10. Online Advertising Markets

You have been asked to design an auction for selling ads on an online social network platform. You wonder what kinds of bids to allow advertisers to place, what allocation and payment rule to use, and how to position ads on a page.

In this chapter, we take a look at the design of auction-based markets for internet advertising. In contrast with traditional advertising conducted through direct mail, TV and newspaper, online ads can be chosen in real-time, and based on information about an individual and about the current context of an individual. This enables better targeting of potential customers, as well as personalized ads.

The revenue from auction-based advertising supports the abundant, diverse content that is available online, this content made possible through the lower costs of production and distribution of content that has driven the democratization of content. These ad markets, with their extensive use automated methods to choose which ads to show and how much to charge, are the result of a number of technological developments:

- *User data*: there is a lot of information available about users, including browsing behavior, physical location, context (e.g., at work, shopping, or reading the news), demographics, and interests. This information comes from multiple sources, including search engines, smartphone apps, and third-party data providers.
- *Measurable actions*: a user's response to an ad, for example whether a user clicks on an ad, the amount of time spent browsing the landing page associated with an ad, or whether an ad leads to a purchase, can be monitored and recorded.
- *Low-latency networks*: networks have low enough latency that real-time bidding, where bids to place ads are made in real time when a user visits a web page, is possible.
- *Machine learning*: advances in machine learning, including methods to train large models with many features, enable accurate estimates of quantities such as the probability that a user will click on an ad.

The flexible pricing provided by auctions is particularly useful when the value of goods is uncertain. In this application, the large, diverse and rapidly changing supply of content makes it difficult to set prices without using an auction. There is also considerable heterogeneity on the buy-side. For example, advertisers vary in regard to the kinds of users they want to reach, and the particular user contexts that are of interest.

In order to avoid latency in the user experience, auctions must be completed in hundreds of milliseconds, literally in less than the blink of an eye. In this short period of time: (1) relevant bids must be retrieved from a data management system (or placed by advertisers), (2) winners must be selected and payments calculated, and (3) the ads that corresponds to winning bids must be retrieved and displayed on a user's device.

Section 10.1 introduces three basic kinds of ads: *sponsored search ads*, such as those that appear on search engines such as Google, *contextual ads*, such as those that appear adjacent to content on platforms such as Facebook, and *display ads*, such as those that appear on the New York Times's website. Although we mainly focus on sponsored search ads, many of the lessons learned are also applicable to the design of systems for contextual and display ads.

In Section 10.2, we introduce the model of *position auctions* that provides the theoretical underpinning of the design of sponsored search auctions. In Section 10.3, we apply the VCG mechanism (see Chapter 8) to sponsored search auction design. Facebook's auction platform uses the VCG mechanism. Section 10.4 introduces the *generalized second-price* (GSP) auction. The Google and Bing search engines use the GSP auction to sell ads. In Section 10.5, we present a case study of the effect of adjusting reserve prices in the GSP auctions used by Yahoo for sponsored search. In Section 10.6, we discuss some advanced topics including the algorithmic problems of bid-pacing and yield optimization, cookie matching, and measuring the value of ads.

10.1. Types of Internet Advertising

Advertising is the dominant revenue stream for internet firms such as Google and Facebook. In this section, we review three distinct kinds of ads, each of which differs in terms of the way in which it is targeted.

10.1.1. Sponsored Search Ads

When a user enters a search query into a web search engine that is monetized through ads, an auction is used to determine which ads to display.

For example, a search for "mountain bikes" completed in Cambridge, MA in February 2013 returned the results shown in Figure 10.1. Ads appear both above and to the right of the search results. The top-right of the page also includes a "shopping panel" which is separately generated. Advertisers in a sponsored search auction can place bids on search keywords, for example a bike shop that sells mountain bikes might want to bid $1 per click for search queries that includes keywords 'professional' and 'mountain' and 'bike.' Advertisers may also target users based on location and time of day and, depending on the policy of the ad platform, use the content of emails or content on collaborative document platforms to guide targeting.

In sponsored search, the bids are placed as *standing bids* with a search engine. This means that the bids persist in the system rather than being placed in real-time, with the bids that match a user's search query retrieved by the auction infrastructure and placed into an auction to determine which ads to show and how much to charge.

In addition to keywords and the bid price, an advertiser's bid can:

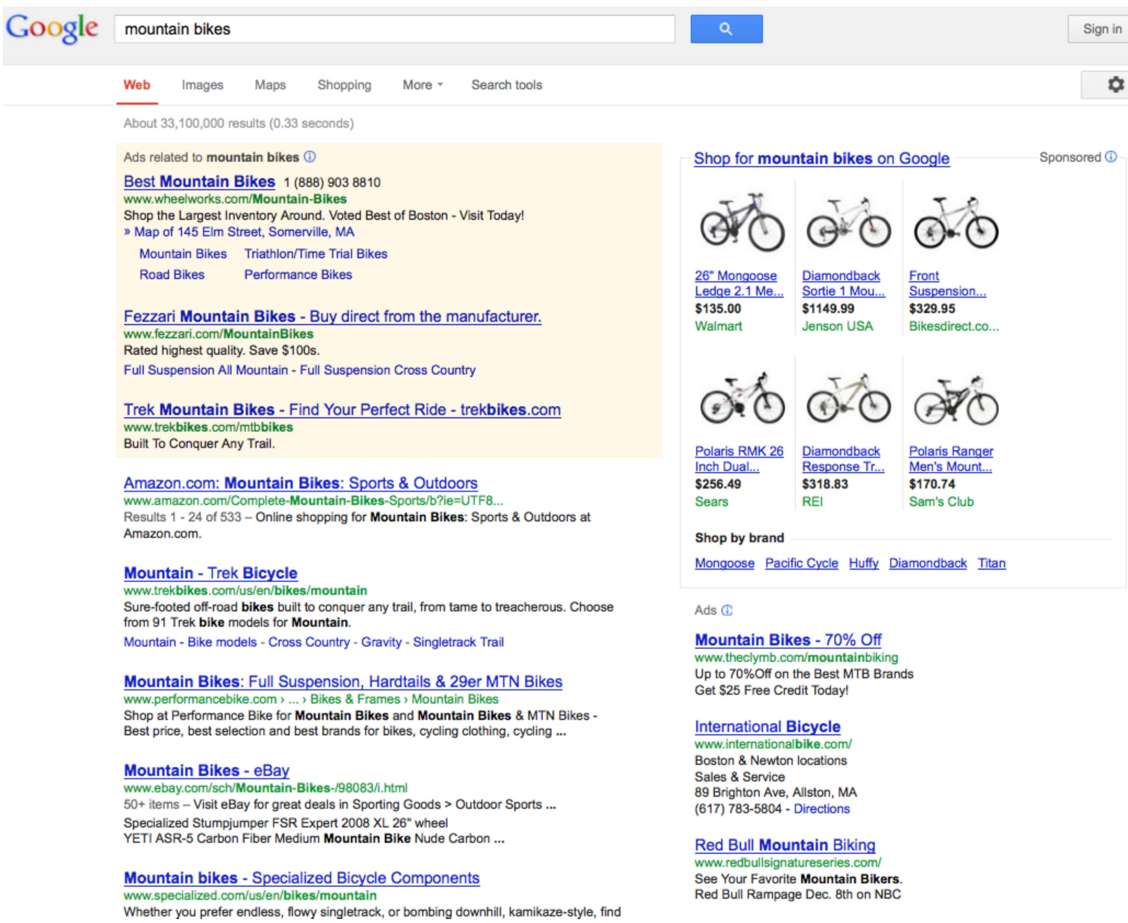


Figure 10.1.: Sponsored search ads for “mountain bikes” on the Google search engine. Ads appear above and to the right of the “organic” results. **xx copy edit: boxes around ads xx**

- Make use of *stop words* such as "injury" that specify a word that must not appear in a user's search query for the bid to match.

- Place a cap on the number of times that an ad can be shown to the same user.

Bids are typically made for the willingness-to-pay in the event of a click. Variations include bids on other user-specific actions; e.g., an add-to-cart action, a purchase, the installation of a smartphone app, or the submission of a product review. Ads may also include the price of a product, a "deep link" to a specific piece of content in a website, a map to show the location of a store or restaurant, or the ability to use VoIP to call the merchant. Figure 10.1 shows some rich ad features, including address and telephone information, links to maps, and deep links to content (e.g., 'Full Suspension All Mountain').

Bids can be coupled with a *maximum daily budget*, with the auction platform automatically reducing the frequency with which bids are matched with search queries or reducing the bid price in order to spend evenly through the day. Rather than a bid per click, this budget can also be coupled with a high-level objective such as "maximize the number of clicks on my ad" or "maximize my yield" (the value generated for every $1 spent). See Section 10.6 for an additional discussion.

Ad platforms provide summary statistics on the performance of an ad campaign, including the number of ads allocated, the number of clicks received, the average position assigned on the search results page (i.e., how close to the top of the page), and the total payments made. This provides an advertiser with information that can be used to guide experimentation in regard to bid amounts, search terms, and targeting.

10.1.2. Contextual Ads

A *contextual ad* is targeted based on the content of a webpage. Bids are typically placed on keywords in the same way as for sponsored search, with the keywords providing a way to specify the webpages for which the bid is valid. Figure 10.2 shows an example of contextual ads on a blog about historic houses. The ads for 'interior design ideas', 'Lowell Overhead Door' and 'New Visualization Tool' are associated with keywords that match the content on the page.

A *content advertising network* collects together many content pages, with each publisher including code on its web server that requests an ad, triggering an auction between relevant standing bids maintained by the ad network when a user visits a web page. Ad networks are operated by search engines, and allow an advertiser to choose to use the same bids for both sponsored search and contextual ads. On a social network platform, context can also come from a user's profile, including "likes" and interests that the platform has automatically associated with a user based on her browsing behavior. Twitter's use of "sponsored tweets" provide another example of a contextual ad. Sponsored tweets can be targeted based on the content in surrounding tweets. Contextual ads can also be social, for example an ad placed by Burberry on Facebook might mention that a particular friend "likes Burberry."

Figure 10.2.: Contextual ads on a blog about historic houses. **xx copy edit: boxes ads xx**

10.1.3. Display Ads

Display ads are images or videos and may appear as banners on web pages or as pop-ups. Display ads are typically used to build brand value or awareness of a product or service. Figure 10.3 shows display ads as they appeared on the *New York Times* web page on February 11, 2013. Display ads are sold in units of *impressions*, where an impression is one view by one person of one ad.

Display ads may be sold through contracts (e.g., 2,000,000 impressions, the month of April, front page only, \$2 per 1,000 impressions). In this case, ads must be matched with page views over a period of time in order to meet the terms of the contract. This is an interesting algorithmic problem (see the chapter notes.)

Display ads may also be sold through real-time bidding on *ad exchanges* (adXs). By placing code provided by the adX on a content server, a publisher can automatically request that the adX runs an auction whenever it serves a webpage. Upon receiving such a request, the adX makes call-outs to request bids from demand-side platforms (DSPs) that represent advertisers. The DSPs decide whether to place bids on behalf of advertisers, as well as which *creative* (ad

Figure 10.3.: Display ads for Marc Jacobs, Capital One, the New York Times, and Tiffany & Co. **xx copy edit: add boxes around ads xx**

content) is most appropriate. Publishers can restrict bidding to particular advertisers, and sell inventory anonymously in order to minimize the cannibalization of sales through other contracts.

The main advantage of real-time bidding over the standing bids used in sponsored search and contextual advertising is that real-time bidding enables *behavioral targeting*, where bids and creatives can be tailored to a user's recent browsing activity. *Re-targeting* is the process of showing an ad to a user who recently visited a website without making a purchase. Behavioral targeting is achieved through *cookie matching* (see Section 10.6.3). Ad exchanges are also used for mobile advertising as well as on social network platforms such as Facebook, enabling the combination of behavioral targeting and contextual targeting.

10.2. Position Auctions

In this section, we introduce the *position auction* for selling multiple ad positions on a page. Let $N = \{1, \ldots, n\}$ denote the set of advertisers. These are advertisers for whom targeting criteria are met; e.g., search query, user location, and time of day in sponsored search, or content and user profile in contextual advertising.

Let $M = \{1, \ldots, m\}$ denote the set of *positions* available on a page, with position 1 the best position and position m the worst position. We assume that $m = n$, with "dummy positions" or "dummy bidders" added as necessary, these positions or bidders having zero value. In a position auction, all advertisers agree on the ranking of positions. We can think of the best position as appearing at the top of the page, the worst position at the bottom.

10.2.1. Advertiser Value

The model of advertiser value makes two main assumptions. First, we assume that it is clicks that generate value for an advertiser, and that all clicks are equal-- every click on an ad has the same value to an advertiser, irrespective of which position the click comes from on the page. The position of an ad on a page only matters to the extent that it affects whether or not there will be a click.

Given this, advertiser i's value for an ad in position $j \in M$ is modeled as,

$$v_{ij} = CTR_{ij} \cdot v_i, \tag{10.1}$$

where CTR_{ij} is the *click-through rate* (CTR) of ad i in position j (the probability of receiving a click) and $v_i \geq 0$ is the advertiser's *value-per-click*. This is the expected value for position j.

Second, we assume that the effect of position on CTR is *separable* from the effect of ad quality. Let $CTR_i \in [0, 1]$ denote ad quality and $pos_j \in [0, 1]$ denote *position effect*. The ad quality is the probability an ad receives a click when it appears in position 1. The quality q_i if ad i may depend on the relevance of the ad to a user, as well as the popularity of a product. For example, the quality for an ad Trek vs Fezzari bikes may be 0.2 and 0.17 respectively, making a Trek ad more likely to be clicked.

By separability, the CTR is modeled as,

$$CTR_{ij} = CTR_i \cdot pos_j. \tag{10.2}$$

The position effect, assumed to be decreasing with position and with $pos_1 = 1$, captures the fall off in clicks as an ad is allocated a lower position.

Putting this together, an advertiser's value for position $j \in M$ is

$$v_{ij} = pos_j(CTR_i \cdot v_i). \tag{10.3}$$

This reveals the central premise of the position auction model. Although advertisers vary in regard to their value for each position they all prefer higher positions on the page than lower positions, and the relative value of two positions is the same for all advertisers (irrespective of whether an ad is high or low quality).

10.2.2. Winner Determination

Early position auctions for sponsored search received bids on impressions and ranked by bid-per-impression (higher bid-per-impression, higher position). Later, bids were on clicks and ads were ranked by bid-per-click. Today, the standard approach is to collect bids on clicks and rank bids according to a product of estimated CTR and bid value (roughly speaking, this is "rank by revenue.")

Let $eCTR_i$ denote the *estimated CTR* (or estimated quality) for ad i in position 1. This estimates comes from machine learning algorithms, and uses historical click data and features of users, advertisers, ads and user queries. Bids are ranked by the product of $eCTR_i$ and *bid-per-click* b_i (with $b_i > 0$).

Relative to earlier approaches, the advantages of this approach are:

1. *Alignment of interests*: advertisers can compete on quality, which is good for user (and thus the platform, since users will come back to the platform) because it promotes higher quality ads and ad targeting.

2. *Remove moral hazard problem*: because payments are made per click, it is in the ad platforms own best interest to accurately estimate the CTR of an ad.

3. *Promote allocative efficiency*: ads with high value per click or high CTRs receive higher positions on the page.

4. *Trust*: advertisers can observe clicks, and thus can validate the amount they are asked to pay.

Given rank-by-revenue, the bid value by advertiser i for position j is

$$b_{i,j} = pos_j(eCTR_i \cdot b_i). \tag{10.4}$$

Let $x = (x_1, \ldots, x_n)$ denote an *allocation* of ads to positions, such that $x_i \in M$ is the position assigned to ad i, and $x_i \neq x_k$ for all $i, k \in N$, $i \neq k$. Let $\hat{v}_i(x)$ $(= b_{i,x_i})$ denote the reported value of advertiser i for allocation x. The *winner determination problem* for position auctions is to allocate advertisers to positions to maximize the total reported value of the allocation, i.e. solve $\max_x \sum_{i \in N} \hat{v}_i(x)$.

Theorem 10.1. *Rank-by-revenue solves the winner determination problem.*

Proof. Assume for contradiction that allocation x is optimal, but has bids 1 and 2 with $eCTR_1 \cdot b_1 > eCTR_2 \cdot b_2$ allocated out of order. Suppose without loss of generality that $x_2 = 1$ and $x_1 = 2$. Consider an alternate allocation x' where the position of these two bids are swapped. The total reported value for the allocation to other bids is the same in x and x'. The increase in reported value to bids 1 and 2 that results from the swap is

$$eCTR_1 \cdot b_1(pos_1 - pos_2) + eCTR_2 \cdot b_2(pos_2 - pos_1) > 0, \qquad (10.5)$$

where the inequality holds because $eCTR_1 \cdot b_1 > eCTR_2 \cdot b_2$ and because $pos_1 > pos_2$. This is a contradiction with the optimality of x. □

To understand the importance of separability for the optimality of rank-by-revenue, suppose otherwise, and that an ad for Trek bikes is not very affected by being in position 1 vs 2, while an ad for Fezzari bikes does significantly less well when allocated position 2. In this case it is possible for the solution to winner determination to put Fezzari in position 1 even if Fezzari's reported value for position 1 is smaller than Trek's. Separability ensures that the relative advantage of one position over another is the same for all advertisers.

Before continuing, we mention here three practical variations on this basic approach to winner-determination:

- Because "clicky" ads may still be low quality (e.g., tricking a user into clicking through false advertising, or because the advertiser is low quality), ads may further be demoted when deemed low quality and even despite having a reasonable CTR (e.g., as measured by the centrality of a landing page on the web graph, or the reputation of a seller.)
- Reserve prices may be used, with ads restricted to those with bid-per-click above a reserve. This can have a useful effect on revenue (see Section 10.5.)
- Bids may be ranked by $(eCTR_i)^\gamma \cdot b_i$, for some *squashing factor* $\gamma \in (0, 1)$, in order to change the relative emphasis given to bid and CTR. Again, this can have a useful effect on revenue (see the chapter notes).

10.3. The VCG Position Auction

Having explained how ads are allocated to positions, the next question is how payments are determined. In this section, we define the VCG mechanism for position auctions. This provides a truthful, value-maximizing auction design.

Definition 10.1 (VCG Position Auction)**.** *Given bids-per-click* $b = (b_1, \ldots, b_n)$*, ordered* $eCTR_1 \cdot b_1 \geq eCTR_2 \cdot b_2 \geq \ldots \geq eCTR_n \cdot b_n$*:*

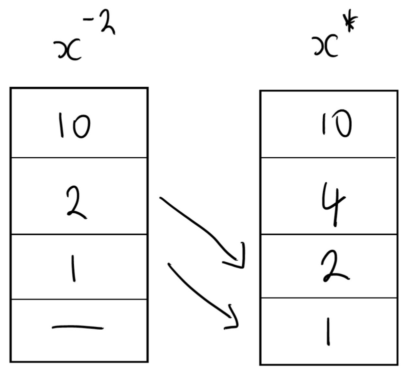


Figure 10.4.: The effect of a bid from an advertiser with bid-per-click \$4 in the VCG position auction.

- *The allocation rule selects allocation x^* that assigns ad 1 to position 1, ad 2 to position 2, $\cdots$, and ad n to position n.*
- *The payment rule charges advertiser i an amount*

$$t_{\mathrm{vcg},i}(b) = \sum_{k \neq i} \hat{v}_k(x^{-i}) - \sum_{k \neq i} \hat{v}_k(x^*) = \sum_{k=i+1}^{n} (pos_{k-1} - pos_k) eCTR_k \cdot b_k, \tag{10.6}$$

where x^{-i} is the allocation that would be made without advertiser i.

In words, ads are allocated to maximize total reported value, and the payment rule charges an advertiser the negative externality on others that results from her bid. This arises from pushing ads with lower rank down by one position, which has the effect of reducing their clicks. See Figure 10.4. The advertiser in position 2 with bid-per-click \$4 pushes the advertiser with bid \$2 down from position 2 to 3, and the advertiser with bid \$1 down from position 3 to 4. Expression (10.6) measures this total effect on all ads that are lower ranked than ad i (i.e., ads $i+1$ through n). The payment for an advertiser allocated a dummy position i is zero, because $pos_k = 0$ for all $k \geq i$.

Theorem 10.2. *The VCG position auction is truthful and efficient when the estimated CTRs are accurate.*

This follows immediately from the standard properties of the VCG mechanism (Theorem 8.6) and Theorem 10.1. It means that truthful bidding of per-click values is a dominant strategy, and that the auction will allocate positions to maximize the total expected value of advertisers.

To convert payment $t_{\mathrm{vcg},i}(b)$ into an equivalent *price-per-click* $p_{\mathrm{vcg},i}(b)$, we require

$$t_{\mathrm{vcg},i}(b) = (eCTR_i \cdot pos_i) p_{\mathrm{vcg},i}(b). \tag{10.7}$$

Rearranging, we set $p_{\text{vcg},i}(b) = t_{\text{vcg},i}(b)/(eCTR_i \cdot pos_i)$, and zero for ads in dummy positions. As long as the estimated CTRs are accurate this does not change the incentive properties, and the VCG position auction with per-click-prices is truthful. When the ads all have the same quality, this simplifies to

$$p_{\text{vcg},i}(b) = \frac{1}{pos_i} \sum_{k=i+1}^{n} (pos_{k-1} - pos_k) b_k, \tag{10.8}$$

and zero for ads in dummy positions.

Example 10.1. *Suppose there are four positions (position effects* 0.2*,* 0.18*,* 0.1 *and* 0*, respectively), and four advertisers (per-click values* $10, 4, 2$ *and* 1*, respectively, and all assumed to have the same quality). If truthful, the ads are assigned to positions* $1, 2, 3$ *and* 4*, respectively. For advertiser* 1*,* $p_{\text{vcg},1}(b) = (1/pos_1)(pos_1 - pos_2)b_2 + (pos_2 - pos_3)b_3 + (pos_3 - pos_4)b_4) = (1/0.2)((0.2 - 0.18)4 + (0.18 - 0.1)2 + (0.1 - 0)1) = 0.34/0.2 = 17/10$*. For advertiser* 2*,* $p_{\text{vcg},2}(b) = (1/pos_2)(pos_2 - pos_3)b_3 + (pos_3 - pos_4)b_4) = (1/0.18)((0.18 - 0.1)2 + (0.1 - 0)1) = 0.26/0.18 = 13/9$*. For advertiser* 3*,* $p_{\text{vcg},3}(b) = (1/pos_3)(pos_3 - pos_4)b_4 = (1/0.1)(0.1 - 0) = 0.1/0.1 = 1$*. This price is equal to the bid of advertiser* 4*. Advertiser* 4*'s price is zero.*

Exercise 10.1 confirms that no advertiser has a useful deviation in Example 10.1.

10.4. The Generalized Second-Price Auction

The GSP auction was introduced as a response to problems with bid cycling in earlier designs, which charged an advertiser the amount of her bid. As discussed in Chapter 1, the pay-your-bid rule resulted in bidding wars and instability. Continually changing bids placed a high load on servers, imposed a cost on advertisers, and resulted in allocative inefficiencies.

In defining the GSP auction, it is convenient to let $b_{n+1} = 0$.

Definition 10.2 (GSP Auction)**.** *Given bids-per-click* $b = (b_1, \ldots, b_n)$*, ordered* $eCTR_1 \cdot b_1 \geq eCTR_2 \cdot b_2 \geq \ldots \geq eCTR_n \cdot b_n$*:*

- *The allocation rule selects allocation* x^* *that assigns ad 1 to position 1, ad 2 to position 2, …, and ad* n *to position* n*.*
- *The payment rule changes advertiser* i *a price-per-click of*

$$p_{\text{gsp},i}(b) = \frac{eCTR_{i+1} \cdot b_{i+1}}{eCTR_i}. \tag{10.9}$$

In words, an advertiser pays the minimal bid amount that she could bid and still be assigned the same position, i.e. $eCTR_i \cdot p_{\text{gsp},i}(b) = eCTR_{i+1} \cdot b_{i+1}$. When the ads all have the same quality, this simplifies to $p_{\text{gsp},i}(b) = b_{i+1}$. The price-per-click does not matter for advertisers allocated to a dummy position because they receive no clicks in any case.

Table 10.1 summarizes the VCG and GSP prices in Example 10.1. Example 10.2 illustrates the GSP prices in the case that advertisers differ in quality.

position	position effect	value-per-click	bid-per-click	VCG price	GSP price
1	0.2	10	10	17/10	4
2	0.18	4	4	13/9	2
3	0.1	2	2	1	1
4	0	1	1	0	0

Table 10.1.: Comparing VCG and GSP prices in an example where all bids have the same quality.

Example 10.2. *Consider an auction with three positions (position effects* $0.2, 0.1,$ *and* 0*) and three advertisers (bids-per-click* 8, 10 *and* 2, *and estimated quality* 1, 0.5 *and* 0.5, *respectively). Given truthful bids, we have* $eCTR_1 \cdot b_1 = 8 > eCTR_2 \cdot b_2 = 5 > eCTR_3 \cdot b_3 = 1$, *and positions* 1, 2 *and* 3 *go to advertisers* 1, 2 *and* 3 *respectively (position* 3 *is the dummy position). We have* $p_{\text{gsp},1}(b) = eCTR_2 \cdot b_2 / eCTR_1 = ((0.5)(10))/1 = 5$. *This is the minimum bid such that the advertiser would retain position 1. Similarly,* $p_{\text{gsp},2}(b) = eCTR_3 \cdot b_3 / eCTR_2 = ((0.5)(2))/0.5 = 2$. $p_{\text{gsp},3}(b) = 0$.

By modifying the first-price payment rule of the early sponsored search auction designs, the GSP auction removes the need to try to bid just above the advertiser in the next position, and has the effect of making the auction more stable.

The VCG and GSP payment rules are identical when selling one or two positions, but differ when selling three or more positions. In working through the effect of this on strategic behavior in GSP, we will make the simplifying assumption that all ads have same quality. The analysis can be generalized to handle quality differences, with the main results qualitatively unchanged.

Theorem 10.3. *The GSP auction is not truthful.*

Proof. Consider the example in Table 10.1. For simplicity, assume $CTR_i = 1$ for all advertisers. Given this, the expected utility for a truthful bid by the advertiser with value \$10 is

$$pos_1(v_1 - p_{\text{gsp},1}(b)) = 0.2(10 - 4) = 1.2. \tag{10.10}$$

By deviating and bidding $b'_1 = 3$, she would be allocated position 2 instead at price $p_{\text{gsp},1}(b'_1, b_{-1}) = 1$, with expected utility

$$pos_2(v_1 - p_{\text{gsp},1}(b'_1, b_{-1})) = 0.18(10 - 2) > 1.2 \tag{10.11}$$

This is a useful deviation, and truthful bidding is not a dominant strategy. □

The GSP auction charges an advertiser as though the effect of her bid is to prevent the advertiser with the next highest bid from being allocated any position at all. In comparison, the VCG position auction's payment rule recognizes that the effect of an advertiser is more moderate, and is to push each lower bid down by one position. For this reason, the GSP auction charges too much for advertisers allocated high positions.

Because advertisers are given feedback by auction platforms about the position they are allocated and the prices of different positions, it is reasonable to study the Nash equilibria of the GSP auction. Let $u_{\text{gsp},i}(b)$ denote the expected utility to advertiser i given bid profile $b = (b_1, \ldots, b_n)$. The expected utility is $u_{\text{gsp},i}(b) = pos_j \cdot CTR_i(v_i - b_{i+1})$, where for advertiser n we assume $b_{n+1} = 0$.

As is standard, bid profile b^* is a Nash equilibrium if no advertiser i has a useful deviation,

$$u_{\text{gsp},i}(b_i^*, b_{-i}^*) \geq u_{\text{gsp},i}(b_i', b_{-i}^*), \quad \forall\, b_i' \neq b_i^*,\ \forall i. \tag{10.12}$$

It is convenient to assume that equilibrium bids are indexed so that $b_1^* \geq b_2^* \geq \ldots b_n^*$, with advertiser i allocated position i. For advertisers in positions below position 1 in this equilibrium, no useful deviation requires:

(P1) *No strictly-improving higher bid*:

$$pos_{i-k}(v_i - b_{i-k}^*) \leq pos_i(v_i - b_{i+1}^*), \quad \forall i \in \{2, \ldots, n\},\ \forall k \in \{1, \ldots, i-1\} \tag{10.13}$$

Condition (P1) ensures that bidding a higher value than b_i^* cannot help. The value on the left-hand side is the expected utility for a bid that moves k positions higher. (P1) precludes an advertiser in a position k with $pos_k > 0$ from having a bid value less than the value of an advertiser allocated to a dummy position (with zero position effect).

For advertisers in positions above position n in this equilibrium, no useful deviation further requires:

(P2) *No strictly-improving lower bid*:

$$pos_{k+i}(v_i - b_{k+i+1}^*) \leq pos_i(v_i - b_{i+1}^*), \quad \forall i \in \{1, \ldots, n-1\},\ \forall k \in \{1, \ldots, n-i\} \tag{10.14}$$

Condition (P2) ensures that bidding a lower value than b_i^* cannot help. The value on the left-hand side is the expected utility for a bid that moves k positions lower. (P2) insists on non-negative utility (and thus prices less than or equal to value), since position n is always available at price zero.

Example 10.3. *Returning to the instance from Example 10.1, Table 10.2 provides a Nash equilibrium. We can check for possible deviations. For the advertiser with value \$10, for example, we have* $0.18(10-2) = 1.44 \geq 0.2(10-4) = 1.2$ *and* $0.18(10-2) = 1.44 \geq 0.1(10-1) = 0.9$*, and there is no useful higher or lower deviation. See Figure 10.5.*

We see from Example 10.3 that a Nash equilibrium of GSP need not be efficient.

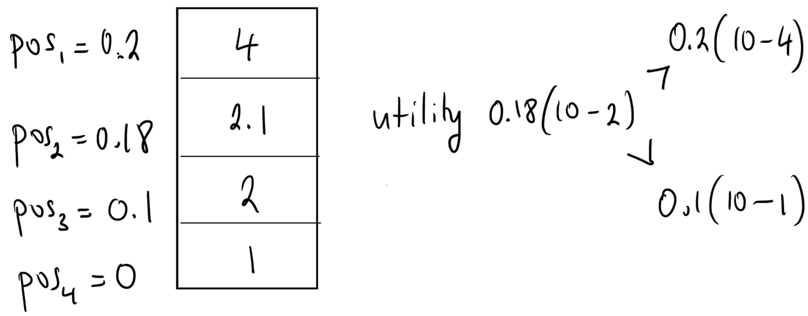


Figure 10.5.: A best response analysis in the GSP auction from the perspective of an advertiser with value $10 per click in position 2.

position	position effect	value-per-click	bid-per-click	GSP price
1	0.2	4	4	2.1
2	0.18	10	2.1	2
3	0.1	2	2	1
4	0	1	1	0

Table 10.2.: A (non value-ordered) Nash equilibrium of the GSP auction.

position	position effect	value-per-click	bid-per-click	GSP price
1	0.2	10	10	b_2
2	0.18	4	b_2	13/9
3	0.1	2	13/9	1
4	0	1	1	0

Table 10.3.: A partially-instantiated EFNE of the GSP auction.

10.4.1. Envy-Free Nash Equilibria

Let us suppose that the advertisers in Example 10.1 receive feedback that position 1 is priced at \$2.1 per click, position 2 at \$2 per click, position 3 at \$1 per click, and position 4 at \$0 per click. Although these are the prices in a Nash equilibrium, the advertiser with value \$10 would "envy" the advertiser in position 1. This is because the expected utility from position 1 at price 2.1 would be $0.2(10 - 2.1) = 1.58 > 0.18(10 - 2) = 1.44$. In actuality, this price of 2.1 is not available to this advertiser, and reflects the effect of her own bid. Rather, position 1 is only available for a price of \$4, and it is this price that played a role in the Nash equilibrium analysis.

Because price feedback reveals the price in each position rather than the precise bids of others we introduce this symmetry requirement as an additional constraint, and understand the properties of Nash equilibria that in addition generate prices that no advertiser envies. While the prices to other advertisers for higher positions may be lower than those available under a unilateral deviation, the prices to other advertisers for lower positions are the same as those that are assumed in a Nash equilibrium analysis.

A bid profile is *envy-free* if no advertiser prefers the position and price of any other advertiser.

Definition 10.3 (Envy-free Nash equilibrium)**.** *An envy-free Nash equilibrium (EFNE) is a bid profile that is a Nash equilibrium and envy-free.*

There is empirical evidence that advertisers play EFNE in sponsored-search auctions (see the chapter notes). An interesting property of any EFNE is that the bids are *value-ordered*, meaning they preserve the same order as the values of advertisers. Because of this, every EFNE is efficient. Moreover, the revenue in any EFNE is at least as much as the revenue in the VCG mechanism. See Exercise 10.1.

Not only do EFNE exist, but there are a large number of these equilibria. Example 10.4 illustrates *balanced bidding*, which will give attention to a particular EFNE.

Example 10.4. *Returning to the instance from Example 10.1, Table 10.3 gives a partially instantiated EFNE. Consider the bid b_2 of the advertiser with value \$4. A bid $b_2 \in (13/9, 10)$ is a best response because the advertiser does not want position 1 at price* \$10, *and the expected utility for position 2 is* $0.18(4-13/9) = 0.46 > 0.1(4-1) = 0.3$, *which is the expected utility for bidding to take position* 3. *But how might the advertiser select a bid b_2 within range* $(13/9, 10)$*?*

A first idea is to bid just below \$10 *because this will make the advertiser in position* 1 *pay a large amount, possibly depleting her budget and driving her out of the auction. But the advertiser in position* 1 *has a possible retaliation, which is to respond by bidding just below* b_2 *(say some bid* $b_1 = b_2 - \epsilon$*, for* $\epsilon > 0$*), which would force the advertiser with value* \$4 *into position* 1 *at a price of* $b_2 - \epsilon$*.*

In balanced bidding, the advertiser with value \$4 would bid as high as possible amongst those bids for which there is no retaliation threat by the next highest advertiser. In this example, the advertiser would select the bid to satisfy

$$0.18(4 - 13/9) = 0.2(4 - b_2), \tag{10.15}$$

which is $b_2 = 17/10$*. At this bid price, the advertiser is exactly indifferent between remaining in position 2 or being forced into position 1.*

Comparing with Table 10.1, we see something interesting: the outcome of the GSP auction in this equilibrium is the same as the truthful outcome of the VCG mechanism.

Definition 10.4 (Balanced bidding)**.** *A bid profile* b *in the GSP auction, ordered so that* $b_1 \geq b_2 \geq \ldots b_n$*, satisfies the balanced bidding property if*

$$pos_{i-1}(v_i - b_i) = pos_i(v_i - b_{i+1}), \tag{10.16}$$

for all advertisers $i \in \{2, \ldots, n\}$*, with* $b_i = v_i$ *for any advertiser allocated a dummy position.*

Balanced bidding is motivated by the dynamic considerations suggested in Example 10.4 of bidding up to increase the price of others but not so much so as to open up an advertiser to retaliation.

When added to the requirements of Nash equilibrium, balanced bidding ensures that an outcome is envy-free. It removes envy for the immediate next higher position, for example, since the point of indifference to retaliation is exactly that at which an advertiser is indifferent between her outcome and that of the next higher advertiser.

We refer to such an equilibrium as a *balanced EFNE*. Because bids will be value-ordered (as they will be in any EFNE), a balanced EFNE can be constructed by setting the bids for advertisers in order of increasing value, working from the bottom position to the top.

Example 10.5. *The bids in Table 10.3 are constructed by repeatedly appealing to the condition for balanced bidding, working from advertiser 4 with value \$1 to advertiser 1 with value \$10. First, set bid* $b_4 = v_4 = 1$ *because this advertiser is allocated a dummy position. Calculate bid* b_3 *to satisfy balanced bidding and make the advertiser indifferent between position* 3 *at \$1 and position* 2 *at* b_3*. For this, we need* $0.1(2 - 1) = 0.1 = 0.18(2 - b_3)$*, and so* $b_3 = 13/9$*. As we have already seen, for* b_2 *we need* $0.18(4 - 13/9) = 0.2(4 - b_2)$*, and so* $b_2 = 17/10$*. For advertiser 1, we set* $b_1 = v_1 = 10$*.*

It is instructive to examine why a Nash equilibrium comes about simply from insisting on value ordering and balanced bidding. See Figure 10.6, which defines balanced-bidding

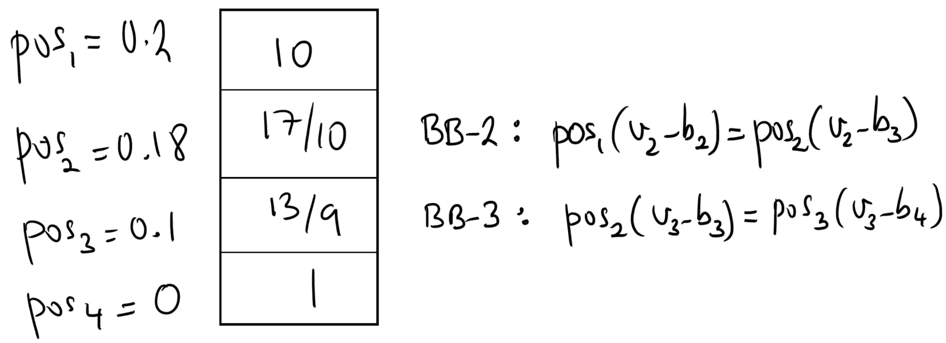


Figure 10.6.: A balanced EFNE of the GSP auction, showing the balanced-bidding requirement for the advertisers in positions 2 and 3.

requirements BB-2 and BB-3 in the running example. We first establish Nash equilibrium property (P2) for advertiser 1. For this, we have

$$pos_1(v_1 \underbrace{-b_2) \quad \geq \quad pos_2(v_1}_{\text{BB-2, } v_1 \geq v_2,\ pos_1 \geq pos_2} - \underbrace{b_3) \quad \geq \quad pos_3(v_1}_{\text{BB-3, } v_1 \geq v_3,\ pos_2 \geq pos_3} - b_4), \tag{10.17}$$

which precludes a useful deviation to position 2 or 3. Let us also establish Nash equilibrium property (P1) for advertiser 3. For this, we have

$$pos_2(v_3 \underbrace{-b_2) \leq pos_2(v_3}_{b_2 \geq b_3} - \underbrace{b_3) = pos_3(v_3}_{\text{BB-3}} - b_4), \tag{10.18}$$

which is the Nash equilibrium requirement in regard to deviating to position 2. We also have,

$$pos_1(v_3 - \underbrace{b_1) \leq pos_1(v_3}_{b_1 \geq b_2} - \underbrace{b_2) \quad \leq \quad pos_2(v_3}_{\text{BB-2, } v_3 \leq v_2,\ pos_1 \geq pos_2} - \underbrace{b_3) = pos_3(v_3}_{\text{BB-3}} - b_4), \tag{10.19}$$

which precludes a useful deviation to position 1.

Theorem 10.4. *Any Nash equilibrium of the GSP auction that satisfies balanced bidding is envy-free, and the outcome in this balanced EFNE is equal to the truthful outcome of the VCG mechanism (assuming estimated CTRs are accurate).*

Proof. We focus here on the proof that payments in the VCG mechanism are equal to those in GSP in the balanced EFNE. The rest of the proof is deferred to Exercise 10.1. We assume for simplicity that that advertiser quality is $CTR_i = 1$ for all advertisers. The proof can be generalized to handle bids with different qualities. We adopt a recursive expression for the VCG payment: $t_{\text{vcg},n}(b) = 0$ and $t_{\text{vcg},i}(b) = (pos_i - pos_{i+1})b_{i+1} + t_{\text{vcg},i+1}(b)$ for advertisers i with $i < n$. Let $t_{\text{gsp},i}(b)$ denote the expected payment of advertiser i in the GSP auction. Let b^* denote the balanced EFNE of the GSP auction.

VCG	GSP	Second price
Twitter, Facebook (including Facebook Exchange)	Google, Bing	Ad exchanges (AppNexus, Twitter MoPub, Google Doubleclick)

Figure 10.7.: The auction designs used by major internet advertising firms.

We prove the equivalence of payments by induction on the identity of the advertiser. The base case for advertiser n is trivial, with $t_{\text{vcg},n}(v) = t_{\text{gsp},n}(b^*) = 0$. For advertiser $i < n$, given the inductive hypothesis that $t_{\text{gsp},i+1}(b^*) = t_{\text{vcg},i+1}(v)$, we have,

$$t_{\text{gsp},i}(b^*) = pos_i \cdot b^*_{i+1} = (pos_i - pos_{i+1})v_{i+1} + pos_{i+1} \cdot b^*_{i+2} \tag{10.20}$$
$$= (pos_i - pos_{i+1})v_{i+1} + t_{\text{gsp},i+1}(b^*) \tag{10.21}$$
$$= (pos_i - pos_{i+1})v_{i+1} + t_{\text{vcg},i+1}(v) = t_{\text{vcg},i}(v), \tag{10.22}$$

where (10.20) follows from (10.16) instantiated to advertiser $i + 1$, (10.21) is by definition, and (10.22) follows from the inductive hypothesis and the recursive definition of VCG payments. □

We have observed this equivalence between the truthful VCG outcome and the balanced EFNE of the GSP auction through the prices in Table 10.3 (with $b_2 = 17/10$) and the prices in Table 10.1.

10.4.2. Design Tradeoffs

Figure 10.7 provides a summary of the different auction designs that are used for internet ad platforms. Although the outcomes of the GSP and VCG position auctions agree in the balanced EFNE of GSP, they differ in other ways. We can list off some reasons as to why the GSP auction was first introduced and is still used for sponsored search and contextual advertising:

- By accident! Some have suggested that the intent of engineers was to deploy the VCG mechanism and designers did not realize that GSP is not VCG.
- Simplicity. The GSP payment rule is easier to explain than the VCG payment rule.
- Economic properties. GSP has the same outcome in the balanced EFNE as the truthful outcome of VCG mechanism.
- Short-term revenue. Fixing the bids, the revenue from GSP is at least that of VCG. A switch from GSP to VCG would have a negative impact on revenue in the short-term and an uncertain effect in the long-term.

- Lower information requirement. The price-per-click in GSP can be computed without knowing the precise CTRs. Rather, it is sufficient to know the quality of each ad (10.9). In comparison, calculating the prices in VCG requires knowledge of the CTRs (10.6).

- Re-engineering cost. Search engines and third party ad-tech platforms have built up considerable infrastructure around GSP. Switching to VCG would cause disruption for example to the various algorithmic systems used to optimize campaign goals of advertisers.

Some reasons to adopt the VCG mechanism (as done for Facebook's contextual advertising platform) or to switch from GSP to using VCG include:

- Flexibility. The VCG mechanism can be used to jointly optimize the ads and organic content shown to a user. This can be done by inferring a user's value for different kinds of content and ads from behavioral patterns, using this to maximize the user value subject to constraints on the number of ads that are displayed. In addition, the VCG can easily accommodate different advertiser and platform goals. For example, the VCG mechanism can be used to accommodate bids such as "my value is \$1 per click for positions 1 and 2, and \$0.80 per click for positions 3 and 4," or to handle bids that each specify values for different amounts of ad space.

- Enable value estimation. Truthful bidding is an equilibrium in VCG, making it easier to understand advertiser values for different user contexts. This information is useful when testing designs in simulation and when setting reserve prices (see Section 10.5.)

- Faster experimentation. Whereas part of the effect of changing a reserve price on the revenue in GSP is indirect, arising through the effect that this has on bids in equilibrium, the truthful equilibrium in VCG is robust to changing these kinds of design parameters. This speeds up experimentation of the effect of design parameters such as revenue for the auction platform.

The single-item, second-price auction is used by most adXs because they sell a single impression at a time, and with pricing per-impression and not per-click because they may not have the infrastructure to be able to track and charge per-click. Rather, the focus is on clearing inventory at low latency and high volume.

10.5. Case-Study: Reserve prices at Yahoo

Yahoo conducted a large experiment in 2008 into setting reserve prices in the GSP auctions used for sponsored search. The experiment involved auctions on around 450,000 distinct keywords. Around 430,000 were randomly placed into a treatment group, and received optimized reserve prices, while the remaining 20,000 were used as a control and retained the default reserve of 10¢ per click.

number of advertisers	reserve 0	reserve 0.5	revenue improvement
single position			
$n = 2$	0.3333	0.4166	25%
$n = 6$	0.7143	0.7165	0.31%
six positions			
$n = 2$	0.1	0.475	375%
$n = 6$	0.8123	1.1764	45%

Figure 10.8.: The theoretical impact of reserve price on expected revenue in an auction with a single position, and an auction with six positions, comparing $n = 2$ and $n = 6$ advertisers.

10.5.1. Theoretical Effect of Reserve Prices

Leaving aside advertiser quality, the GSP position auction with *reserve price* $r \geq 0$ on clicks works as follows:

- Receive bids-per-click, $b = (b_1, \ldots, b_n)$.
- Allocate a position to bids $b_i \geq r$, in decreasing order of bid price.
- Charge an allocated advertiser i a price-per-click $p_{\text{gsp},i}(b) = \max(b_{i+1}, r)$, where b_{i+1} is the bid price of the next highest bid.

The effect on the balanced EFNE is to set a floor on bids, so that advertisers with $v_i \geq r$ will bid $b_i \geq r$ in equilibrium. With this modification, the balanced EFNE can still be constructed by working from low value to high value, setting bid values to satisfy the balanced-bidding condition. The outcome of the GSP with reserve in the balanced EFNE is equal to that of the VCG position auction with reserve. See Exercise 10.2.

Because an advertiser's private information is a single number (the value-per-click), the domain behaves similarly to that of a single-item auction for the purpose of maximizing expected revenue. Factoring the effect of CTR, we can model values-per-impression for the top position as sampled IID from a distribution with cumulative distribution function F and density function f. The optimal auction ranks bids by decreasing value and adopts a reserve price of $r = \phi^{-1}(0)$, where $\phi(v_i) = v_i - (1 - F(v_i))/f(v_i)$ is the virtual value function (see Chapter 7). This assumes that distribution F is regular.

Let us assume the per-impression values for the top position are distributed *uniform*$(0, 1)$. Given this, the virtual value $\phi(v_i) = 2v_i - 1$, and the optimal reserve price is $r = \phi^{-1}(0) = 0.5$. We assume that the position effect depends on a discounting parameter δ, with $0 < \delta < 1$, so that $pos_1 = 1, pos_2 = \delta, pos_3 = \delta^2$, and so on. We assume $\delta = 0.7$.

Figure 10.8 illustrates the theoretical effect on revenue of adopting an optimal reserve. The expected revenue can be computed as the *expected total virtual value*, where the effect of the reserve is that only bids with non-negative virtual value are allocated. This follows from Theorem 7.3, which also holds for position auctions, since they are single-parameter auctions.

To understand the analysis, consider an auction with two advertisers. With a single position for sale, the expected revenue without a reserve is the expected value of the maximum statistic on two samples uniformly distributed on $[-1, 1]$, which is $(2/3)(2) - 1 = 1/3$, noting that the maximum virtual value is $2(1) - 1 = 1$ and the minimum virtual value is $2(0) - 1 = -1$. With the optimal reserve, there are three cases to consider:

		expected virtual value
probability 1/4:	both virtual values < 0	0
probability 1/2:	one virtual value above 0, one below 0	1/2
probability 1/4:	both virtual values ≥ 0	2/3

In the first case, both values are below the reserve, no ad is allocated, and the realized virtual value is zero. In the second case, one ad is above the reserve, and the expected virtual value is 1/2 (since the virtual value of this ad is $U(0, 1)$). For the third case, both ads are above the reserve, and the expected virtual value is the expected maximum statistic on two samples uniformly distributed on $[0, 1]$, and thus 2/3. Combining, the total expected virtual value is $1/4(0) + 1/2(1/2) + 1/4(2/3) = 5/12 \approx 0.4166$, and this is the expected revenue. With six positions for sale, the expected revenue with two advertisers and without a reserve is the expected value of maximum statistic on two samples uniformly distributed on $[-1, 1]$ plus δ times the expected value of the 2nd-highest statistic. This is $((2/3)(2) - 1) + \delta((1/3)(2) - 1) = 1/3 - \delta/3 = 0.1$. for discounting factor $\delta = 0.7$. With the optimal reserve, there are three cases to consider:

		expected virtual value
probability 1/4:	both virtual values < 0	0
probability 1/2:	one virtual value above 0, one below 0	1/2
probability 1/4:	both virtual values ≥ 0	9/10

The analysis for the first and second cases is unchanged. For the third case, the advertiser with the maximum value receives position one, while the advertiser with the second-highest value receives position two. Conditioned on virtual values being uniform on $[0, 1]$, the expected virtual value of the first advertiser is 2/3, while the expected virtual value of the second advertiser is δ times 1/3. Combining, the total expected virtual value is $1/4(0)+1/2(1/2)+1/4(2/3+\delta(1/3)) \approx 0.475$, and this is the expected revenue.

With six positions for sale, the reserve price has a substantial effect on expected revenue. Even with six advertisers, the effect is to increase revenue by 45%. There is less competition with six positions for sale than one, and the reserve has more effect. This is meaningful, because the median number of advertisers competing on a keyword at the time of the Yahoo experiment was 5.5.

10.5.2. Experimental Results

The distribution on values of advertisers was needed in order to calculate the virtual value and set optimal reserve prices. The values for any particular keyword were assumed to distributed according to a log-normal distribution. Given this, the problem was to estimate the mean

Percentile	10%	25%	Median	75%	90%
Optimal reserve	9¢	12¢	20¢	37¢	72¢

Figure 10.9.: Statistics for the distribution of estimated, optimal reserve prices on the 430,000 keyword auctions in the treatment group.

and standard-deviation parameters and the number of advertisers (assumed constant). For this, the parameters and number of advertisers were initialized to some values, and then the following three steps were iterated until an optimal fit of parameters was identified:

1. For the current estimate of the mean, standard-deviation, and number of advertisers, sample value profiles, and for each value profile compute the balanced EFNE, given a reserve price of 10¢ per click (the typical, historical reserve).

2. Given these samples of balanced EFNE, compare the average number of advertisers allocated a position, the average bid in position two and below, and the standard deviation of bid price in position two and below with observed bid data on the keyword.

3. Adjust the mean, standard-deviation and number of advertisers in a direction that improves the match between the generated statistics and the statistics in the observed bid data, and go back to step (1).

This was repeated for every keyword. This estimation approach makes two simplifying assumptions: it ignored advertiser-specific quality factors, and used the same estimated position effect for all keywords. Given the estimated distribution on values, the optimal reserve prices were computed, looking for the value at which the virtual value would be zero. Figure 10.9 presents a summary of these reserves for the keywords in the treatment group. The estimated, optimal reserve price was higher than the historical reserve price of 10¢ for almost 90% of the keywords.

Because Yahoo's management was concerned about changing the reserve prices by too much too quickly, the reserve price that was adopted for each keyword in the treatment group was a linear combination of the estimated, optimal reserve price and the historical reserve, with a weight of between 0.4 and 0.6 given to the optimal reserve. This was justified in part by a simulation, in which the effect of reserve price on revenue was relatively flat from around 1/2 of the optimal reserve price upwards.

A conservative estimate is that the adoption of these new reserve prices improved revenue by at least 2.7% during the experiment. If extended to the entire market this would have increased revenue by hundreds of millions (US$) a year. From this perspective, the experiment was viewed as a success. One difficulty in analyzing the data was that the number of searches per keyword proved quite volatile, for example because of the effect of events such as Mother's day, Father's day, the NFL season, the Beijing Olympics, and the start of the new school year. This made it hard to isolate the effect of changing reserve prices from other changes occurring in the broader market.

The effect on revenue was not uniform across keywords. In fact, the effect was negative on keywords with less than ten searches a day (representing almost 80% of the keywords, but only 27% of the total search volume, and generating only 19% of the revenue), but positive otherwise. Overall, the positive effect on high-volume keywords outweighed the negative effect on low-volume keywords. The effect on revenue was especially positive for keywords with a high optimal reserve and a small number of advertisers; e.g., there was a 9% improvement in revenue on keywords for which the the optimal reserve price was 20¢ or greater.

10.6. Advanced Topics

In this section, we discuss some of the algorithmic challenges facing advertisers in internet advertising markets. These problems relate to budget effects and yield optimization and broaden out from our treatment of a single auction to thinking about the broader ad market. We also provide some more detail about cookie matching, which is used to allow re-targeting in real-time adXs. In closing, we will also return to the perennial question of how to measure the effectiveness of advertising.

10.6.1. Bid Pacing to a Single Audience

Let us first suppose that an advertiser has decided to spend a daily budget of $B > 0$ on an "audience", for example an audience on Facebook comprised of people who live in Boston, who like Hillary Clinton, and who are aged 20-25. *Bid pacing* is the problem of deciding how much to bid during the course of a day, seeking to spend B while maximizing profit. To keep things simple, we assume that the auctions are each for a single position (and thus single-item, second-price auctions). Furthermore, we assume that an advertiser has no information that suggests that the supply of users or competition with other advertisers varies in a systematic way during the day. Let $S > 0$ denote the total supply of user impressions, and thus auctions during the day.

To determine the optimal bid, an advertiser needs an expression for the expected payment that she will make for different bid values. This is because it is important to understand the effect on budget. Let z denote the maximum bid placed by other advertisers. Let G denote the distribution function on this maximum bid, and assume for simplicity that the distribution is the same throughout the day.

Given this, the expected payment $t_i(b_i)$ for bid b_i is

$$t_i(b_i) = G(b_i) \times \mathbb{E}_z[z \mid z \leq b_i] CTR_i^{\text{win}}, \tag{10.23}$$

where CTR_i^{win} is the CTR conditioned on winning the auction. This is the product of the probability the bid wins and the expected value of the highest outside bid conditioned on winning and the CTR conditioned on winning. The expected total amount spent in a day is

$$\min(B, t_i(b_i) \cdot S). \tag{10.24}$$

We assume $t_i(v_i) \cdot S > B$, so that the advertiser would expect to hit her budget constraint before the end of the day if bidding truthfully.

One possible bid pacing strategy is "bid high, bid sometimes":

Strategy 1: Bid at value-per-click v_i, with probability $B/(t_i(v_i) \cdot S)$.

A second bid pacing strategy is "bid low, bid always":

Strategy 2: Bid at amount-per-click $b_i < v_i$ in every auction, where $t_i(b_i) \cdot S = B$.

To make these approaches robust, the parameters (bid probability or bid amount) can be adjusted during the day based on realized spend, with the remaining budget and remaining supply adopted in place of B and S, respectively.

Because the expected amount spent is the expected number of clicks multiplied by the expected price-per-click, the expected number of clicks purchased Q under each strategy is:

$$Q(\text{strategy 1}) = \frac{B}{\mathbb{E}_z[z \mid z \leq v_i]} < \frac{B}{\mathbb{E}_z[z \mid z \leq b_i]} = Q(\text{strategy 2}), \tag{10.25}$$

and strategy 2 purchases more clicks. Thus, under the assumption every click brings the same value v_i, irrespective of the price at which it is purchased, then the "bid low, bid always" strategy dominates that of the "bid high, bid sometimes" strategy.

Note, though, that while the assumption that every click brings the same value is reasonable for a single auction, it may be less reasonable for auctions that are in effect "selected" through different bidding strategies. Consider Figure 10.10, which illustrates the effect of competition between two advertisers i and j. Suppose that advertiser j targets some of the same audience as advertiser i, and that bid i under-cuts j, so that $b_i < b_j$. Suppose advertiser i wins some of the auctions that she targets that j does not target, while losing the auctions that they both target. If the auctions that j targets tend to have higher value-per-click to advertiser i, then j is "cream-skimming" and the effect of "bid low, bid always" may be to win less useful clicks.

For this reason, "bid high, bid sometimes" may be more effective and especially when targeting is quite coarse, so that the danger of "bid low, bid always" is that other advertisers have cream-skimmed all the good users (leaving only low value users with low price). A possible response is to use more precise targeting, which can then be coupled with "bid high, bid sometimes."

10.6.2. Yield Maximization across Audiences

An orthogonal problem facing an advertiser is how to bid across multiple audiences in order to maximize *yield* (= (value-cost)/cost). For example, a yield of 0.2 indicates that a profit of \$0.20 is made on every \$1 spent. For a daily budget of B, and assuming value is greater than cost, the problem of yield maximization is equivalent to maximizing total value subject to hitting the budget B. The yield maximization problem is related to the classical 0-1-KNAPSACK problem.

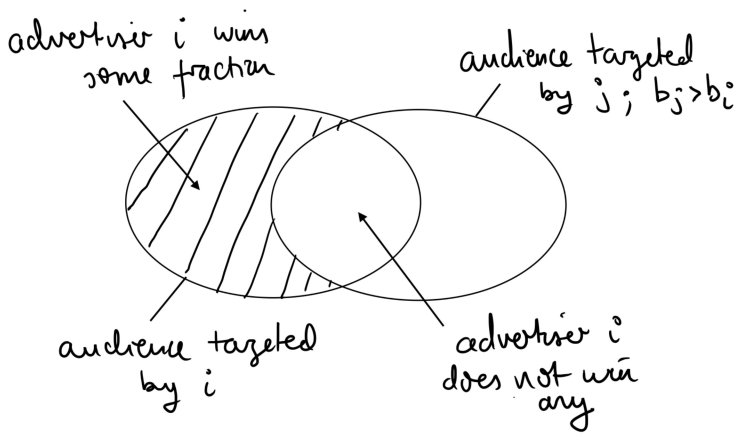


Figure 10.10.: Advertisers i and j target an overlapping audience, and advertiser j out-bids advertiser i. The effect is that j cream-skims the good users.

Definition 10.5 (0-1-KNAPSACK problem)**.** *There are a set of objects $k \in K$, each with value $r_k > 0$ and size $\ell_k > 0$. There is a bag (knapsack) of capacity $C > 0$. Find a set of objects with maximum total value, subject to the constraint that their total size does not exceed the capacity.*

A greedy algorithm for 0-1-KNAPSACK sorts the objects by decreasing 'bang-per-buck,' or score r_k/ℓ_k, and select them greedily while the bag still has capacity, skipping over any object that does not fit given the selections made so far.

Example 10.6. *Consider a* 0-1-KNAPSACK *problem with capacity $C = 8$ and four objects:*

object	1	2	3	4
(value,size)	$(5, 8)$	$(2.1, 3)$	$(1.9, 3)$	$(4, 6)$
bang-per-buck	0.62	0.70	0.63	0.67
select?	*no*	*yes*	*yes*	*no*

The greedy algorithm assigns scores as shown (e.g., $0.62 \approx 5/8$), and sorts the objects into order $2, 4, 3$ and 1. Given this, it selects 2, skips 4 because it does not fit ($6 > 8 - 3$), selects 3, and skips 1 ($8 > 8 - 3 - 3$). The total value is $1.9 + 2.1 = 4$, compared to a value of 5 for the optimal solution, which would just select object 1.

In formulating the yield-maximization problem, we assume that for each audience (say j), the advertiser knows the following:

- $Q_j(b_{ij})$: the expected number of clicks acquired as a function of bid b_{ij}.
- $p_j(b_{ij})$: the expected price-per-click as a function of bid b_{ij}.

	audience 1		audience 2	
possible bid	1	2	3	4
bid-per-click	1.5	3	1	2
value-per-click	2.1	5.4	1.1	3
price-per-click	1	2	0.5	1.5
number of clicks acquired	60	100	250	300
(value,size)	(126,60)	(540,200)	(275,125)	(900,450)
score	2.1	2.7	2.2	2
fraction selected	0	1	0.4	0

Figure 10.11.: Four choices of bids, two to each audience, in a yield maximization problem facing an advertiser.

- $v_j(b_{ij})$: the expected value-per-click as a function of bid b_{ij} (as in pacing, this may depend on the bid amount).

The daily budget B plays the role of the capacity of the knapsack. We assume for simplicity that the set of possible bids for each audience is finite. Each possible bid value b_{ij} corresponds to a distinct object k, with value $r_k = Q_j(b_{ij}) \cdot v_j(b_{ij})$ and size $\ell_k = Q_j(b_{ij}) \cdot p_j(b_{ij})$. The score for this object is $r_k/\ell_k = v_j(b_{ij})/p_j(b_{ij})$.

Unlike the 0-1-KNAPSACK problem, a fraction of an object can be selected (representing bidding for a part of the day). This can be handled by modifying the greedy algorithm to accept a fraction of the first object that cannot be accepted in full without exceeding the capacity.

In addition, it only makes sense to select a single bid per audience. Suppose that bid b_{ij} has already been selected and audience j. If the next best bid is for the same audience, then we replace bid b_{ij} with this bid if the amount that can be spent is greater.

Example 10.7. *There are two audiences, the daily budget* $B = 250$*, and two possible bids for each audience. See Figure 10.11. The greedy algorithm would assign scores as shown (e.g.,* $2.1 = 2.1/1$*), and rank the four bids in order* $2, 3, 1$ *then* 4*. Given this, it selects bid* 2 *(*$200 < 250$*), and then fraction* $0.4 = (250 - 200)/125$ *of object* 3*. The total value is* $540 + (0.4)275 = 650$*, providing a yield of* $(650 - 250)/250 = 1.6$*, and optimal for this instance.*

The algorithm is not always optimal. To see this, suppose instead that bid 1 *generates value-per-click of* 5 *rather than* 2.1*. In this case, the score would be 5, and the sorted bids would be* $1, 2, 3$*, and* 4*. Given this, we'd first select bid* 1 *(*$60 < 250$*), then later select bid* 2 *in place of* 1 *since this enables more spend, and finally accept fraction 0.4 of bid* 3*. The total value would be \$650, whereas accepting bid* 1 *and* $0.42 \approx (250 - 60)/450$ *of bid* 4 *would provide value* $(5)(60) + (190/450)(900) = 680 > 650$*.*

In practice, the problem facing an advertiser is not just one of optimizing bids and budgets across multiple audiences. In addition, an advertiser also needs to learn information about the market, including the value associated with different bids, the cost associated with different bids, and the available volume. See the chapter notes.

10.6.3. Cookie Matching

A cookie is a file that can be written to a user's machine by her web browser when she visits a webpage. Recall that cookie matching enables an adX to map its own cookie for a user with a DSP's cookie, so that this cookie can be communicated to the DSP when making a call-out and asking for a bid.

Let $C_z(u)$ denote the cookie owned by party z on user u's machine. The sequence of steps in cookie matching is as follows:

- A user u visits `marcjacobs.com`. Marc Jacobs requests that its DSP serves a *pixel* (a tiny, invisible webpage) alongside the web content. This allows the DSP to read its cookie $C_{\text{DSP}}(u)$ on the user's machine, or write one to the machine if the user is not known to the DSP. The DSP can now associate its cookie with Marc Jacobs (as well as additional context about the kind of page served by Marc Jacobs, and the time the page is served). Along with behavioral context for other advertiser clients of the DSP, the effect might be:

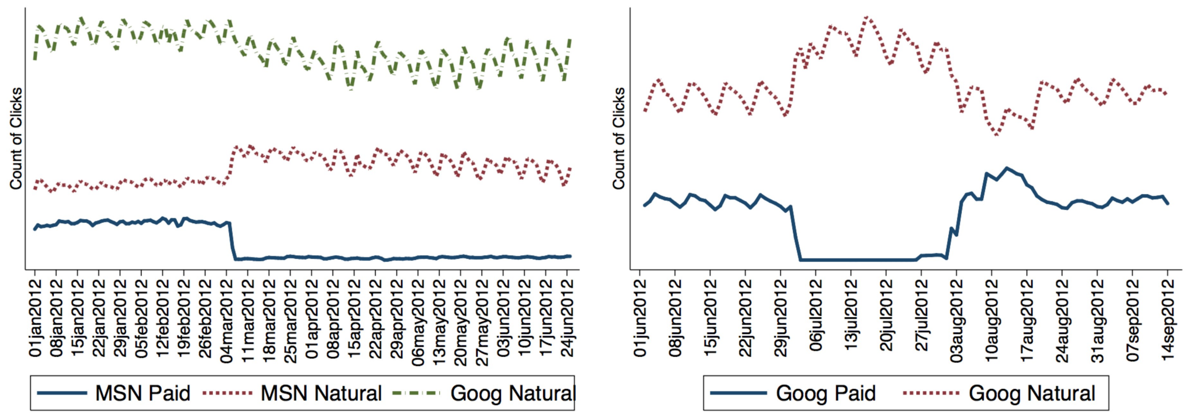
$$C_{\text{DSP}}(u) \mapsto \{(\text{Marc Jacobs, add-to-cart, 03/12/2016}), (\text{Hyatt}, \ldots), \ldots\} \tag{10.26}$$

- At the same time, the DSP also redirects the user's browser to the adX, which has the effect of sending cookie $C_{\text{DSP}}(u)$ to the adX. The adX can now read its own cookie, $C_{\text{adX}}(u)$, or write one if the user is not known to the adX. The adX can now associate (or *match*) its cookie with the cookie of the DSP, adding this to that of other DSPs associated with this user:

$$C_{\text{adX}}(u) \mapsto \{C_{\text{DSP}}(u), \ldots, \}. \tag{10.27}$$

- Suppose that user u visits the New York Times at some time in the future, and that the New York Times uses the adX to decide which ads to display. The web server includes a pixel for the adX in the view for a user, allowing the adX to read cookie $C_{\text{adX}}(u)$. The adX looks up the DSP and sends the matched cookie, $C_{\text{DSP}}(u)$, when making a real-time call-out to the DSP and requesting that it make a bid. The DSP can now know that the user recently visited Marc Jacobs without completing a purchase and bid appropriately, perhaps with a personalized ad (e.g., an image of a jacket that the user had viewed, and a deep link).

The ad ecosystem also includes *data aggregators*, who trade in behavioral data as well as demographic and socioeconomic data about users. For example, the *New Yorker* magazine might decide to generate more revenue by selling the ability to track people who visit the New Yorker site to other advertisers. In this way, a DSP could choose to target an ad for *Vanity Fair* to a user who reads the *New Yorker*. Firms such as BlueKai and Axciom facilitate these kinds of trades. This ability to track users across the web and across multiple devices, coupled with granular targeting of ads, raises privacy concerns. We return to this topic in Chapter 30.



(a) MSN Test **(b) Google Test**

Figure 10.12.: Brand Keyword Substitution at eBay. (Blake et al. 2015)

10.6.4. Effectiveness of Online Advertising

The average revenue from a page of sponsored search ads has generally been more than 10x larger than that available from other advertising models (premium inventory display ads aside). This is because of the very specific information about a user's current intent that comes from a search query. For example, recent searches might reveal whether or not a user is at work, whether she is interested in making a purchase, or planning her weekend. Estimates suggest that as many as 1-in-6 searches on Google result in a click on an ad. Behavioral targeting and contextual ads are closing the gap to sponsored search, though, generating increasing amounts of revenue for ad auction platforms.

But what about the value to advertisers? A famous quote attributed to retailer John Wannamaker states that "I know that half the money I spend on advertising is wasted, but I can never find out which half." In the context of sponsored search, what is important is not the number of clicks but the number of incremental clicks (and purchases) that arise as a result of advertising.

In 2012, eBay conducted a large-scale experiment to measure the effectiveness of sponsored search ads. The effect is that as of 2013 eBay is no longer using sponsored search. Prior to the experiment, eBay had been bidding on brand searches that included its name, e.g., "ebay shoes." eBay halted these ads on Yahoo and Bing (then MSN) while continuing to bid on Google (this providing the control.) The click traffic and attributed sales were basically unaffected, suggesting that ads were simply diverting users from clicking on organic search results, which are free to eBay. See Figure 10.12, which shows the substitution from paid clicks to organic clicks (for MSN along with a subsequent test on Google.)

eBay had also been placing ads on nonbrand searches such as "shoes" (managing over 100 million keywords and keyword combinations.) The experiment turned off bids for nonbrand searches in selected geographic areas on Google, switching off 30% of eBay's bidding activity for 60 days. The average effect was a very small and statistically insignificant effect on sales. Advertising was effective in increasing new registered users and purchases made by infrequent purchasers, eBay estimated the overall return on investment to be negative.

eBay's experiment raises the question about whether sponsored search advertising is effective for firms that already have good brand recognition, for whom consumers may already find the site through organic search results or direct navigation. eBay is the second largest e-tailer and an established brand. Subsequent research conducted through a large-scale, fully randomized experiment on Bing showed that smaller companies can benefit from sponsored search advertising (e.g., because they gain traffic from top competitors.)

10.7. Chapter Notes

Goto.com introduced auctions for search advertising in February 1998. The model allowed per-click bidding and bids were used to determine the rank of search results. By 2001, then re-branded as Overture, "organic" (non auction) search results were provided adjacent to auction results. Bidding per-click meant that an advertiser only made a payment when a click on the advertisement was received.

Google introduced the use of estimated CTRs in 2002, ranking bids not by bid amount but by the product of eCTR and bid price-per-click (rank-by-revenue). Google also adopted the GSP payment rule. The estimate that as many as 1-in-6 searches result in a paid click comes from Lahaie (2006).

Yahoo also adopted the GSP auction a little later in 2002. Overture was acquired by Yahoo in the second half of 2003. By 2007, Yahoo had adopted rank-by-revenue in place of rank-by-bid. The squashing idea is introduced in Lahaie (2006), and provides a practical approach to improving revenue. Today, Microsoft's Bing search engine also uses rank-by-revenue and the GSP payment rule. See Varian (2006) and Fain and Pedersen (2006) for additional historical notes.

The analysis of balanced bidding and the balanced EFNE in the GSP auction follows Edelman *et al.* (2007) (and the terminology follows that of Edelman and Schwarz (2010), who study bidding dynamics); see also Varian (2007). The instability of bids in earlier, first-price auction designs is discussed in Edelman and Ostrovsky (2007).

Revenue equivalence (see Chapter 6) can be generalized to multi-item auctions such as position auctions, but does not apply to our analysis of the GSP and VCG position auctions because we study a complete information (Nash) equilibrium of the GSP auction. Revenue equivalence applies when comparing the incomplete information equilibria of auctions. Gomes and Sweeney (2014) show that a Bayes-Nash equilibrium (BNE) may not exist in the GSP auction. By way of contrast, the *generalized first price auction* (GFP), which charges the amount of an advertiser's bid in the event of a click, has a unique and efficient BNE; see Chawla and Hartline (2013).

Muthukrishnan (2009) provides a nice description of ad exchanges along with open challenges. His notes on adXs also inspired the treatment of cookie matching in this chapter (personal communication). Our use of "adX" as an abbreviation for an ad exchange is not intended to refer to DoubleClick's ad exchange.

For references on bid pacing, see Ghosh *et al.* (2009) and Lang *et al.* (2012). For references on yield optimization, see Muthukrishnan *et al.* (2010) and Feldman *et al.* (2008). Other challenges in internet advertising relate to the design of algorithms for matching uncertain supply with uncertain demand; e.g., display ads with bids that demand a particular number of user impressions; see Mehta *et al.* (2007), Buchbinder *et al.* (2007) and Vee *et al.* (2010).

The eBay study on the effectiveness of sponsored search advertising is due to Blake *et al.* (2015). The study on the effectiveness of sponsored search advertising for smaller firms is due to Simonov *et al.* (2015).

10.8. Comprehension Questions and Exercises

10.8.1. Comprehension Questions

c10.1 Why does separability between quality and position effect simplify the problem of winner determination?

c10.2 Give three reasons why we study the balanced EFNE of the GSP auction.

c10.3 From an ad platform's perspective, what is one advantage and one disadvantage of using the GSP auction vs. the VCG position auction?

10.8.2. Exercises

10.1 The GSP position auction

a) Verify that the bids in Example 10.3 form a Nash equilibrium.

b) Verify that the bids in Example 10.5 are a balanced EFNE.

c) Verify that there are no useful deviations from truthful bidding in the VCG in Example 10.1.

d) Prove that bids satisfy the value-ordering property in any EFNE.

e) Prove that the revenue in every EFNE of the GSP auction is at least as large as the revenue from truthful bidding in the VCG mechanism.

f) Prove that a Nash equilibrium that satisfies balanced bidding is envy-free.

10.2 VCG vs GSP auction

a) Fixing the bids, prove that the expected payment by each advertiser in the GSP auction is at least that in the VCG position auction.

b) Why is this a misleading way to compare the revenue properties of the two auctions?

c) Prove that the outcome of the GSP auction with balanced bidding and a reserve price r is equal to the outcome of the VCG with a reserve price r. [Hint: a reserve in the VCG can be modeled by introducing a dummy agent representing the seller, always willing to buy any number of positions at price r.]

10.3 Give another example, other than that in Example 10.3, where a Nash equilibrium of the GSP auction is not value ordered and thus not efficient.

10.4 Suppose a position auction that uses the same winner-determination rule as GSP and VCG (ranking bids by $CTR_i \cdot b_i$), but has a different payment rule.

a) Use the payment identity characterization from Chapter 8 to derive a payment rule to make the auction truthful.

b) How does this payment rule compare to that in the VCG mechanism?

10.5 Provide an example to demonstrate that a first-price position auction may have no (pure strategy) Nash equilibrium.

