Pricing Guidance in Ad Sale Negotiations: The PrintAds Example

Adam Isaac Juda
juda@google.com

S. Muthukrishnan
muthu@google.com

Ashish Rastogi
rastogi@google.com

ABSTRACT

We consider negotiations between publishers and advertisers in a marketplace for ads. Motivated by Google's online PrintAds system which is such a marketplace, we focus on the role of the market runner in improving market efficiency. We abstract the problem of *pricing guidance* where the market runner provides an initial price-point for negotiations based on data analysis. The problem is nuanced because the market runner can not fully reveal the price data for any of the publishers. We introduce two solutions for pricing guidance, the first using clustering and the second using support vector machines, and present experimental evaluation of our methods. Pricing guidance by the market runner is a novel direction, and we expect more research in the future.

1. INTRODUCTION

Prices for goods and services are determined in different ways. Traditionally, prices are set by the seller or provider, such as in retail stores (per item) or telephone bills (per call or as a flat fee per month). A different way to set prices is via an auction, and this is increasingly used on the Internet, including in ad sales. The focus of this paper is on yet another way prices are determined, via *negotiation*, which is used extensively in marketplaces. In particular, we consider online marketplaces where parties negotiate over the Web. The motivating example for us is that of Google's AdSense for Print product (heretofore referred as PrintAds) [1] - a marketplace for newspaper publishers and advertisers to negotiate prices for ads that appear in print. In what follows, we will provide an overview of PrintAds, motivate the problem of providing pricing guidance to the parties, and offer a solution to this problem.

PrintAds. PrintAds is a marketplace for print publishers and advertisers. Advertisers log into the PrintAds system and are helped by various targeting tools to identify desirable publishers (e.g., finding a New York metro daily newspaper with circulation above 10,000 readers). Then, they enter the details of the ad they wish to publish, such as its size, the desired section where they wish their ad to appear etc. Finally, advertisers specify an offer price they are willing to pay for the ad to run. When they submit an offer, it is sent to the publisher, who may accept the offer or decline. If the publisher accepts, the ad eventually runs and payment rendered; else, a new round of negotiation begins. [1]

The two primary parties within this marketplace are the publishers and the advertisers. Publishers partake in this market in order to reach a larger scale of potential advertisers than would otherwise be possible through their direct sale channels, while advertisers partake in the market as it more readily enables them to negotiate with multiple publishers easily. However, an important third participant within this market is Google as the *market runner*. The market runner ensures negotiations proceed smoothly and seeks to make the market as efficient as possible. Efficiency is desired by the market runner not only for altruistic reasons, but also as the market runner receives compensation only when transactions are successfully completed, and so has economic incentives to increase overall value creation.

Pricing in PrintAds. Traditionally ads in print have been sold via direct sales teams. While rate cards exist with posted prices, negotiations typically occur between the individual advertisers and the publisher, resulting in contracts that vary widely in price, different from the rate card maximums. These characteristics of contracts can also be present within PrintAds. However, because the typical PrintAds advertiser is new to advertising within newspapers, [2] negotiations tend to proceed less smoothly than offline negotiations: some offers made by advertisers may be viewed as exceedingly low by publishers, some publishers do not respond to offers tendered, etc. In order to smooth the negotiation process, the market runner needs to provide better guidance about typical prices of print inventory. This is a nuanced role. Publishers seek higher prices, close to their rate cards, but also do not want to forego the incremental income that this additional sales channel can provide. Advertisers like lower prices, but want to secure that price without having to bear the costs of entering a negotiation with the direct sales

---

[1] The system and its variants are described in more detail in Section 2.

[2] Many PrintAds customers are more familiar with Google's AdWords program, where they construct small, text-only advertisements, and pay very small amounts of money on a per click basis. This is in contrast to advertising in a newspaper where creatives are richer and payments tend to be significantly higher.

channel offline. For both parties to operate under similar expectations around a reasonable set of prices, the market runner needs to provide pricing guidance; however, the market runner does not get to set prices as these are ultimately determined by negotiations between the two parties. The goal of the market runner is only to help negotiations reach an agreement. Additionally, the market runner has to be careful that it does not compromise any private information of the publisher or advertiser during negotiation. This introduces a game-theoretic perspective to the design of such guidance.

Our Contributions. We approach the problem of providing pricing guidance as one of data analysis. As an analogy, consider how buyers make offers on real estate. There is tremendous strategy in how sellers arrive at the list price by considering comparable sale data, both to target a significant group of buyers as well as to spur a bidding war. Likewise, in a buyers' market, the buyers have an eventual price they are willing to pay, but strategize while making rounds of offers. The PrintAds marketplace offers far more complexities over this analogy. Publishers' inventory of ad space is flexible, but market demand at times can be thin, and online negotiations can run to many rounds. Our contributions are as follows.

- We model the Google PrintAds marketplace and abstract the *pricing guidance* problem. In particular, we address a concern that price prediction should not overfit any particular publisher's data.

- We propose two algorithms based on data analysis to solve the pricing guidance problem. One is based on clustering while the other is based on Support Vector Machines (SVM) with suitable kernels. Both do offline data analysis for model building, but work fast online for providing guidance.

- We present empirical analysis of our methods with data from the PrintAds system over the past year, both comparing them and summarizing some of the resulting trends when the system uses our methods.

We have over an year of experience with designing, launching and operating the pricing guidance system in Google PrintAds, and we believe that the principles we have abstracted apply broadly to any online negotiation system with a market runner. While negotiations have been studied extensively in Economics and Business, typically, the focus is on the two (or more) parties that are part of the negotiation. In online marketplaces, our work shows the role of a distinct party, i.e., the market runner, and brings focus to a nuanced task they need to perform, in particular in uncertain markets. This direction is novel and we believe will lead to more research in this area within Computer Science. In Section 8, we pose some extensions of the PrintAds negotiation system for further research.

2. THE ONLINE PRINTADS SYSTEM

Google's PrintAds [1] is an online marketplace for publishers in print media such as newspaper and magazine publishers, and their advertisers. Traditionally, ads in print media

are sold by negotiation between sales and marketing teams working for publishers and advertisers respectively. This human-intensive process has inefficiencies, e.g., each party has to separately approach each counter-party to negotiate terms. A standard solution is to create a marketplace and infuse some efficiency. Google's PrintAds system is such a marketplace. Consider the following example:

An advertiser wishes to run an Ad Campaign with The Morning Call, a newspaper in Allentown, PA with the following specifications: "A 1 column by 7 inch black and white on Sunday, Tuesday, Wednesday and Friday during the week of December 19, 2007.

Figure 2 shows the interface advertisers use to make offers to publishers. After selecting a set of publishers via a targeting tool (not the focus of this paper), the advertisers focus on The Morning Call and specify the features for the ad that they wish to run. These include: the section of the newspaper in which that ad is to appear, the size of the ad and the days of the week on which the ad is to appear.

The advertiser is asked to provide a "bid" for the ad. The interface (Figure 2) displays a spectrum for the range of acceptable offer prices for the ad. The minimum offer price is set at some fixed reserve price (called *Google min.* in the figure), which is to rule out spam bids. The list price on the publisher-issued rate card is the maximum offer price (marked as *Listed rate* in the figure). Once the advertiser submits her offer, the offer details are sent to the publisher, who may either accept the offer, or reject the bid as too low. The advertiser then has the opportunity to revise her bid and goes through the process again. Once an offer is accepted by the publisher, the ad is run with the agreed upon specifications. The advertiser and the publisher may pull out of the negotiations prior to the offer being accepted, but once accepted, the ad is guaranteed to run. Unlike Google's AdSense system that is used to price online ads, pricing in PrintAds is the outcome of negotiations and it is not based on auction.

Before we proceed, a remark about the price measurement unit in the PrintAds market. The publishers and the advertisers normalize the price as a *cost per milli per column inch* (CPMPCI), defined as:

$$\text{CPMPCI} = \frac{\text{offer in dollars}}{\frac{\text{circulation}}{1000} \times \text{ad size in column-inches}} \quad (1)$$

An important aspect of the PrintAds marketplace is its many uncertainties. First, the PrintAds market has *elastic* supply: the publishers can always print more pages or reduce certain articles to accommodate more advertisements. Second, the PrintAds system is a new sales channel in addition to all of the sales channels currently used by the publishers, including their own sales team. This introduces channel conflicts that make the design of the system formidable. Third, due to a variety of reasons, prices vary widely, even for similar advertisements. It is known (at least among large advertisers) that they can negotiate to prices below the list prices. Finally, many of the advertisers in the PrintAds system are more familiar with Google's AdWords program, and less with the print media. Thus they are uncertain about the

Figure 1: Online PrintAds System: the advertiser is making an offer for a 1 column by 7 inch ad in *The Morning Call (Allentown, PA)*. The advertiser wishes the ad to appear on the following days of the week: *Sunday, Tuesday, Wednesday and Friday*. For the weekday edition, the advertiser offers the publisher $277 per issue and for the weekend edition (that has larger circulation), the advertiser bids $470 per issue.

typical prices. This gets amplified because the advertisers need to generate suitable offers while only given the Google min and the Listed rate, which are typically far apart.

Such uncertainties lead to many frustrations. In one scenario, an advertiser makes a very low offer, close to Google min. The publisher is dissatisfied and does not respond to the offer, which in turn, disappoints the advertiser. As a result, both parties leave the system. In a different scenario, the advertiser makes a very high offer, close to the Listed rate. The Publisher accepts, but the advertiser gets small return on investment and eventually rejects the PrintAds system. Besides these two extremes, there are other unstable points, such as, the negotiation goes on for several rounds, exhausting both parties.

These uncertainties and frustrations were evident in the system at the beginning, prompting us to ask how the negotiations could be made more efficient and successful since Google as the market runner has incentives to make the system more effective. One solution will comprise the market runner buying inventory from publishers in bulk, and reselling at a more predictable market price to the advertisers, thereby being the risk taker. This and other solutions that radically changed the market were unacceptable to the parties involved, and our constraint was to work within the description of the PrintAds system above.

3. PRICING GUIDANCE

Our approach to improving the PrintAds market relies on a crucial, but simple observation, that there is an *asymmetry* of roles and information.

The publisher has access to various signals for determining how much an advertiser may be willing to pay for an advertisement, and indeed may not even sell the spot if the market yielded low prices. Indeed, publishers can fill-up empty spots with their own advertisements (so called "house ads"), and can increase or reduce the number of pages, size and number of articles etc in the publication to accommodate varying number of slots.

On the other hand, advertisers, in particular small or new advertisers, are ill-suited to guess the price a publisher may be willing to accept. Still, the advertisers are responsible for making initial offers in each round.

Breaking the asymmetry in the roles would alter the system. For example, reverse negotiations (where publishers make offers to show the ads), or alternate mechanisms (such as auctions) would essentially lead us to rebuild the PrintAds system. Instead, working within the operational PrintAds system, we isolated the underlying asymmetry of information as the key bottleneck.

In the absence of good information, the typical advertiser has taken a very conservative approach when bidding, submitting an extremely low initial offer (in comparison to the typical closing price for such an offer). While some publishers seriously consider such offers and respond to advertisers, others ignore low offers received, which presumably reduces advertiser satisfaction. Still in our discussions with them, the publishers preferred getting offers to forbidding low offers (eg., via an explicit min price), because this was valuable information for them.

Our solution to break the asymmetry in information is to provide guidance to the advertisers on suitable prices to offer, thereby revealing more information to them. More precisely, we abstract *pricing guidance* as providing a suitable start price for ad offers, carefully chosen between the Google min and Listed rate. Thus, with pricing guidance, the advertisers will see a screen as shown in the example in Figure 3. Figure 3 shows the offer portion of Figure 2 annotated with
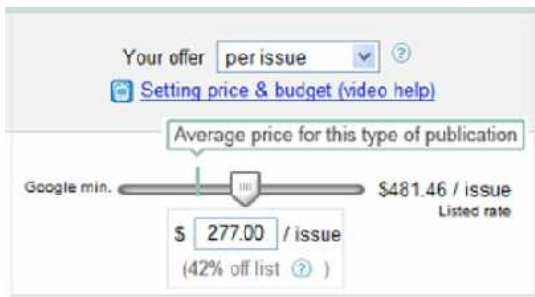
Figure 2: Pricing guidance is shown as a blue marker above. Advertisers take the guidance into account when determining their offer price.

pricing guidance in the PrintAds system. The green bar marked with the "average price for this type of publication" [3] shows the offer price that is recommended by the guidance system. Now the advertiser can use this guidance in determining an offer price (ignore it, adopt it, offer higher or lower, etc). Notice that pricing guidance is concerned with the negotiation mechanism between the advertisers and the publishers and unlike minimum reserve-prices in the auction-based online AdSense ad pricing, pricing guidance does not enforce any constraints on the final closing price of the offer (if accepted).

4. APPROACHES TO PRICING GUIDANCE

We discuss a few approaches which were considered when we began formulating the problem. The shortcomings of these approaches helped guide us towards the solution we ultimately propose.

- *Fixed % Guidance.* Fixing the guidance at $x\%$ of the rate card rate is a simple form of guidance. However, with any fixed strategy like this, the advertiser will quickly figure out that the guidance is a fixed relative number and will thereafter ignore it.

- *Price Determination.* When an advertiser wishes to make an offer to a particular publisher, we might consider the history of all the offers to that publisher thus far and use that to "predict" the ultimate or beginning price for the offer using machine learning methods. But in nascent markets such as PrintAds, publishers are reluctant to reveal their sales data in detail. This is a concern as PrintAds is an alternative channel to the established sales teams; advertisers who buy from the sales teams could check prices on the PrintAds system, and possibly request sales teams lower their rates. In equilibrium, the two channels (sales teams and online PrintAds) should converge to prices that differ only in the premium the sales teams offer, but short term this might affect revenues. Additionally, since the system reveals no information about the advertisers to the publishers, it is not fair to have it reveal any publisher-specific information to the advertisers.

- *Multipliers.* A popular approach inspired by sponsored search is to learn "multipliers" [8]. We can imagine an

[3] This language was chosen so that guidance is accessible to the advertisers.

advertiser-specific multiplier α_a for advertiser a and a publisher-specific multiplier ρ_p for publisher p, [4] with pricing guidance for an advertiser a with publisher p being determined by $\alpha_a \rho_p$, e.g., as a fraction of the maximum price from the rate card. This has the desirable effect of summarizing the effect of any advertiser or publisher, the multipliers can be learned over time, etc. An engineering hurdle here is that the data is sparse with few publishers, advertisers and advertisements (thousands rather than millions of keywords as in sponsored search); hence, machine learning is a challenge. The bigger issue is game-theoretic. Since there are only a few publishers of interest to a given advertiser, advertisers could quickly learn the relative multipliers of different publishers and use that to strategize negotiations with sales teams.

5. OUR PRICING GUIDANCE SOLUTION

We view the problem as one of data analysis and propose novel solutions that are dynamic, preserve publisher's private information during their negotiation, and accommodate well for markets of various sizes and different number of participants.

5.1 Formal Problem

We formulate the pricing guidance problem as follows. Our input is the set of accepted offers O. Each offer $o_i \in O$ has six features: circulation$_i$ (circulation), is_color$_i$ (B/W or Color), day$_i$, (whether the ad is to appear on a weekday or a weekend publication), columns$_i$, height$_i$, is_college$_i$ (whether the publication is a college publication or not) and a corresponding accepted CPMPCI ρ_o. There are other features in the system, but the ones above were the most useful. This data may be preprocessed and analyzed. When an advertiser enters the PrintAds system to put an offer, they present a *query* offer q. The problem then is to return a guidance CPMPCI ρ_q for the query. This is scaled by the area and shown as the blue guidance marker as in Figure 3. A small detail is that if the calculation above yields a price larger than the maximum listed rate (for e.g. $481.46 in the example in Fig 3), then the listed rate is taken as the guidance.

5.2 Clustering-based Solution

Our approach here is to cluster accepted offers in O based on one or more features, and summarize CPMPCI's in each cluster by a single "summary" CPMPCI value. Then, query offer q is mapped to its cluster whose summary CPMPCI value is returned as ρ_q. The underlying premise is that we can find clusters such that the summary CPMPCI's will be good pricing guidance while at the same time, each cluster will be sufficiently heterogeneous in terms of the publishers (made precise later) so it does not reveal any single publisher's data. Note that the requirement of having sufficiently many distinct publishers in each cluster is a *non-standard* requirement of clustering methods. Using insights and empirical analysis, we chose to cluster accepted offers

[4] As an analogy, in sponsored search, there are advertiser-specific (α's) and position-specific (β's) multipliers which are used to scale the bids: $\alpha_a \beta_p$ is the scale for effective bid. In fact, this is built into the auction for selling such ads and advertisers keenly track their multipliers.

in O along the circulation of the publication (some other features such as ad size were not discriminating). This was a good choice because the PrintAds system has publications with a wide variety of circulations: from national newspapers with large circulation to campus and local newspapers with very circulation of the order of tens of thousand.

Let I denote the set of publishers, and let $C = \{c[i], i \in I\}$ be the set of all publishers' circulations. Assume that elements in $C = \{c[1], \ldots, c[N]\}$ are sorted such that $c[i] < c[j]$ when $i < j$. Note that $|C| < |I|$ since there may be multiple publishers with the same circulation.[5]

Publications Partitioning Problem (PPP). Given array $A[1, N]$ where $A[i]$ is the *set* of all accepted offers with Publication ID that has circulation $c[i]$. We are also given two integer parameters k and ℓ capturing the number of partitions and the publisher heterogeneity in each partition, respectively.

The problem is to partition the array A into at most k pieces given by $i_0 = 0 < i_1 < i_2 \ldots < i_k = N + 1$. We say an offer o_a is in jth partition and denote $o_a \in J$ if the $circulation_a \in [c[i_j] + 1, c[i_{j+1}]]$. Each partition must satisfy two conditions.

- The number of distinct circulations in each partition must be at least ℓ, that is, $|\{i | A[i] \neq \phi\}| \geq \ell$, where ϕ is the empty set.

- Let μ_J be the *median* ρ_i of offers $o_i \in J$. Then, the partition should minimize *total error* given by $\sum_j \sum_{o_i \in J} |\rho_i - \mu_J|$.

Note that the second condition above is similar to the histogram problem studied well in databases and statistics. In fact, the median is known to minimize the L_1 error within each partition (eg, see [9]). Our problem above differs from the standard histogram problem by having the first condition.

Now we briefly sketch an efficient algorithm to solve PPP. The solution relies on Dynamic Programming. As is usual, we only focus on determining the optimum error; the actual partition with this error can be determined in the standard way from this calculation. Define $E(i, j)$ to be the optimum solution for $A[1, i]$ with j partitions satisfying the conditions above for $i \leq N$ and $j \leq k$ Then, $E[i, j] = \min_{1 < a < j} E[1, a] + \sum_{o_b \in I} ||\rho_b - \mu_I|$, where $I = [a + 1, j]$ and I has at least ℓ distinct circulation sizes. Using careful data structures (details omitted due to space constraints), this gives:

THEOREM 1. There is an $O(N^2 k + m \log m)$ time algorithm to solve PPP where m is the total number of accepted offers.

_____
[5]Circulation numbers reported by the publishers are often rounded off to the nearest thousand and this may result in multiple publishers having the same circulation.

5.3 Regression-based Solution
We view pricing guidance as a prediction problem and describe our regression-based solution.

SVM Regression. Support vector machines (SVMs) are widely used to learn linear classifiers by minimizing a combination of training error and a regularization term [7]. The following overview will suffice for us. In SVM regression, we receive m training samples $(x_1, y_1), \ldots, (x_m, y_m) \in X \times \mathbb{R}$, where each x_i is drawn from a fixed (but unknown) distribution D over X in an i.i.d. fashion and y_i is the corresponding target value. The problem is to find a hypothesis $h : X \mapsto \mathbb{R}$ that predicts y_x for an $x \in X$ drawn according to D. SVM-based regression minimizes the L_1-loss: define $c_1(h, x)$ as the absolute error of hypothesis h on $x \in X$. That is, $c_1(h, x) = |h(x) - y_x|$. A key advantage of the support vector method is that it is "kernelizable" – i.e. one can project the set X (the space of offers in our case) into a (potentially higher) dimension feature space through the use of kernel functions (so long as they are positive definite symmetric) and consider linear classifiers in that space. This gives more powerful hypotheses classes and is effective in practice. The SVM regression produces a hypothesis $h(x) = K(w, x) + \rho$, that minimizes the following objective function:

$$F(h, S) = \frac{1}{2}\|w\|_K^2 + \frac{C}{m} \sum_{i=1}^{m} \max\{0, c_1(h, x_i) - \epsilon\},$$

where $K(\cdot, \cdot)$ is the kernel function and ρ is a bias term and $\max\{0, c_1(h, x) - \epsilon\}$ is the ϵ-insensitive loss that the hypothesis h makes on a point x. The resulting hypothesis after minimizing the above objective function can be expressed through a set of (say l) "support vectors", $\{v_1, \ldots, v_l\} \subseteq \{x_1, \ldots, x_m\}$ and has the form $h(x) = \sum_{i=1}^{l} w_i K(v_i, x) - b$, where w_i is weight associated with support vector v_i and b is a bias term.

Pricing Guidance using SVMs. We used RBF kernels [7] and trained one prediction function h per publisher p by *excluding* any offers with p. The set of offers O denotes the input space X described above. The target value y_x for an instance x corresponds to the CPMPCI ρ for an offer $o \in O$. The output of our training process is thus a family of prediction functions $\{h_p : p \in P\}$. Excluding the offers of p from h_p ensures that p's offer do not impact pricing guidance and only data available from the market (rest of the offers) which is fair. The SMO algorithm is widely used to train SVMs and its complexity is $O(m^3)$ where m is the number of offers in the training set. Making a prediction on query q involves making l kernel computations and $l + 1$ arithmetic operations where l is the number of support vectors.

6. EXPERIMENTAL EVALUATION
6.1 The Data
The data was collected from Google's PrintAds system over a period of six months and consisted of thousands of offers. We excluded offers that corresponded to advertising packs or pre-negotiated contracts between the advertisers and publishers. For SVM regression, we normalized the features: boolean features were mapped to $\{+1, -1\}$ and numerical features were scaled to values in $[-1, 1]$. In addition,

since the circulation$_i$'s were orders of magnitude larger than the other features, we took its logarithm before normalizing. This significantly improved classifier accuracy. One feature that (surprisingly) did not affect the accepted price was whether or not the offer was placed in a bundle.

6.2 Comparison of Two Approaches

We compared the different approaches for Pricing Guidance. We compared them based on their accuracy of predicting the price at which offers were concluded at the end of the negotiation. Good accuracy means that even though our methods limit the influence of a publisher's data on pricing guidance, still, there is enough information in the overall market (even excluding that publisher) that pricing guidance makes the market efficient.

We considered two measures of accuracy, the absolute error $c_1(h, x)$ and the squared error $c_2(h, x)$, given by $c_1(h, x) = |h(x) - y|$, $c_2(h, x) = (h(x) - y)^2$. The error of the learning algorithm on a set S is denoted by the mean absolute error (MAE) $\widehat{R}_1(h, S)$ and the mean squared error (MSE) $\widehat{R}_2(h, S)$ given by:

$$\text{MAE}(h, S) = \frac{1}{|S|} \sum_{x \in S} c_1(h, x), \quad \text{MSE}(h, S) = \frac{1}{|S|} \sum_{x \in S} c_2(h, x).$$

The learning parameters C, γ for the SVM-regression methods were determined by sweeping through a range of values. The average error across the each one of the 10 folds, together with the standard deviation of this measurement are reported in Figure 3.

Note that while regression by itself does not improve the accuracy of prediction over the clustering-based solution, a careful choice of the Kernel function yields an improvement over the Clustering-based solution ($\sim 30\%$). The approach is also more "stable" in that it has a lower standard deviation. Furthermore, these conclusions are independent of the choice of the error function (i.e. mean squared error or mean absolute error).

6.3 The Influence of Pricing Guidance

Pricing guidance was introduced to PrintAds on December, 15th, 2007. We compare the CPMPCI-normalized prices accepted through the PrintAds system from September 2, 2007 through December 14th, 2007 to those prices accepted from December 15th, 2007 through January 31st, 2008, evaluating the influence that pricing guidance had on accepted prices. In particular, we are interested in determining two things: first, the extent to which pricing guidance is influencing the average prices accepted by newspapers; and second, the extent to which pricing guidance is influencing the variance of prices accepted by publishers. Given that pricing guidance is an aggregate representation of prices typically accepted by publishers in the past, pricing guidance would ideally have an immaterial influence on average prices. However, given that the goal of pricing guidance is to provide more transparency around prices, one would anticipate that variance of prices should decrease after pricing guidance is introduced (as advertisers would no longer be submitting inappropriately high or low bids).

Qualitatively, Figures 4(a) and 4(b) show the influence that pricing guidance has on PrintAds. Figure 4(a) is a histogram over publishers of the relative change in the average prices accepted after pricing guidance was introduced versus before (where publishers were omitted from the histogram if they had fewer than 10 accepted insertions either before or after pricing guidance was introduced). The average change in accepted prices among publishers after pricing guidance was introduced was 14.3% (standard deviation of 76.0%). However, this increase in price is at least partially attributed to the mix of newspaper articles accepted after pricing guidance was introduced. A fixed-effect regression model was run where normalized accepted price was regressed again the following variables: a binary variable indicating if the offer accepted after pricing guidance was introduced, a binary variable indicating if the offer for the Sunday edition of the paper, the height of the advertisement and the column-width of the advertisement. Within this regression, we were unable to reject the null hypothesis; therefore, we can not say with statistical certainty that pricing guidance was in fact influencing the average price accepted.

Figure 4(b) tells a much different story, and is a histogram over publishers of the relative change in the standard deviation of prices accepted after pricing guidance was introduced versus before. The typical publication saw a decrease in the standard deviation of prices accepted after pricing guidance was introduced. (Due to the bias that outlier publications introduced, the average standard deviation change is omitted.) In combination with the change in average prices, the directions of the influences of pricing guidance are clear. While pricing guidance did not meaningfully impact the average prices accepted by publishers, it did decrease the variance of prices accepted, providing a more uniform price experience across advertisers for any given publisher.

While the aggregate influence pricing guidance is well-behaved, we were further curious to see if there were biases among publishers for whether pricing guidance had a positive or negative influence on their change in prices. Indeed we found one. Statistical regressions were run where the change in average prices accepted was regressed against the absolute value of average prices accepted. We were able to reject the null hypothesis that the absolute value of accepted prices has no correlation to the impact of pricing guidance. We discovered that publications that had typically been accepting lowing prices were accepting relatively higher prices after pricing guidance was introduced, whereas publications that had typically been accepting higher prices were accepting relatively lower prices after pricing guidance was introduced. Interpreting this finding is nontrivial. If one believes that normalized newspaper advertising is a commodity good, then pricing guidance causing normalized prices to approach each other across publishers is a positive attribute of pricing guidance. Alternatively, if one believes that publishers offer substantially different value propositions on a normalized basis, then pricing guidance serving as an attractor may be a detrimental attribute of pricing guidance.

7. RELATED WORK

Several popular, online markets exist where a market runner brings large sets of buyers and sellers together and facilitates

Method	Mean Squared Error	Mean Absolute Error
Clustering-based Solution	0.119 ± 0.043	0.194 ± 0.034
SVR with Linear Kernels	0.124 ± 0.046	0.196 ± 0.023
SVR with RBF Kernels	0.086 ± 0.022	0.163 ± 0.014

Figure 3: A comparison of the different approaches for Pricing Guidance.
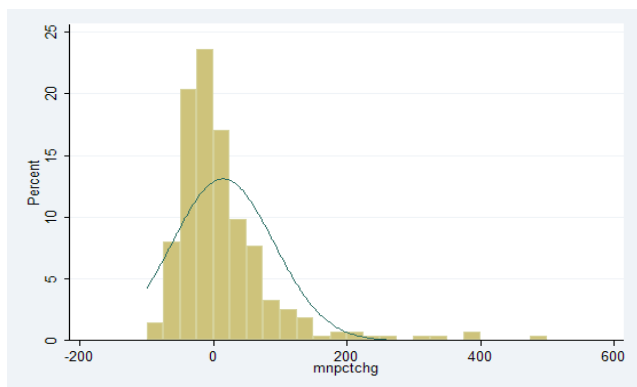
(a) Histogram over publishers of change in normalized average prices accepted after pricing guidance versus before.
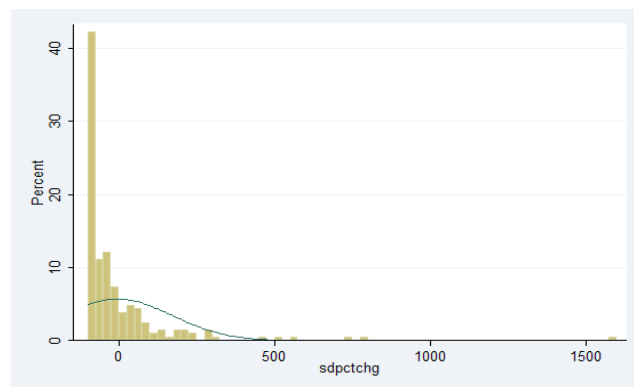
(b) Histogram over publishers of change in standard deviation of normalized prices accepted after pricing guidance versus before.

Figure 4: Changes in prices among publishers after pricing guidance was introduced to PrintAds.

transactions, including eBay [6] and Amazon [7]. However, these markets have unique characteristics that make their pricing very different than the PrintAds system. On eBay, just as within PrintAds, many goods are priced dynamically. However, there is sufficient demand for individual products on eBay that sellers of goods can establish prices via an auction. Alternatively, within PrintAds, the exponential number of possible ad insertions creates an environment where there is insufficient demand for any one item for an auction mechanism to be appropriate. On Amazon, just as within PrintAds, there are many buyers and sellers who may not otherwise be able to find each other without Amazon's assistance. However, items sold on Amazon are typically sold at a post price, as many items sold are typically also available through standard retails outlets. Alternatively, within PrintAds, there is no readily viable secondary market, and prices are negotiated.

The study of price negotiations is extensive in Economics, Business, Political Sciences and Psychology, and includes the role of information in price negotiations, game theory of bargaining, behavioral aspects of negotiations including threats, reputation, and appetite for risk and patience, and so on (for an overview, see [3, 2, 4]). The focus in such studies is on the two (or more parties) that are involved in negotiations. Our emphasis is on the role of the market runner, who is not one of the negotiating parties. This aspect may be seen as a "structural" aspect in Economics parlance, or as related to arbitration or mediation. Our formulation of the role of the market runner as an automated agent that merely initiates the starting point of price negotiations is a lightweight approach to the elaborate roles studied in literature for mediation, and arose out of the unique challenges

of the PrintAds marketplace. Note that pricing guidance is quite different from determining the "true" price of the goods (ads), a task Economists study extensively.

As discussed, pricing guidance provided in PrintAds has a game-theoretic aspect, where we wish for pricing guidance to be not directly dependent on the prices a publisher has previously accepted. If prices were dependent, we could be unfairly providing the publisher's private information to an advertiser and jeopardizing alternate sales channels of the publisher; this would also introduce a potential opportunity for the publisher to game the system. Our work developed methods where we determine a good estimate of the starting price for negotiating with a publisher based on the behavior of *other* or a *group* of publishers. These methods generate good guidance and are resistant to manipulation. At the simplest level, such methods have elements of incentive-compatible allocations where the outcome is not determined by one's bids, or more closely, the ideas in prior-free auctions [6] where values of a subset of bidders is used to set a reserve price for the remainder. However, the specifics and context of our methods are significantly different. Our pricing guidance methods may be thought of as impacting first/last mover advantages in price theory [5].

8. CONCLUDING REMARKS

We have abstracted a pricing guidance problem motivated by negotiations in an online marketplace, and in particular Google's PrintAds system. Our data-analysis method is determined dynamically and differently for each publisher, but also limits the ability for advertisers and publishers to manipulate the guidance. Our methods have been in use for more than a year, with empirical behavior suggesting that the guidance reduced the variance of normalized prices publishers are accepting, while not dramatically changing the

---

[6] www.ebay.com

[7] www.amazon.com

average accepted normalized price in aggregate.

Future Work. A deeper analyses of pricing guidance will be of interest. For example, this includes the use of offers that did not ever get accepted ("failed negotiations") and study of the game theory and dynamics of the players and the negotiation ("How often did advertisers use the suggested price and how?" "How did the number of negotiation rounds and total time of negotiation change?"). In addition, we have identified variations of the marketplace which gives more directions for research.

- When publishers decline the offer, the publishers can make a counteroffer. Now, the dynamics change. In particular, the market runner may consider building a guidance for counteroffers, a "dual" problem to the guidance problem we have studied here.

- The market runner can try to arbitrage by selling the slot to the advertiser at some price and then negotiate with the publisher for determining the price for that ad, causing the market runner to take risk. Devising the strategy for the market runner in this scenario will be of interest.

- Rather than providing pricing guidance as a single number representing a typical price, one could instead provide a range of possible prices that could be accepted, or a probability that their offer will be accepted, rejected, or counter-offered.

- An advertiser may be interested in running a campaign requiring the participation of several newspapers. This acquisition problem would be combinatorial in nature, likely requiring a more centralized solution in order to ensure that efficient transactions occur.

9. ACKNOWLEDGMENTS

10. REFERENCES

[1] Google Print Ads system. `http://adwords.google.com/support/bin/answer.py?answer=70508&ctx=sibling`

[2] A. Roth. *Game-theoetic models of bargaining.* Edited by A. Roth. Cambridge University Press, 1985.

[3] S. Brams. *Negotiation Games. Applying Game Theory to Bargaining and Arbitration.* Routledge Publ, 2003.

[4] A. Muthoo. *Bargaining Theory with Applications* Cambridge University Press, 1999.

[5] J. Hirshleifer and A. Glazer. *Price theory and applications.* Cambridge Univ. Press, 2005.

[6] N. Balcan, A. Blum, J. Hartline and Y Mansour. Reducing Mechanism Design to Algorithm Design via Machine Learning. *JCSS*, 2008.

[7] V. Vapnik. *Statistical Learning Theory.* Wiley, 1998.

[8] H. Varian. Position Auctions. *International Journal of Industrial Organization*, 2006.

[9] S. Guha. On the Space-time of Optimal, Approximation and Streaming Algorithms for Synopsis Construction Problems. Journal version. `http://www.cis.upenn.edu/~sudipto/mypapers/space.pdf`