA offer, p_i is the predicted probability of an action/conversion.

The (predicted) expected payout for offer i is

$$e_i \equiv p_i b_i.$$

Let

$$w = \arg\max_i e_i.$$

Then offer w wins the auction. (In case of a tie, select w uniformly at random from indices of tied expected payouts.) Let

$$s = \arg\max_{i \neq w} e_i.$$

Call offer s the second-place offer. The charge for the winning offer is

$$a_w = \min \left\{ \frac{p_s b_s}{p_w} + \varepsilon, b_w \right\},$$

where ε is the minimum bid increment, usually \$0.01. If offer w is a CPM offer, then the advertiser is charged a_w. For a CPC offer, the advertiser is charged a_w if the ad is clicked. For a CPA offer, the advertised is charge a_w if showing the ad leads to a conversion.

Let p_w^* be the actual probability of action for which p_w is an estimate. Then the expected revenue from the auction is

$$r = a_w p_w^*.$$

In most cases, the auction provider and the publisher get a fixed share of the revenue. So both have an interest in increasing revenue.

3 Impact of Estimating Probabilities on Revenue and Fairness

In this section, we analyze how estimating probabilities affects auction revenue and fairness. Estimated probabilities of clicks and conversions are a component of estimated eCPMs. Auction winners are chosen based on estimated eCPMs. Expected auction revenues are based on actual eCPMs. So estimated probabilities produce variances in estimated eCPMs, which in turn can produce unfairness or lost revenue in auctions.

3.1 Estimating Probabilities of Actions

To begin, consider how estimates of probability vary from true probabilities. There are many methods to estimate probabilities of clicks and conversions. The simplest method is Bernoulli sampling, where the fraction of auction wins that result in a click or conversion is the estimated probability. Generally, methods begin with a prior based on results for similar ads and content. Then, Bernoulli sampling is used to modify the estimated probability. In essence, most methods try to "partially borrow" samples from other ads and content, for which there is plenty of data, and then tune the estimate based on actual performance of the ad on the same or similar content. As samples accumulate, the estimated probability is based more and more on Bernoulli sampling. The following analysis focuses on Bernoulli sampling, but the general principles also apply to more complex prediction methods.

Let p^* be the actual probability of action for a performance ad. Let p be the estimate of probability based on Bernoulli sampling. The estimate p has a binomial distribution, with mean p^*. If the ad is shown n times, resulting in k actions, then

$$p \equiv \frac{k}{n}.$$

The estimate p is unbiased:

$$\mu(p) = p^*.$$

The variance is

$$\sigma^2(p) = \frac{np^*(1-p^*)}{n^2} = \frac{p^*(1-p^*)}{n}.$$

The standard deviation is

$$\sigma(p) = \frac{\sqrt{p^*(1-p^*)}}{\sqrt{n}}.$$

For click prediction, p^* is on the order of 0.01. For conversion prediction, it is on the order of 0.001. In both cases, the square root of $1-p^*$ is very close to one. So, by the definition of standard deviation

$$\mathbb{E}|p-p^*| \approx \frac{\sqrt{p^*}}{\sqrt{n}}.$$

Now consider how the estimate p affects estimated eCPM. Let v be the estimated eCPM for an offer. Let v^* be the actual eCPM. Let b be the bid for the offer. There is no bias:

$$\mathbb{E}(v-v^*) = \mathbb{E}(bp-bp^*) = b\mathbb{E}(p-p^*) = 0.$$

For the standard deviation,

$$E|v-v^*| = E|bp-bp^*| = bE|p-p^*| \approx b\frac{\sqrt{p^*}}{\sqrt{n}}.$$

Consider this as a fraction of actual eCPM:

$$E\frac{|v - v^*|}{v^*} \approx \frac{b\frac{\sqrt{p^*}}{\sqrt{n}}}{bp^*} = \frac{1}{\sqrt{np^*}}.$$

So the relative error due to estimating eCPM grows as p^* shrinks. As a result, CPA ads are likely to have much more inaccurate estimates than CPC ads. For example, with $n = 10,000$ samples and $p^* = 0.001$ for a CPA ad, the estimated eCPM is expected to differ from the actual eCPM by about 33%. In contrast, for a CPC ad with $p^* = 0.01$ and the same number of samples, the expected relative error is only about 10%. (For a CPM ad the relative error is 0%, since the probability of payout is known with certainty.)

The gap in accuracy between CPA and CPC ads can be even worse in practice than in these examples. Conversions for one ad may be based on different actions than conversions on other ads. So using conversions from one ad to estimate conversion probabilities for other ads is usually less effective than doing so for click probabilities.

Another way to view the formula above is that to keep relative error constant, if the probability of action p^* shrinks by some factor, then the number of auctions needed to learn p^* must grow by the same factor. For our examples with $p^* = 0.01$ for CPC ads and $p^* = 0.001$ for CPA ads, ten times as many learning auctions must be devoted to each CPA ad as to each CPC ad in order to have the same expected relative error for both.

3.2 Impact on Revenue and Fairness

The auction selects a winner by maximizing estimated eCPM. A fair auction would maximize actual eCPM. So it makes sense to ask whether there are systematic differences between offers selected by maximizing estimated eCPM and offers selected by maximizing actual eCPM.

Consider an auction with two offers having the same actual eCPM, v^*. Suppose one offer is based on an estimated probability, and its estimated eCPM has probability $1/2$ of being above v^* and probability $1/2$ of being below v^*. Suppose the other offer has CPM pricing, so the "estimated" probability is exactly v^*. Each ad has probability $1/2$ of being the auction winner.

Now consider an auction with twenty offers, all having actual eCPM v^*. Suppose ten offers are based on estimated probabilities, and each estimated eCPM independently has probability $1/2$ of being above v^* and probability $1/2$ of being below v^*. Suppose the other ten offers have CPM pricing. What is the probability that a CPM offer will win an auction? For this to happen, all ten offers based on estimated probabilities must be below v^*. So the probability is $1/1024$. Yet the CPM offers have the same eCPM as the other offers, and they make up half the offers. Obviously, this is unfair. These examples illustrate that offers based on estimated probabilities gain an unfair advantage as a class when there are more of them in an auction. In general, offers based on estimates with higher variances gain an advantage as a class.

Estimated probabilities also affect revenue. Recall that the expected revenue from each auction is

$$r = a_w p_w^*,$$

which is the charge to the winner based on estimated probabilities times the winner's actual probability of action. Recall that the charge is

$$a_w = \min\left\{\frac{p_s b_s}{p_w} + \varepsilon, b_w\right\}.$$

For analysis, ignore the added ε. Then

$$a_w = \frac{p_s b_s}{p_w}.$$

So expected revenue is

$$r = \frac{p_s b_s}{p_w} p_w^*.$$

Define r^* to be the ideal expected revenue - the expected revenue if the auction could select winning and second-place offers based on true probabilities of actions rather than based on estimates. Define the ideal winning index

$$w^* = \arg\max_i \; p_i^* b_i.$$

Define the ideal second-place index

$$s^* = \arg\max_{i \neq w} \; p_i^* b_i.$$

Then the ideal expected revenue is

$$r^* = \frac{p_{s^*}^* b_{s^*}}{p_{w^*}^*} p_{w^*}^*.$$

As an aside, the ideal expected revenue is the revenue gained from the second-price auction if there is no uncertainty about response probabilities. However, it is not necessarily the optimal revenue for the publisher, because the second-price auction mechanism is not necessarily revenue-optimizing (Myerson 1981, Lahie and Pennock 2007.)

Define the *revenue impact* to be the portion of ideal expected revenue lost due to using expected probabilities:

$$R = \frac{r^* - r}{r^*}.$$

Table 1 shows results of simulations to determine the impact on revenue from using estimated probabilities. Each entry in the table is the average revenue impact, expressed as a percentage, over one million simulated auctions. Each auction has ten CPM offers, ten CPC offers, and ten CPA offers. Each CPM offer has $p^* = 1.0$; each CPC offer has $p^* = 0.01$, and each CPA offer has $p^* = 0.001$. Probabilities of actions are estimated independently for each offer. Random values are drawn according to Bernoulli distributions to simulate using from $n = 10,000$ up to $n = 100,000$ auctions to learn each probability of action. For each offer in each auction, the bid is drawn independently from a distribution that makes the distribution of actual offer values a normal with mean \$1 and standard deviation \$0.10.

Table 1: Number of Learning Auctions Per Offer - Impact on Revenue

		CPA									
	learning auctions (000's)	**10**	**20**	**30**	**40**	**50**	**60**	**70**	**80**	**90**	**100**
CPC	**10**	18.6	14.0	11.5	9.9	8.9	8.2	7.6	7.2	6.9	6.6
	20	19.0	14.3	11.6	9.9	8.6	7.7	7.0	6.4	5.9	5.6
	30	19.1	14.5	11.7	9.9	8.6	7.6	6.8	6.2	5.7	5.3
	40	19.2	14.5	11.8	10.0	8.6	7.6	6.8	6.2	5.6	5.2
	50	19.3	14.6	11.9	10.0	8.6	7.6	6.8	6.1	5.6	5.2
	60	19.3	14.7	11.9	10.0	8.6	7.6	6.8	6.1	5.6	5.1
	70	19.3	14.7	11.9	10.0	8.7	7.6	6.8	6.1	5.6	5.1
	80	19.3	14.7	11.9	10.1	8.7	7.6	6.8	6.1	5.6	5.1
	90	19.3	14.7	12.0	10.1	8.7	7.6	6.8	6.1	5.6	5.1
	100	19.3	14.7	12.0	10.1	8.7	7.6	6.8	6.1	5.5	5.1

Consider the results in Table 1. Note that there is a significant (18.6%) impact on revenue when $n = 10,000$ auctions are used to estimate each probability of action. Using more auctions to learn, and hence using better probability estimates, significantly reduces revenue impact. In addition, the revenue impact from marginal learning varies based on the amount of learning already and on the probability of action:

- Decreasing differences between successive entries in each row illustrate that there are diminishing returns from marginal learning for the CPA offers in this simulation.
- Given a budget of learning auctions, how the learning auctions are partitioned among offers affects revenue. To get a sense of the effects of different partitions, compare values in entries on the same rise-to-the-right diagonals in Table 1.
- Marginal learning for offers with higher probabilities of action can actually decrease revenue. To see this, observe that the values increase down the first few rows of Table 1. As predicted values of CPC offers converge to actual values, CPA offers with poorly predicted values dominate the auctions.

Table 2 shows data on fairness collected from the simulations that produced Table 1. Each entry shows the percentage of simulated auctions won by the ideal winner, i.e., the percentage with $w = w^*$. Note that:

- Marginal learning increases fairness.
- There are diminishing returns from marginal learning.
- Marginal learning auctions contribute more to fairness when applied to offers with less accurate predictions of values.

Table 2: Number of Learning Auctions Per Offer - Impact on Fairness

		CPA									
	learning auctions (000's)	**10**	**20**	**30**	**40**	**50**	**60**	**70**	**80**	**90**	**100**
CPC	**10**	11.6	16.7	20.6	23.7	26	27.9	29.4	30.7	31.8	32.8
	20	11.7	17.3	21.9	25.5	28.6	31.1	33.2	34.9	36.4	37.8
	30	11.7	17.5	22.3	26.3	29.5	32.3	34.6	36.7	38.4	39.9
	40	11.8	17.7	22.5	26.7	30.2	33.1	35.5	37.7	39.6	41.2
	50	11.8	17.7	22.7	27.0	30.6	33.6	36.1	38.4	40.3	41.9
	60	11.8	17.8	22.9	27.2	30.8	33.9	36.6	38.8	40.8	42.5
	70	11.8	17.8	23.0	27.3	31.0	34.2	36.8	39.1	41.1	42.8
	80	11.9	17.9	23.1	27.5	31.1	34.4	37.0	39.4	41.5	43.4
	90	11.8	17.9	23.1	27.6	31.3	34.5	37.3	39.8	41.7	43.6
	100	11.9	18.0	23.1	27.6	31.4	34.6	37.5	39.8	42.0	43.8

4 A Basic Method to Improve Revenue and Fairness

The problem with estimating probabilities of action is that an offer with an over-estimated probability is likely to win the auction. If there are several predicted offers in an auction, then the offers with the greatest predicted values are likely to have over-predicted values, even if the individual estimates are unbiased. So one way to address the problem is to reduce all estimated probabilities. If the reduction is not too severe, then, on average, it will improve the value estimates for the offers with greatest predicted values. These are the offers of interest for revenue and fairness, since they are the candidates for winner and runner-up in the auction.

Here is a method to reduce probability estimates:

- Subtract a fraction d of each probability estimate to form an adjusted estimate: $\hat{p} = p\,(1 - d)$.
- Use the adjusted estimates in auctions to determine the winner, runner up, and amount to charge the winner if there is an action.
- Use separate reduction factors d for CPC and CPA offers.
- Experiment with reduction factors to optimize revenue, fairness, or a combination of them.

Table 3 shows the impact on revenue from using different reduction factors for CPC and CPA probability estimates. The impact on revenue is measured as percent of ideal revenue lost by using reduced probability estimates. Each entry is the average over one million simulated auctions. For each auction, five CPM, five CPC, and five

CPA offers are randomly generated, as follows. Actual values for offers are drawn independently from a normal distribution with mean \$1.00 and standard deviation \$0.10. For CPC offers $p^* = 0.01$, and for CPA offers $p^* = 0.001$. Simulated training based $n = 10,000$ learning auctions determines estimates p of probabilities of action. (To simulate training, p is drawn at random from a Bernoulli distribution based on n and p^*.)

Table 3: Adjustments to Probability Estimates - Impact on Revenue

		CPA									
	penalties d	**0**	**0.1**	**0.2**	**0.3**	**0.4**	**0.5**	**0.6**	**0.7**	**0.8**	**0.9**
CPC	**0**	16.4	12.6	9.1	7.0	6.4	6.4	6.4	6.4	6.4	6.4
	0.1	17.7	13.7	9.2	6.0	4.9	4.8	4.8	4.8	4.8	4.8
	0.2	18.3	14.5	10.2	7.1	5.9	5.9	5.9	5.9	5.9	5.9
	0.3	18.4	14.8	10.6	7.6	6.5	6.5	6.6	6.6	6.6	6.6
	0.4	18.5	14.8	10.6	7.6	6.6	6.6	6.6	6.6	6.6	6.6
	0.5	18.4	14.8	10.6	7.6	6.6	6.6	6.6	6.6	6.6	6.6
	0.6	18.4	14.8	10.6	7.6	6.6	6.6	6.6	6.6	6.6	6.6
	0.7	18.5	14.8	10.6	7.6	6.6	6.6	6.6	6.6	6.6	6.6
	0.8	18.5	14.8	10.6	7.6	6.6	6.6	6.6	6.6	6.6	6.6
	0.9	18.4	14.8	10.6	7.6	6.6	6.6	6.6	6.6	6.6	6.6

Consider the results in Table 3. Note that:

- Selecting revenue-optimizing penalties for CPC and CPA bids increases revenue by over 10%.
- The revenue-optimizing penalties for CPA range from 0.5 to 0.9. For CPA offers, with $n = 10,000$ and $p^* = 0.001$, the standard deviation in the estimate of p is about one third of p^*. Using a penalty of 0.5, p would need to be about twice p^* for an offer to be competitive after the correction. So p would need to be about three standard deviations above its mean. Given that only about 0.3% of samples lie more than three standard deviations from the mean of a normal distribution, which approximates a Bernoulli distribution (Feller 1968), the penalties 0.5 and above are disqualifying almost all CPA offers.

Table 4 shows the impact on fairness from using penalties to adjust CPC and CPA probability estimates. Similar to Table 2, each entry shows the fraction of auctions won by the ideal winner for the simulated auctions for each corresponding entry in Table 3. Note that maximum fairness (shown in Table 4) occurs for penalties that nearly maximize revenue (shown in Table 3.)

Tables 5 and 6 show the impact on revenue and fairness from simulations similar to those behind Tables 3 and 4, except with $n = 50,000$ simulated learning auctions for each CPA offer. (As in the earlier simulations, $n = 10,000$ simulated learning

Table 4: Adjustments to Probability Estimates - Impact on Fairness

	CPA										
	penalties d	**0**	**0.1**	**0.2**	**0.3**	**0.4**	**0.5**	**0.6**	**0.7**	**0.8**	**0.9**
CPC	**0**	21.3	27.5	34.1	38.2	39.6	39.2	39.1	39.3	39.2	39.2
	0.1	20.2	27.2	35.3	41.2	43.0	42.7	42.5	42.6	42.6	42.5
	0.2	18.8	24.6	31.1	35.8	36.9	36.3	36.1	36.2	36.2	36.1
	0.3	18.4	23.8	29.7	33.7	34.4	33.8	33.6	33.6	33.5	33.5
	0.4	18.4	23.7	29.7	33.7	34.2	33.6	33.4	33.3	33.4	33.4
	0.5	18.4	23.7	29.7	33.7	34.2	33.5	33.3	33.3	33.3	33.4
	0.6	18.4	23.8	29.7	33.7	34.2	33.5	33.4	33.4	33.3	33.4
	0.7	18.4	23.8	29.7	33.7	34.1	33.7	33.4	33.3	33.3	33.3
	0.8	18.4	23.8	29.6	33.6	34.2	33.6	33.3	33.4	33.4	33.3
	0.9	18.5	23.8	29.7	33.7	34.2	33.6	33.4	33.4	33.3	33.3

auctions are used for each CPC offer.) With the added learning, the revenue and fairness-optimizing adjustments for CPA offers are reduced.

Table 5: Adjustments with Medium CPA Learning - Impact on Revenue

	CPA					
	penalties d	**0**	**0.1**	**0.2**	**0.3**	**0.4**
CPC	**0**	7.6	5.9	6.0	6.3	6.4
	0.1	8.2	5.0	4.3	4.7	4.8
	0.2	9.0	5.8	5.2	5.8	5.9
	0.3	9.2	6.1	5.7	6.4	6.6
	0.4	9.2	6.1	5.8	6.4	6.6

Tables 7 and 8 are similar to Tables 5 and 6, except that they use $n = 100,000$ simulated learning auctions for each CPA offer. (As in the earlier simulations, $n = 10,000$ simulated learning auctions are used for CPC offers.) This makes the expected relative error in probability estimates the same for CPA and CPC offers, since their probabilities of action differ by a factor of ten. Note that the extra learning again decreases the revenue- and fairness-optimizing adjustments for CPA offers, and it makes the optimal adjustment the same for CPC and CPA offers.

This section focused on exploring how fairness and revenue are influenced by probabilities of action and by different amounts of learning. We introduced a simple method to adjust probabilities of action and showed that, for simulated auctions, the method can improve revenue and fairness. The next section develops a more sophisticated

Table 6: Adjustments with Medium CPA Learning - Impact on Fairness

		CPA				
	penalties d	**0**	**0.1**	**0.2**	**0.3**	**0.4**
CPC	**0**	36.0	41.6	40.8	39.4	39.2
	0.1	35.6	45.1	45.4	43.0	42.5
	0.2	32.6	40.6	39.8	36.9	36.1
	0.3	31.8	39.0	37.5	34.2	33.5
	0.4	31.8	38.9	37.3	34.1	33.4

Table 7: Adjustments with High CPA Learning - Impact on Fairness

		CPA				
	penalties d	**0**	**0.1**	**0.2**	**0.3**	**0.4**
CPC	**0**	5.7	5.5	6.2	6.4	6.4
	0.1	5.5	3.8	4.4	4.8	4.8
	0.2	6.2	4.4	5.3	5.9	5.9
	0.3	6.4	4.8	5.9	6.5	6.6
	0.4	6.4	4.8	5.9	6.6	6.6

method to adjust estimated probabilities - one that uses the amount of learning behind the estimated probabilities as an input.

5 Adjust Probabilities Based on Standard Deviations of Probability Estimates

Another method to adjust probabilities is to subtract a multiple of the estimated standard deviation of the difference between p and p^* from p. We use a standard deviation as the unit of adjustment because different estimates (for different values of n and p) have roughly equal probabilities of deviating by equal numbers of standard deviations, based on the similarity between Bernoulli distributions and the normal distribution (Feller 1968). We estimate the standard deviation

$$\sqrt{\frac{p^*\left(1-p^*\right)}{n}},$$

by using the estimate p in place of (the unknown) p^*. So the adjusted probability estimate is

$$\hat{p} = p - c\sqrt{\frac{p\left(1-p\right)}{n}},$$

Table 8: Adjustments with High CPA Learning - Impact on Fairness

		CPA				
	penalties d	**0**	**0.1**	**0.2**	**0.3**	**0.4**
CPC	**0**	42.3	43.8	40.2	39.2	39.2
	0.1	43.6	49.3	44.6	42.7	42.5
	0.2	40.1	44.6	38.8	36.3	36.1
	0.3	39.1	42.7	36.3	33.8	33.6
	0.4	39.1	42.6	36.1	33.6	33.4

where c is a coefficient selected to optimize revenue, fairness, or a combination of them. This probability adjustment is similar to the Wilson correction for estimating binomial proportions (Wilson 1927, Brown et. al. 2001.) Use the adjusted probability estimate in place of p to select a winner, a runner up, and set the charge to the winner in case of an action in each auction.

Figures 1 to 4 show the impact of adjustments based on standard deviations in simulated auctions. There are two scenarios - low and high CPA learning. In low CPA learning, p is drawn based on a Bernoulli distribution that simulates learning with n = 10,000 auctions, for both CPC and CPA offers. For high CPA learning, n = 100,000 for CPA offers, and n = 10,000 for CPC offers. For each scenario, each plotted value is based on the same 100,000 simulated auctions, using a different adjustment coefficient c. Each auction has ten randomly generated offers. Each offer is selected from a uniform distribution over CPM, CPC, and CPA offer types. CPC offers have p^*=0.01. CPA offers have $p^* = 0.001$. Actual values for offers are drawn uniformly at random from a normal distribution with mean \$1.00 and standard deviation \$0.10. Figure 1 shows the average and sample standard deviation over the 100,000 simulated auctions of revenue impact,

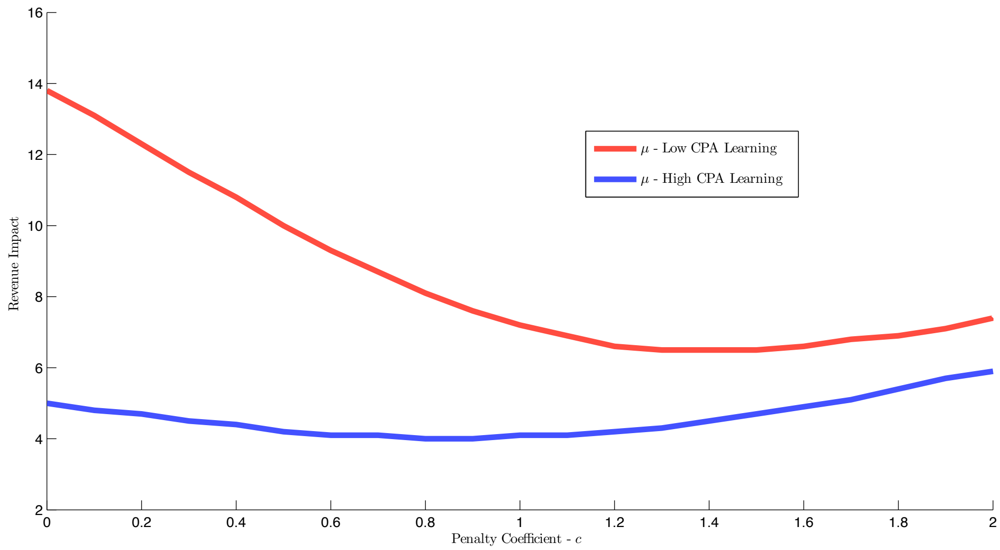
$$\hat{R} = \frac{r^* - \hat{r}}{r^*},$$

as a percentage. Figure 2 shows fairness measured as percentage of auctions won by the ideal winner. Figures 3 and 4 show the distribution of wins over auction types.

For low CPA learning, note that:

- Figure 1 shows that using a revenue-maximizing value of c increases revenue by over 7%. (The revenue-maximizing value is $c = 1.4$.)
- Figures 1 and 2 show that revenue-maximizing values of c are close to fairness-maximizing values of c, and vice versa.
- Figure 3 shows that neither revenue-maximizing nor fairness-maximizing values of c come close to distributing wins equally over the offer types. Using no adjustment ($c = 0$) causes CPA offers to win the lion's share of auctions. On the other hand, using revenue-maximizing and fairness-maximizing values of c causes CPM offers to dominate.

Figure 1: Effect of Penalty Coefficient on Revenue Impact



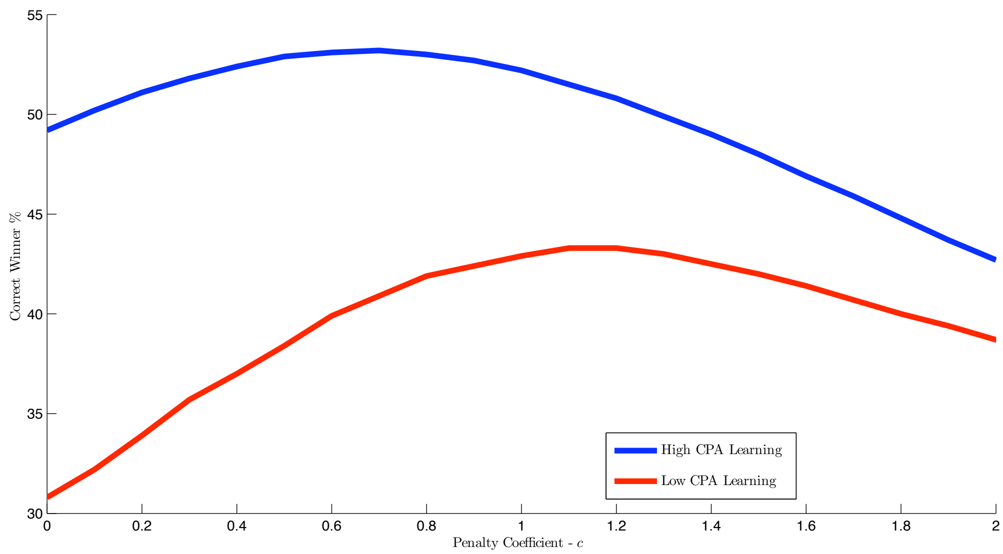
- The sample standard deviations of the revenue impact curves in Figure 1 are approximately as large as the mean revenue impact. This means that the revenue impact varies significantly from auction to auction. Still, the results are statistically significant - with 100,000 simulated auctions, the standard deviation of the estimate of the mean is approximately the sample standard deviations divided by about 300.

For high CPA learning, note that:

- In Figure 1, the revenue impact without any adjustment to probability estimates ($c = 0$) is 5%, compared to almost 14% for low CPA learning.
- The revenue improvement from using a revenue-maximizing value of c is 1%. (The revenue maximizing value is $c = 0.8$.)
- Figure 4 shows that using a revenue-maximizing value of c causes CPM offers to win more than their share of auctions. In contrast to Figure 3, the CPC and CPA shares of wins are almost equal for each value of c. This is because CPC and CPA estimates have nearly equal standard deviations in the high CPA learning scenario.

Figures 5 and 6 show how the number of offers in each auction affects revenue and fairness with adjustments based on standard deviations. Both figures are based on the same simulations. For each number of offers in $4, 8, \ldots, 40$, the data for all values of c are based on the same one million auctions. For each auction, the specified number of offers are generated independently at random. For each offer, whether it is CPM, CPC, or CPA is determined uniformly at random. For CPM offers, $p^* = 1.0$. For

Figure 2: Impact of Penalty Coefficient on Choosing Correct Winner



CPC offers, p^* is drawn uniformly at random from $\{0.005, 0.01, 0.02, 0.05\}$. For CPA offers, p^* is drawn uniformly at random from $\{0.0002, 0.0005, 0.001, 0.002\}$. For all offer types, the number n of simulated learning auctions is drawn uniformly at random from $\{5000, 10,000, 50,000, 100,000\}$. Then the estimated probability of action p is determined by drawing from a Bernoulli distribution based on p^* and n. Actual values are drawn at random from a Gaussian with mean \$1.00 and standard deviation \$0.10, and bids are set by dividing actual values by probabilities of action p^*.

As in Figures 1 and 2, revenue impact is the fraction of ideal revenue lost due to using estimated probabilities, and fairness is measured by the fraction of auctions won by the offer with highest actual value. Note that:

- As auction size grows the optimal adjustment coefficients c increase.
- The increase in optimal adjustment coefficients diminishes as auction size grows.
- Revenue and fairness impacts change smoothly with changes in the adjustment coefficient.

Table 9 summarizes the results from Figures 5 and 6. For each auction size, there are columns on revenue impact and columns on fairness. The revenue impact columns show revenue impact without an adjustment, optimal revenue impact, and the revenue-optimizing value of the adjustment coefficient. Differences between revenue impact without adjustment and optimal revenue impact indicate the potential to increase revenue by adjusting estimated probabilities. Note that:

- The potential revenue increase grows with auction size.

Figure 3: Impact of Penalty Coefficient on Type of Winner - Low CPA Learning

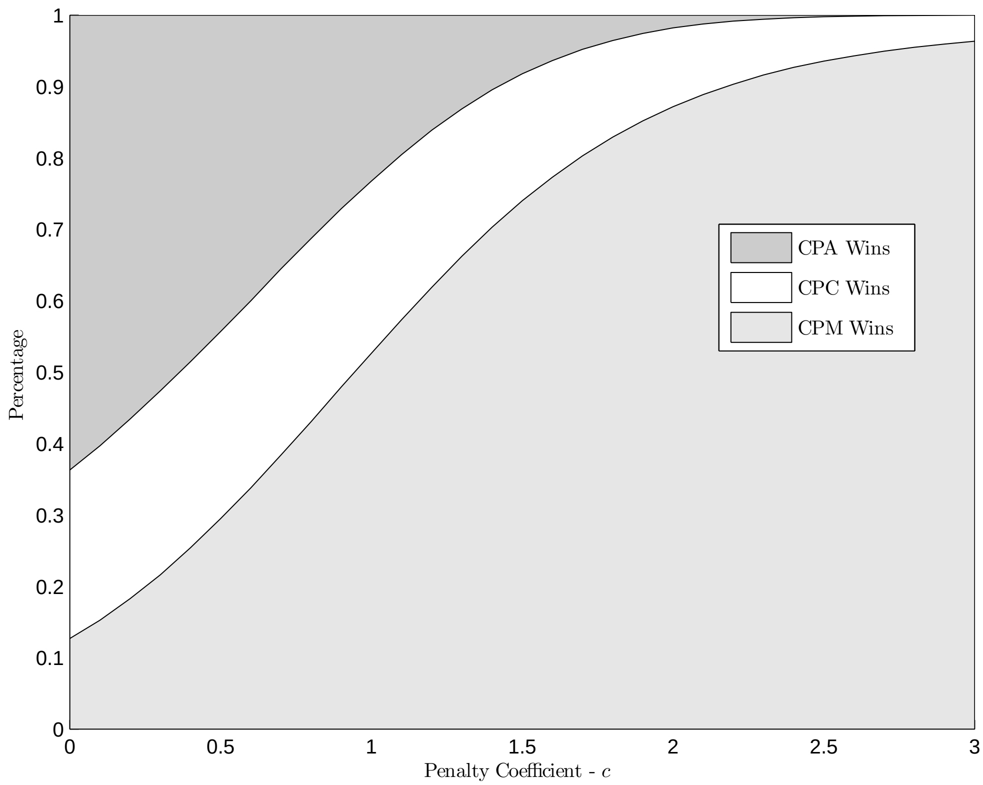


Figure 4: Impact of Penalty Coefficient on Type of Winner - High CPA Learning

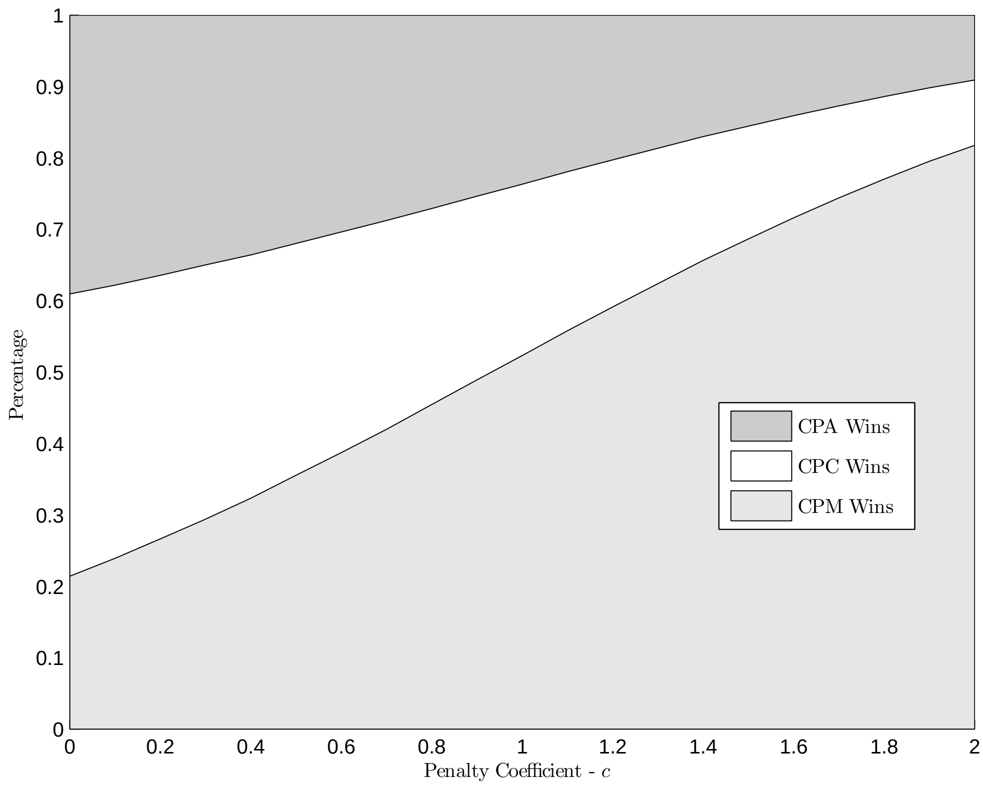


Figure 5: Revenue Impact Over Auction Sizes

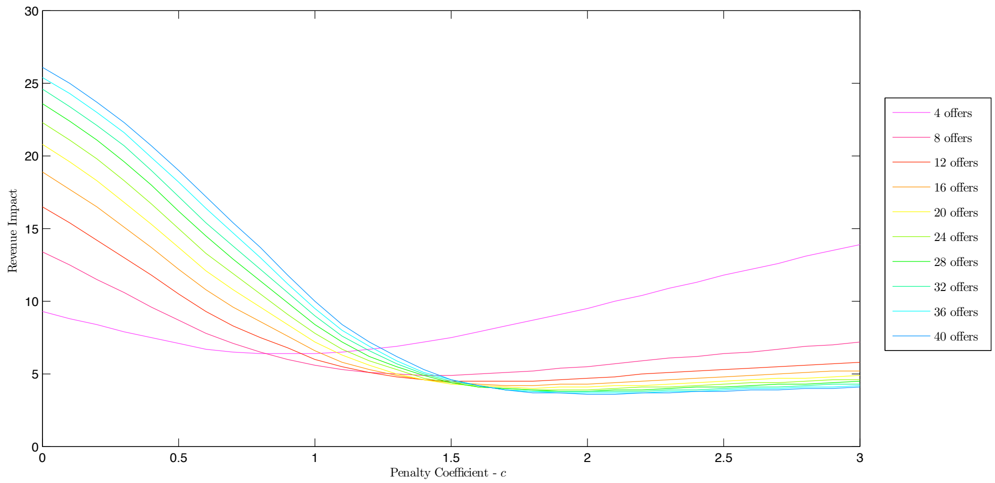


Figure 6: Impact on Fairness Over Auction Sizes

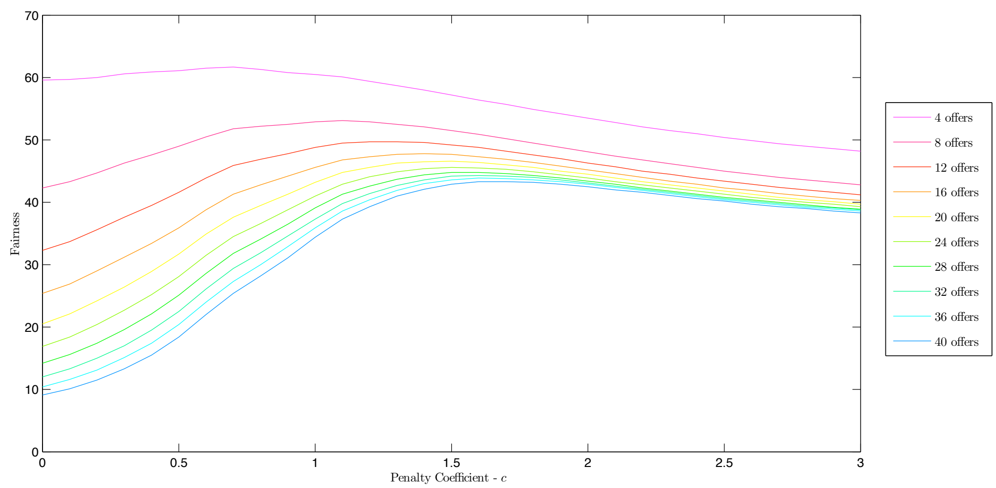


Table 9: Optimal Adjustment Coefficients Over Auction Sizes

Offers	**Revenue Impact % (c=0)**	**Optimal Revenue Imcapt %**	**Revenue Optimizing c**	**Correct Winner % (c=0)**	**Optimal Correct Winner %**	**Fairness Optimizing c**
4	9.3	6.4	0.9	59.6	61.7	0.7
8	13.4	4.9	1.4	42.3	53.1	1.1
12	16.5	4.5	1.6	32.3	49.7	1.3
16	18.9	4.2	1.7	25.4	47.8	1.4
20	20.8	4.0	1.8	20.5	46.6	1.5
24	22.3	3.9	1.9	16.9	45.6	1.5
28	23.6	3.8	1.9	14.2	44.8	1.6
32	24.6	3.8	2.0	12.0	44.3	1.6
36	25.4	3.7	2.0	10.4	43.9	1.6
40	26.1	3.6	2.0	9.1	43.3	1.7

- Revenue impact without adjustment grows with auction size - more offers create more opportunities for some offer to have a large gap between estimated and actual probability of action.
- Revenue impact with optimal adjustment shrinks with auction size - more offers mean more offers with high estimated values that are close to actual values after adjustment.

The columns on fairness show fairness without adjustment, optimal fairness, and the fairness-optimizing adjustment. The differences between fairness without adjustment and fairness with the optimal adjustment indicate the potential to increase fairness. Note that:

- The potential increase in fairness grows with auction size.
- Fairness without adjustment decreases dramatically with auction size - more offers mean more offers that have large gaps between estimated and actual values.
- Fairness with optimal adjustment also decreases with auction size - having more offers makes it more difficult to choose the very best offer, even with adjustments to estimated probabilities.

6 Search Auctions

In this section, we change focus from auctions for display advertisements to auctions for search advertisements. In the search auctions we describe here, there are multiple slots for ads on each page. So there are multiple auction winners. The auction orders offers by estimated eCPM (breaking ties randomly.) The most desirable slot is awarded to the first offer, the second most desirable slot is awarded to the second offer, and so on. The charge for each winning offer is based on the estimated eCPM of the next offer, with the intention of charging the first winner the second price, the second winner the third price, and so on. Typically, search auctions have only CPC pricing.

6.1 Notation and Model

We will use notation similar to that for display auctions. As in display auctions, for offer i, let b_i be the bid, p_i be the estimated probability of action, and p_i^* be the actual probability of action. Also, the estimated eCPM is $b_i p_i$ and the actual eCPM is $b_i p_i^*$.

We will use a generalized second-price auction model (Edelman et. al. 2007, Varian 2009.) Let m be the number of offers, and let k be the number of ad slots. For $j \in \{1, \ldots, m\}$, let w_j be the original index $i \in \{1, \ldots, m\}$ of the offer in position j after ordering offers by estimated eCPM. For example, w_1 is the index of the winner of the first ad slot. Similarly, let w_j^* be the original index i of the offer in position j after ordering offers by actual eCPM. Then expected auction revenue is

$$r = \sum_{j=1}^{k} \frac{b_{w_{j+1}} p_{w_{j+1}}}{p_{w_j}} p_{w_j}^*.$$

Similarly, ideal auction revenue is

$$r^* = \sum_{j=1}^{k} \frac{b_{w_{j+1}^*} p_{w_{j+1}^*}^*}{p_{w_j^*}^*} p_{w_j^*}^* = \sum_{j=1}^{k} b_{w_{j+1}^*} p_{w_{j+1}^*}^*.$$

As in display, define the revenue impact for search as

$$R = \frac{r^* - r}{r^*}.$$

In practice, response rates decrease as an ad moves from more to less desirable slots (Varian 2007, Blumrosen et. al. 2008, Kempe and Mahdian 2008, Gomes 2008.) For simplicity, we ignore this effect in our simulations in this section. Including this effect would increase the revenue impact from the early ad slots and decrease the impact from the later slots. When the effect is strong, the revenue from the top slot overwhelms the revenue from other slots, resembling the single-slot display auction. When the effect is weak, the effects on revenue resemble those in this section.

6.2 Adjusting Estimated Probabilities

This section explores two methods to increase revenue by adjusting estimated probabilities in search auctions. Both methods use adjustments based on standard deviations between estimated and actual probabilities of action. The first method uses the same adjustment to determine winners and prices for all ad slots. The second method uses different adjustments for different ad slots.

6.2.1 Using the Same Adjustment Across Ad Slots

In this subsection, estimated probabilities p_i are adjusted to form adjusted estimates $\hat{p}_i$ according to the formula

$$\hat{p} = p - c\sqrt{\frac{p\,(1-p)}{n}},$$

where c is a constant selected to optimize revenue, fairness, or a combination of them, and n is the number of auctions used to form the estimate p. The adjusted probability estimates $\hat{p}$ are used in place of estimates p to order the offers, select winners, and determine charges. Define $\hat{r}_s$ to be the revenue using these adjusted estimates. Define

$$\hat{R}_s = \frac{r^* - \hat{r}_s}{r^*}$$

to be the associated revenue impact when using these adjusted estimates.

Table 10 shows results of simulations to determine the revenue impact from these adjustments to the probability estimates. Each column is based on a set of 10,000 simulated auctions. Each offer is generated independently, with:

- actual value determined by a Gaussian with mean \$1 and standard deviation \$0.10.
- p^* selected uniformly at random from $\{0.005, 0.01, 0.015, \ldots 0.05\}$.
- p drawn at random based on a Bernoulli distribution simulating n learning auctions, with n selected uniformly at random from $\{1000, \ldots, 10,000\}$

For each set of 10,000 simulated auctions, the average revenue and revenue impact is calculated without adjustments and with adjustments based on using a value of c, called c^*, that optimizes revenue. The value c^* is determined using gradient descent based on a different random set of auctions than those used in the test.

From Table 10, notice that:

- The optimal coefficient c^* is increasing in the number of offers. This is due to the fact that the bias for the top five offers increases with the number of offers.
- When there are 6 offers and 5 winners, the optimal coefficient is negative. With only one more offer than the number of slots, under-estimation of values becomes a larger problem than over-estimation.
- Based on the bottom row, the adjustment generally has more impact as the number of offers increases. The exception is when there are six offers for five slots. In this case, the adjustment compensates for under-estimates, and both the absolute value of the adjustment and the revenue impact are larger than for ten offers.

6.2.2 Using Different Adjustments for Different Slots

In this subsection, estimated probabilities are adjusted differently for different ad slots. In each case, each adjusted estimate $\hat{p}$ is

$$\hat{p} = p - c\sqrt{\frac{p\,(1-p)}{n}},$$

Table 10: Revenue Impact from the Same Adjustments Across Ad Slots

		Number of Offers					
		6	**10**	**15**	**20**	**25**	**30**
	r^*	4.87	5.20	5.38	5.48	5.55	5.61
$c = 0$	r	4.69	5.00	5.13	5.19	5.23	5.26
	R	3.77	3.91	4.63	5.28	5.81	6.21
Constant c	c^*	-0.46	0.15	0.63	1.00	1.14	1.16
	$\hat{r}_s$	4.70	5.00	5.14	5.23	5.28	5.33
	$\hat{R}_s$	3.59	3.86	4.32	4.57	4.81	4.95
$\mathbf{R - \hat{R}_s}$		**0.18**	**0.09**	**0.31**	**0.71**	**1.00**	**1.26**

but different values of c are used for different ad slots. Let $\mathbf{c} = (c_1, \ldots, c_k)$ be the sequence of c-values for ad slots. Then the auction procedure is as follows. Start with slot 1. Adjust probabilities of action for all offers using c_1. Order by adjusted estimated eCPMs to determine a winner for the first slot and a charge based on the second offer in the ordering. Remove the winner. Then repeat this process, using c_2 for the second slot, c_3 for the third slot, and so on.

Define $\hat{r}_d$ to be the expected revenue using this procedure. Define

$$\hat{R}_d = \frac{r^* - \hat{r}_d}{r^*}$$

to be the revenue impact.

Table 11 shows results of simulations to determine the revenue impact. Each column is based on the same set of 10,000 simulated auctions as the corresponding column in Table 10. For each column, an optimal value of $\mathbf{c}$, called $\mathbf{c}^*$, is computed using gradient descent over a different set of auctions than those used for the results shown in Table 11.

From Table 11, notice that:

- With six offers and five slots, the optimal adjustments for the top slots decrease probabilities to compensate for over-estimates, and the optimal adjustments for the remaining slots increase probabilities to compensate for under-estimates.
- For each number of offers, the optimal adjustments are largest for the top slots. Hence, adjustments increase as bias increases for each successive slot. (The bias for the winner is greater than the bias for the runner up, and so on.)
- Slot-by-slot, optimal adjustments increase with number of offers, as expected.

Table 11: Revenue Impact from Different Adjustments Across Ad Slots

		Number of Offers					
		6	**10**	**15**	**20**	**25**	**30**
	r^*	4.87	5.20	5.38	5.48	5.55	5.61
$c = 0$	r	4.69	5.00	5.13	5.19	5.23	5.26
	R	3.77	3.91	4.63	5.28	5.81	6.21
Variable c	$\mathbf{c}^*$	0.41	0.78	1.04	1.23	1.37	1.46
		0.10	0.54	0.87	1.07	1.24	1.31
		−0.31	0.26	0.75	0.94	1.18	1.22
		−0.68	0.12	0.67	0.86	1.14	1.16
		−1.15	−0.01	0.58	0.83	1.11	1.12
	$\hat{r}$	4.72	5.01	5.15	5.23	5.29	5.33
	$\hat{R}_d$	3.22	3.72	4.23	4.54	4.77	4.91
$\mathbf{R} - \hat{\mathbf{R}}_\mathbf{d}$		**0.55**	**0.19**	**0.40**	**0.74**	**1.04**	**1.30**

- Observe the bottom row. For ten and more offers, the value of using the adjustment increases with the number of offers. For six offers, the value is higher than for ten and twenty offers, because the procedure makes an effective adjustment for the strong negative bias found in later ad slots.

Table 12 compares using the same adjustment for all slots to using different adjustments for different slots. From the bottom row, notice that the advantage of using different adjustments is greater when there are fewer offers. When there are many offers, the optimal adjustments for different slots are closer to each other, so there is less advantage in allowing different coefficients.

7 Conclusion

This paper explores the impact of using estimates of offer values in an auction. We have shown that using estimates introduces a bias that can significantly reduce revenue and fairness. This paper also outlines a few methods to correct for the bias, improving revenue and fairness. The methods adjust estimated values of offers to be used in auctions. The adjustments are based on number of offers in each auction and the

Table 12: Comparison of Same and Different Coefficients

	Number of Offers					
	6	**10**	**15**	**20**	**25**	**30**
$\hat{R}_s$	3.59	3.86	4.32	4.57	4.81	4.95
$\hat{R}_d$	3.22	3.72	4.23	4.54	4.77	4.91
$\mathbf{\hat{R}_s - \hat{R}_d}$	**0.37**	**0.14**	**0.11**	**0.03**	**0.04**	**0.04**

amount of learning for each offer. The general technique is to correct for the bias by adjusting estimated offer values based on the number of competitive offers in each auction and based on information about the distributions of actual values for each of the estimated values.

The methods in this paper have free parameters. To apply the methods in practice, it is possible to use simulations to select starting points for the parameters. Then use statistical optimization techniques (Box et. al. 2005) to adjust the parameters, optimizing for any desired combination of revenue and fairness. Fortunately, the simulations in this paper indicate that revenue-optimal parameter settings are similar to fairness-optimal ones.

For the parameter c in the methods that adjust based on standard deviation, it makes sense to use values of c based on the number k of offers in each auction, as in classical shrinkage methods such as Wilson's correction (Wilson 1927, Brown et. al. 2001) or James-Stein estimation (James and Stein 1961, Stein 1955.) For some ad calls, many ad offers are eligible but most are uncompetitive. It may be useful to determine a subset of offers that are competitive and then apply shrinkage based on the number of competitive offers instead of the total number of eligible offers. (The simulations in this paper are based on auctions in which all or almost all offers are competitive.) Most shrinkage methods are designed to minimize average error over the quantities being estimated (Brown 1966, Bock 1975.) For fairness in auctions that select a single winner, it would be interesting to explore whether there are estimators that tend to select the variable with highest actual mean directly, rather than first applying shrinkage methods and then selecting the maximum estimate.

In general, it is possible to use any of a variety of machine learning approaches to determine functional forms for the adjustments and set parameters for those forms. Inputs can include the number of offers, their estimated values, and any available information about the distributions of actual values, such as how much frequencies of action have varied over time for each offer or for sets of offers. Since online advertising marketplaces hold many auctions, the amount of data needed for machine learning approaches is available to them. One such approach is to use Bayesian principles (Berger 1985, Duda and Hart 1971, Gelman et. al. 2004.), basing adjustments on priors developed using empirical data from past auctions.

The simulations for search advertising auctions indicate that using different correction factors for different ad slots can improve revenue and fairness. It would be interesting to explore whether a similar tactic can improve methods to correct for uncertainty in portfolio allocations for financial markets (Jorion 1986, Jobson et. al. 1979, Lintner 1965.) For example, it may be useful to apply one correction to all available investments, select one or a few investments to receive a portion of the resource allocation, remove those investments, apply a weaker correction to those remaining, and then select among them to receive the remaining resource allocation.

References

Bax, E. & R. Gopalakrishnan. 2009. "Randomness and uncertainty in portfolio methods for inventory allocation in online advertising." *submitted to ACM Transactions on the Web* .

Berger, J. O. 1985. *Statistical Decision Theory and Bayesian Analysis (2nd Edition)*. Springer.

Blumrosen, L., J. D. Hartline & S. Nong. 2008. "Position auctions and non-uniform conversion rates." *ACM EC Workshop on Advertisement Auctions* .

Bock, M. E. 1975. "Minimax estimators of the mean of a multivariate distribution." *Annals of Statistics* 3(1):209–218.

Box, GEP, WG Hunter & JS Hunter. 2005. *Statistics for Experimenters: Design, Innovation and Discovery*. Wiley.

Brown, L. D. 1966. "On the inadmissibility of invariant estimators of one or more location parameters." *Annals of Mathematical Statistics* 37:1087–1136.

Brown, L. D., T. T. Cai & A. DasGupta. 2001. "Interval estimation for a binomial proportion." *Statistical Science* 16(2):101–133.

David, H. A. & H. N. Nagaraja. 2003. *Order Statistics (3rd Edition)*. Wiley.

Devroye, L., L. Györfi & G. Lugosi. 1996. *A Probabilistic Theory of Pattern Recognition*. Springer.

Duda, R.O., P.E. Hart et al. 1973. *Pattern Classification and Scene Analysis*. Wiley.

Edelman, B., M. Ostrovsky & M. Schwarz. 2007. "Internet advertising and the generalized second-price auction: selling billions of dollars worth of keywords." *American Economic Review* 97:242–259.

Feller, W. 1968. *An Introduction to Probability Theory and Its Applications, Volume 1*. Wiley.

Galton, F. 1886. "Regression towards mediocrity in hereditary stature." *Journal of the Anthropological Institute of Great Britain and Ireland* pp. 246–263.

Gelman, A. 2004. *Bayesian Data Analysis*. CRC Press.

Gomes, R., N. Immorlica & V. Markakis. 2008. "Externalities in keyword auctions: an empirical and theoretical assessment." *ACM EC Workshop on Advertisement Auctions* .

James, W. & C. Stein. 1961. "Estimation with quadratic loss." *Proc. Fourth Berkeley Symp. Math. Statist. Prob.* 1:361–379.

Jobson, J. D., B. Korkie & V. Ratti. 1979. "Improved estimation for markowitz portfolios using James-Stein type estimators." *Proc. of the American Stat. Assoc., Business and Economics Statistics Section* 41:279–284.

Jorion, P. 1986. "Bayes-Stein estimation for portfolio analysis." *Journal of Financial and Quantitative Analysis* pp. 279–292.

Kempe, D. & M. Mahdian. 2008. "A cascade model for externalities in sponsored search." *ACM EC Workshop on Advertisement Auctions* .

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

Lahie, S. & D. M. Pennock. 2007. "Revenue analysis of a family of ranking rules for keyword auctions." *ACM Conference on Electronic Commerce* pp. 50–56.

Lintner, J. 1965. "The valuation of risk assests and the selection of risky investments in stock portfolios and capital budgets." *The Review of Economics and Statistics* 47(1):13–39.

Milgrom, P.R. 2004. *Putting Auction Theory to Work.* Cambridge University Press.

Myerson, R. B. 1981. "Optimal auction design." *Mathematics of Operations Research* 6(1):58–73.

Samuels, M.L. 1991. "Statistical reversion toward the mean: More universal than regression toward the mean." *American Statistician* pp. 344–346.

Stein, C. 1955. "Inadmissibility of the usual estimator for the mean of a multivariate normal distribution." *Proceedings of the Third Berkeley Symp. on Prob. and Stat.* 1:197–206.

Thaler, R. H. 1988. "Anomalies: the winner's curse." *The Journal of Economic Perspectives* 2(1):191–202.

Vapnik, V. 1998. *Statistical Learning Theory.* Wiley.

Vapnik, V. N. & A. Ya. Chervonenkis. 1971. "On the Uniform Convergence of Relative Frequencies of Events to Their Probabilities." *Theory of Probability and its Applications* 16(2):264–280.

Varian, H. R. 2006. "Position auctions." *International Journal of Industrial Organization* 25:1163–1178.

Varian, H. R. 2009. "Online ad auctions." *American Economic Review* 99:430–434.

Wilson, E. B. 1927. "Probable inference, the law of succession, and statistical inference." *Journal of the American Statistical Association* 22:209–212.

Wilson, R. B. 1969. "Competitive bidding with disparate information." *Management Science* 15(7):446–448.