Comparing Different Yahoo Sponsored Search Auctions: A Regression Discontinuity Design Approach

Jia Yuan*
Department of Economics
University of Minnesota

Abstract

The sponsored search auction is a successful pricing mechanism which helps search engine companies sell navigation service to advertisers. Correctly understanding the performance differences among different types of sponsored search auctions will not only affect the multibillion dollar revenue of search engine companies, but it will also help develop more superior sponsored search auctions in the future. For the two popular sponsored search auctions—the Generalized First Price (GFP) auction and the Generalized Second Price (GSP) auction—current consensus in both the industry and academia is that the GSP auction is more stable and more efficient than the GFP auction. Specifically, in the GSP auction, bidders are less likely to "game the system", meaning that an individual bidder will change his bid less frequently; his bid range will be smaller; and a bidder with a higher value will be more likely to win a higher and better slot. This paper examines this prevailing belief using a Regression Discontinuity Design (RDD) approach and finds that after bidders switch to the GSP auction, they actually bid 36% more frequently and increase their daily bid range by \$1.31. To compare efficiency differences, this paper constructs an efficiency index and shows that the GSP auction mechanism is at least 4% more efficient. This paper contributes to the sponsored search auction literature in two aspects. First, this paper provides empirical evidence contradicting the current consensus about the GFP auction and the GSP auction. Second, this is the first research to construct an efficiency index and empirically evaluate the efficiency improvement of the GSP auction.

Keywords: Online advertising, sponsored search auction, regression discontinuity

*I am grateful to Patrick Bajari for his continuous guidance and support. I thank Tom Holmes, Kyoo il Kim, Minjung Park, Amil Petrin, Robert Town and all other participants at the Applied Microeconomics Workshop for comments and suggestions. I also thank the Yahoo! research department for providing bid data. Remaining errors are mine. Address for correspondence: 4-101 Hanson Hall, 1925 Fourth Street South, Minneapolis, MN, 55455. Email: yuanjia@econ.umn.edu

1 Introduction

The sponsored search auction has played an indispensable role in the success of search engine giants like Yahoo! and Google. For example, Yahoo!'s first half-year revenue in 2008 was \$3.62 billion and at least 50% of that revenue came from the sponsored search auction.[1] For Google, its first half-year revenue in 2008 was \$10.55 billion with 97% of this revenue generated by the sponsored search auctions.[2] Actually, the sponsored search auction is not only crucial to search engine companies, but it is also "vital to the success of many other small business" such as bid management software firms, bidding campaign consulting firms, and key word selecting firms, etc. (See Jansen and Mullen (2008).)

The sponsored search auction is a pricing mechanism which helps search engine companies sell navigation services to advertisers. When addressing search requests, search engines display both the search results and advertisers' web links, which are called sponsored links. These sponsored links attempt to navigate potential customers to specific product web sites. Because this targeting of potential costumers has proven effective, advertisers are willing to pay in order to obtain an ideal placement for their web link on a search result page. Search engine companies invented the sponsored search auction to sell these sponsored link placements.

The sponsored search auction was first introduced in 1998 by Goto for Yahoo!. Since then, search engine designers have upgraded the mechanism several times. The purpose of replacing an old sponsored search auction with a new one is "to bring more stability to the auction bidding, increase profits, and help reduce strategic bidding". (See Jansen and Mullen (2008).) One of the major transformations the sponsored search auction has undergone was Yahoo!'s switch from the Generalized First Price (GFP) auction to the Generalized Second Price (GSP) auction.[3] This auction rule change, which took place on June 26, 2002, is

[1] See Yahoo! Quarterly Report on Form 10-Q to SEC for the quarter ended June 30, 2008 and Borgers, Cox, Pesendorfer and Petricek (2007).

[2] See Google Inc Quarterly Report on Form 10-Q to SEC for the quarter ended June 30, 2008 and Borgers et al. (2007).

[3] During 2002, the Yahoo! sponsored search auction was managed by a company named Overture, which

generally believed to have been a success by both the industry and academia in the sense that "superior designs" have replaced the "inefficient market institutions". (See Edelman, Ostrovsky and Schwarz (2008) and Jansen and Mullen (2008).) The GSP auction is believed to be more efficient because while using it, bidders will be less likely to "game the system". This means that an individual bidder will change his bid less frequently and that his bid range will be smaller; moreover, a bidder with a higher value will be more likely to win a higher and better position with a higher amount of clicks.

Correctly understanding and evaluating how different sponsored search auctions perform is important for both economists and the search engine industry. Having the correct answers will not only affect the multibillion dollar revenue of search engine companies, but it will also help develop more superior sponsored search auctions in the future. This paper examines the prevailing belief that the GSP auction is superior to the GFP auction using bid data collected from Yahoo! sponsored search auctions in 1000 markets from between June 15, 2002 and June 14, 2003.

Yahoo!'s auction rule change on June 26, 2002 provides an opportunity to compare the performances of the two auction mechanisms in a treatment effect framework. Specifically, all the bidders after June 26, 2002 would face a treatment of the GSP auction. Thus, estimating the performance differences of the two auction systems will be turned into identifying the average treatment effect. However, in the standard treatment effect framework, the identification usually depends on strong assumptions on the comparison group and control group. This paper avoids this challenge by using a Regression Discontinuity Design (RDD) approach, which enables us to maintain relatively simple and reasonable assumptions to obtain identification.

Recently, there has been a renaissance of the RDD method to estimate the treatment effect. RDD is a special case of treatment effect analysis, usually applied under circum-

later was acquired by Yahoo!. Without causing confusion, this paper does not distinguish these two names and will always use Yahoo! sponsored search auction.

Under the new auction rule, bidders could choose either the GFP auction or the GSP auction to submit their bids.

stances where the treatment probability function displays a sharp and observable discontinuity jump at some cutoff point of an observable variable called selection variable. Hahn, Todd and Van der Klaauw (2001) and Lee (2008) established the identification conditions for RDD, and now RDD has been broadly applied to estimate the treatment effect in many economic contexts. Van der Klaauw (2002) estimates the effect of financial aid offers on college enrollment through a RDD approach by exploiting the discontinuity in the financial aid assignment rule. Angrist and Lavy (1999) observed that in the Israeli public school system, the Maimonides' rule of "maximum class of 40" generated an exogenous source of variation in the classes, and the author used this variation to estimate the class size effect on scholastic achievement. Lee (2008) applies RDD to estimate incumbency advantage in U.S. House elections by exploiting the fact that candidates receive the treatment of winning the election when the vote share is bigger than $\frac{1}{2}$. Chen and Van der Klaauw (2008) use RDD to estimate the work disincentive effects of the disability insurance program. This paper extends the application of RDD to the Yahoo! sponsored search auction.

Contrary to conventional wisdom, the estimation results show that the bidding behavior under the GSP auction was less stable than thought. The daily frequency with which an individual bidder changed his bid increased by 6.8 times, representing a 36% increase. In addition, the daily bid range of each bidder increased by $1.31. Plus, the daily maximum value of the bids submitted by each individual bidder increased by 55 cents. All the above estimates suggest that the GSP auction did not reduce the "strategic bidding behaviors" as believed by most economists and developers in the industry.

These findings have important implications for the current sponsored search auction theory literature. Recent theories on the GSP auction, including Edelman et al. (2008), Varian (2006) and Athey and Ellison (2007), are basically based on a static game theory structure. Edelman et al. (2008) and Varian (2006) argue that this game framework "describes the basic properties of the prices observed in Google's ad auction reasonably accurately." However, Google is not using a pure GSP auction analyzed by the theories, and the above evidence actually shows that the bidding behaviors in the GSP auction are even more volatile and more aggressive than those under the GFP auction. This may suggest that our theoretical understanding about bidding behavior and equilibrium properties under the sponsored search auction from a stable framework, which also provided the guidance for the latter sponsored search auction upgrade, might not be well founded.

This paper also estimates the efficiency improvement, which the GSP auction brought to the auction market by replacing the GFP auction, as suggested by the literature. To measure efficiency, I first construct an index measure based on the following idea: a more efficient auction system should help the bidder with the higher value obtain the higher slot more often. If the auction is fully efficient, bidders with higher values should always dominate the bidders with lower values, and we should observe that the probability that higher value advertisement rank higher than always be 1. The less efficient the mechanism is, the smaller this probability will be. Therefore, this relative ranking between two bidders can be used as an index to measure the efficiency of the auction mechanism.

The challenge of identifying the efficiency improvement is that bidders' true values were unobservable. However, we observe the following facts. If the new system can improve the bidding efficiency, on average, the probability index bigger than $\frac{1}{2}$ in the old system will be even bigger than $\frac{1}{2}$ in the new auction system; and a probability index smaller than $\frac{1}{2}$ in the old system will be even smaller than $\frac{1}{2}$ in the new system. Based on these observations, we propose an estimation strategy and find that the new auction mechanism is at least 4% more efficient. In other words, the GSP auction system gives the advertiser with a higher value a 4% better chance to obtain a higher slot.

This paper contributes to the sponsored search auction literature in two aspects. First, in the past there was no empirical analysis to compare and evaluate the performances of different sponsored search auctions. In past literature, the comparison between the two popular auctions—the GFP auction and the GSP auction—was illustrated purely by hypothetical examples, which will be discussed in detail in section 3. This paper, however, provides solid empirical evidence contradicting the current beliefs about the comparison between the GFP auction and GSP auction.

Second, this research constructs an efficiency index and it is also the first to empirically evaluate the efficiency improvement of the GSP auction. Understanding and evaluating how efficiently the auction system allocates link placements is both an important and challenging question, especially when each bidder's true value in the auction is unobservable. This paper turns measuring efficiency into comparing the relative ranking between two bidders and is the first to identify the efficiency improvement brought by the GSP auction.

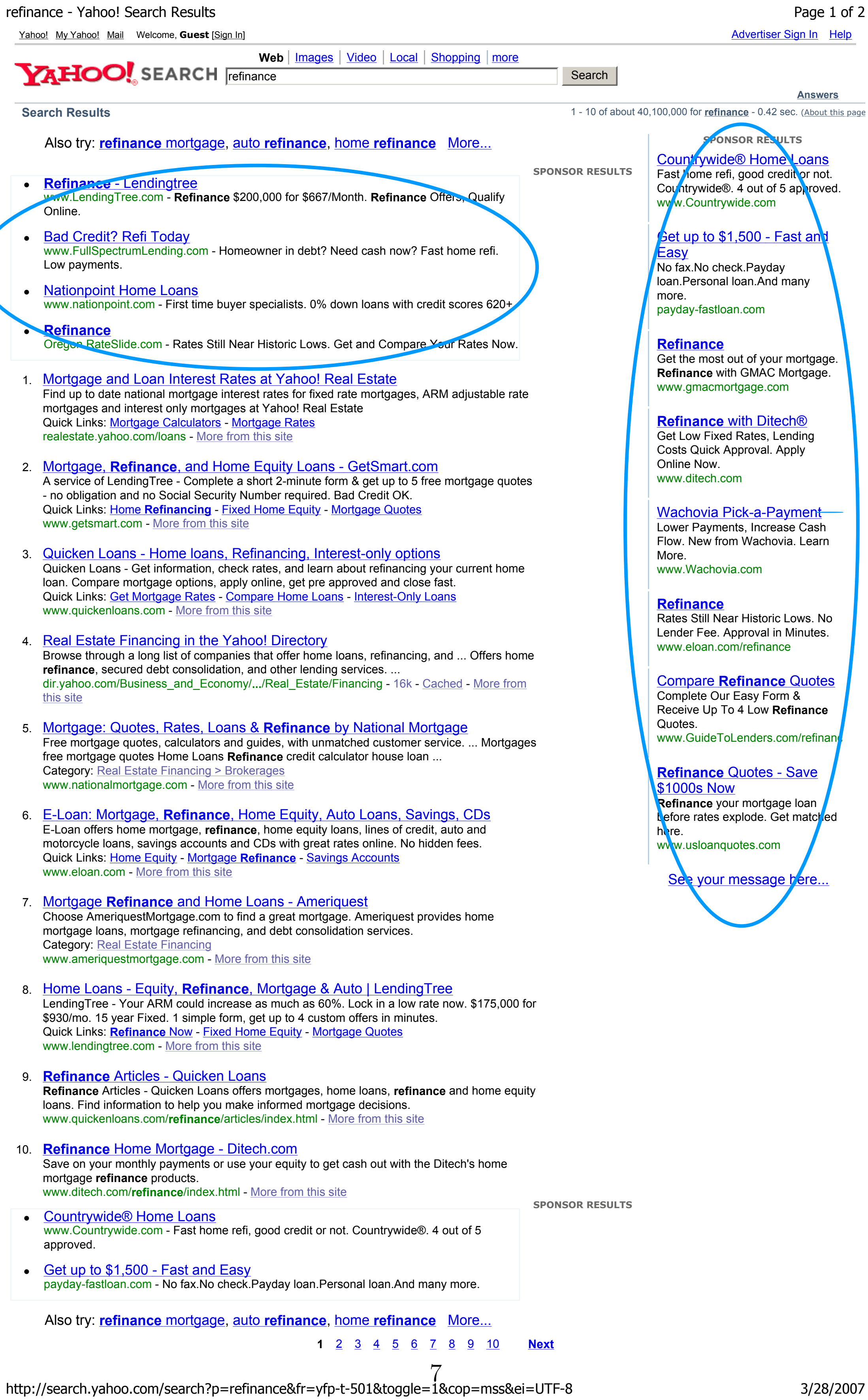
The paper is organized as follows. Section 2 introduces the Yahoo! sponsored search auction. Section 3 briefly surveys the sponsored search auction literature and especially examines the conventional wisdom about the performance of the GFP auction and GSP auction. Section 5 sets up the RDD model. Section 4 introduces the data and presents the simple statistics and OLS regression results. Section 6 presents the RDD estimation results. Section 7 constructs an efficiency index and evaluates efficiency improvement of the GSP auction over the GFP auction. Section 8 concludes.

2 Yahoo! Sponsored Search Auction

In the search engine industry, there are three key players: the advertisers, the search engines and the potential customers. Search engines navigate potential customers to advertisers' product web sites by displaying their web links when potential customers conduct keyword search requests. These advertisers' links are called sponsored links. Sponsored links distinguish themselves from the organic (non-sponsored) web search results by whether or not a fee is paid to the search engine company.

Figure 1 shows an example of sponsored links for the key word "refinance". When someone uses Yahoo! to search for information about "refinance", the search engine will display search results along with sponsored links, which are circled in Figure 1. Usually

Figure 1: Sponsored Links for the Keyword "Refinance"



refinance - Yahoo! Search Results Page 1 of 2

Yahoo! My Yahoo! Mail Welcome, **Guest** [Sign In] Advertiser Sign In Help

Web | Images | Video | Local | Shopping | more

YAHOO! SEARCH refinance Search

Answers

Search Results 1 - 10 of about 40,100,000 for **refinance** - 0.42 sec. (About this page

Also try: **refinance** mortgage, auto **refinance**, home **refinance** More...

SPONSOR RESULTS

- **Refinance** - Lendingtree
  www.LendingTree.com - **Refinance** $200,000 for $667/Month. **Refinance** Offers, Qualify Online.
- Bad Credit? Refi Today
  www.FullSpectrumLending.com - Homeowner in debt? Need cash now? Fast home refi. Low payments.
- Nationpoint Home Loans
  www.nationpoint.com - First time buyer specialists. 0% down loans with credit scores 620+
- **Refinance**
  Oregon.RateSlide.com - Rates Still Near Historic Lows. Get and Compare Your Rates Now.

1. Mortgage and Loan Interest Rates at Yahoo! Real Estate
   Find up to date national mortgage interest rates for fixed rate mortgages, ARM adjustable rate mortgages and interest only mortgages at Yahoo! Real Estate
   Quick Links: Mortgage Calculators - Mortgage Rates
   realestate.yahoo.com/loans - More from this site
2. Mortgage, **Refinance**, and Home Equity Loans - GetSmart.com
   A service of LendingTree - Complete a short 2-minute form & get up to 5 free mortgage quotes - no obligation and no Social Security Number required. Bad Credit OK.
   Quick Links: Home **Refinancing** - Fixed Home Equity - Mortgage Quotes
   www.getsmart.com - More from this site
3. Quicken Loans - Home loans, Refinancing, Interest-only options
   Quicken Loans - Get information, check rates, and learn about refinancing your current home loan. Compare mortgage options, apply online, get pre approved and close fast.
   Quick Links: Get Mortgage Rates - Compare Home Loans - Interest-Only Loans
   www.quickenloans.com - More from this site
4. Real Estate Financing in the Yahoo! Directory
   Browse through a long list of companies that offer home loans, refinancing, and ... Offers home **refinance**, secured debt consolidation, and other lending services. ...
   dir.yahoo.com/Business_and_Economy/.../Real_Estate/Financing - 16k - Cached - More from this site
5. Mortgage: Quotes, Rates, Loans & **Refinance** by National Mortgage
   Free mortgage quotes, calculators and guides, with unmatched customer service. ... Mortgages free mortgage quotes Home Loans **Refinance** credit calculator house loan ...
   Category: Real Estate Financing > Brokerages
   www.nationalmortgage.com - More from this site
6. E-Loan: Mortgage, **Refinance**, Home Equity, Auto Loans, Savings, CDs
   E-Loan offers home mortgage, **refinance**, home equity loans, lines of credit, auto and motorcycle loans, savings accounts and CDs with great rates online. No hidden fees.
   Quick Links: Home Equity - Mortgage **Refinance** - Savings Accounts
   www.eloan.com - More from this site
7. Mortgage **Refinance** and Home Loans - Ameriquest
   Choose AmeriquestMortgage.com to find a great mortgage. Ameriquest provides home mortgage loans, mortgage refinancing, and debt consolidation services.
   Category: Real Estate Financing
   www.ameriquestmortgage.com - More from this site
8. Home Loans - Equity, **Refinance**, Mortgage & Auto | LendingTree
   LendingTree - Your ARM could increase as much as 60%. Lock in a low rate now. $175,000 for $930/mo. 15 year Fixed. 1 simple form, get up to 4 custom offers in minutes.
   Quick Links: **Refinance** Now - Fixed Home Equity - Mortgage Quotes
   www.lendingtree.com - More from this site
9. **Refinance** Articles - Quicken Loans
   Refinance Articles - Quicken Loans offers mortgages, home loans, **refinance** and home equity loans. Find information to help you make informed mortgage decisions.
   www.quickenloans.com/**refinance**/articles/index.html - More from this site
10. **Refinance** Home Mortgage - Ditech.com
    Save on your monthly payments or use your equity to get cash out with the Ditech's home mortgage **refinance** products.
    www.ditech.com/**refinance**/index.html - More from this site

SPONSOR RESULTS

- Countrywide® Home Loans
  www.Countrywide.com - Fast home refi, good credit or not. Countrywide®. 4 out of 5 approved.
- Get up to $1,500 - Fast and Easy
  payday-fastloan.com - No fax.No check.Payday loan.Personal loan.And many more.

Also try: **refinance** mortgage, auto **refinance**, home **refinance** More...

1 2 3 4 5 6 7 8 9 10 **Next**

SPONSOR RESULTS

Countrywide® Home Loans
Fast home refi, good credit or not. Countrywide®. 4 out of 5 approved.
www.Countrywide.com

Get up to $1,500 - Fast and Easy
No fax.No check.Payday loan.Personal loan.And many more.
payday-fastloan.com

Refinance
Get the most out of your mortgage. **Refinance** with GMAC Mortgage.
www.gmacmortgage.com

Refinance with Ditech®
Get Low Fixed Rates, Lending Costs Quick Approval. Apply Online Now.
www.ditech.com

Wachovia Pick-a-Payment
Lower Payments, Increase Cash Flow. New from Wachovia. Learn More.
www.Wachovia.com

Refinance
Rates Still Near Historic Lows. No Lender Fee. Approval in Minutes.
www.eloan.com/refinance

Compare **Refinance** Quotes
Complete Our Easy Form & Receive Up To 4 Low **Refinance** Quotes.
www.GuideToLenders.com/refinanc

Refinance Quotes - Save $1000s Now
Refinance your mortgage loan before rates explode. Get matched here.
www.usloanquotes.com

See your message here...

http://search.yahoo.com/search?p=refinance&fr=yfp-t-501&toggle=1&cop=mss&ei=UTF-8 3/28/2007 around 10 sponsored links, located on the top and on the right of each page, will be displayed.

Advertisers are interested in buying these link slots for their product web sites because they may target the potential customers more efficiently. In 1998, Goto first introduced the sponsored search auction in the search engine industry to sell these link slots.[4]

The sponsored search auction is a multi-object dynamic auction in which each individual advertiser bids for the ideal slot for his web site. Sponsored search auctions usually have the following common features. First, all the link slots are auctioned at the same time. As shown in Figure 1, there were at least 12 sponsored link slots being auctioned at that time. Second, the auction is dynamic with an infinite time horizon. Each bidder can change or withdraw his bid at any time, which will be immediately reflected in the slot placement. Third, all search engines share a common payment rule: pay per click (PPC), which means that whenever there is a click on the sponsored link, the bidder will pay Yahoo! once. And lastly, in Yahoo!'s sponsored search auction, all the information, including bids and slot placement, is public information, which can be observed by all the bidders directly.

In keeping with the keyword search for Figure 2 "Refinance", shows all bidders' bids and slot allocation information as it was captured by a free public web site.[5] The bid range is from $16.13 to $7.49 and each bidder's position is determined solely by his bid. As can been seen, "LendingTree" had the highest bid; therefore, this advertisement was placed at the highest slot as shown in Figure 1.

Designing efficient auction rules regarding how the advertisers pay the search engine and how the search engine allocates the link slots among the advertisers is a key challenge faced by the search engine designers because the decision to adopt different forms of sponsored search auctions has an important impact on the success of search engine companies. In the past 6 years, Yahoo! upgraded its sponsored search auction several times hoping to find a better auction mechanism to bring more stable bidding behaviors and higher auction

[4]Goto was later renamed to Overture and acquired by Yahoo! in late 2003.

[5]The free bid check website is http://keyword.secretstohighprofit.com/default.aspx. Figure 1 and Figure 2 were captured at the same time on March 28, 2007.

Figure 2: Bids and Rankings

Keyword Dynamo - Overture Bids and Suggestion Tool — Page 1 of 1

Keyword: economics

Search

☐ Overture Keyword Suggestion
☑ Overture Bids

Keywords Dynamo

Welcome to the NEW Keyword Dynamo Tool. The only Tool where you can view Overture Bids for US Market. Due to it popularity and usage overture has blocked

View Bids

Type in a search term and we'll show you the Max Bids and listings for that term.

refinance

Search Cancel

1. **Refinance - Lendingtree**
   $300,000 For Only $1,000/month. Refinance Today. Bad Credit Options.
   www.LendingTree.com
   (Advertiser's Max Bid: $16.13)

2. **Bad Credit? Refi Today**
   Homeowner in debt? Need cash now? Fast home refi. Low payments.
   www.FullSpectrumLending.com
   (Advertiser's Max Bid: $12.53)

3. **Nationpoint Home Loans**
   First time buyer specialists. 0% down loans with credit scores of 620+.
   www.nationpoint.com
   (Advertiser's Max Bid: $9.70)

4. **Refinance**
   Rates Still Near Historic Lows. Get and Compare Your Rates Now.
   Oregon.RateSlide.com
   (Advertiser's Max Bid: $9.55)

5. **Countrywide® Home Loans**
   Refi to combine 1st mortgage & debt. Low payments with a 40-year loan.
   www.Countrywide.com
   (Advertiser's Max Bid: $8.75)

6. **Get up to $1,500 - Fast and Easy**
   No fax.No check.Payday loan.Personal loan.And many more.
   payday-fastloan.com
   (Advertiser's Max Bid: $8.42)

7. **Refinance**
   Refinance and reduce your payments upto 60%.
   www.floridaslowestrates.com
   (Advertiser's Max Bid: $7.49)

8. **Refinance**
   Get the most out of your mortgage. Refinance with GMAC Mortgage.
   www.gmacmortgage.com

http://keyword.secretstohighprofit.com/default.aspx — 3/28/2007 revenue.

Before June 26, 2002, a bidder in the Yahoo! sponsored search auction paid Yahoo! his bid multiplied by the number of the clicks on his web site. For example, if a bidder bid \$3 and his web site received 3000 clicks, the bidder would have to pay Yahoo! \$9,000. The literature calls this type of sponsored search auction "Generalized First Price (GFP) Auction" to distinguish it from the standard first price auction.

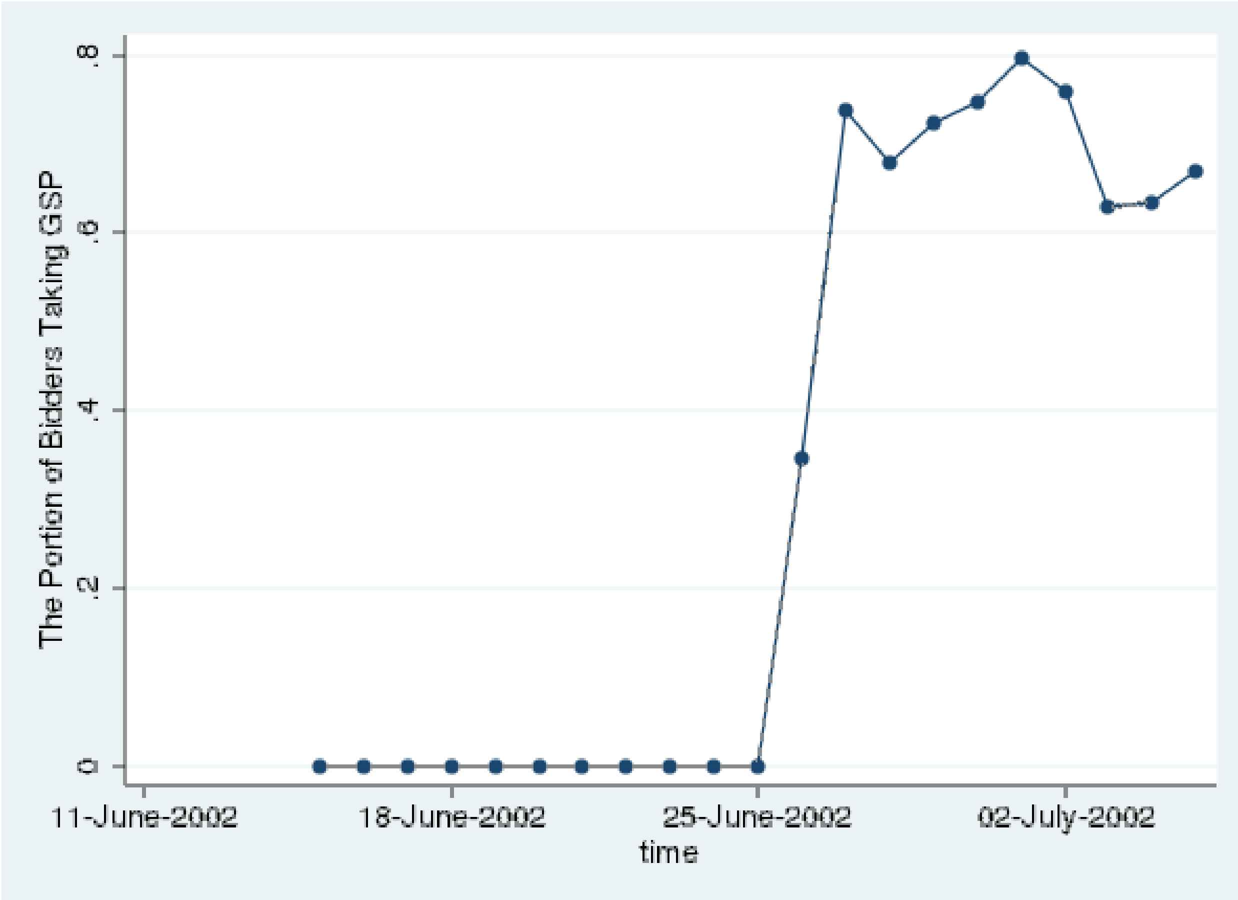
On June 26, 2002, Yahoo! upgraded its Generalized First Price (GFP) Auction to a Generalized Second Price (GSP) Auction. In this new auction system, the web site placement was still determined solely by a bidder's bid, but each bidder, instead of paying his own bid per click, only had to pay 0.01 more than the next highest bid below his. For example, if two bidders bid \$0.4 and \$0.6, respectively, in the old bidding system, the winner would pay \$0.6 per click received; however, in the GSP auction system, he would be charged at a rate of \$0.41.

The most recent Yahoo! sponsored search auction upgrade took place in 2007. Before May 2007, slot allocation was determined only by bidders' bids. The bidder with higher bids got higher link slots as shown in Figure 1 and Figure 2. After May 2007, Yahoo! sponsored search auctions no longer determined slot allocation solely based on bidders' bids, but also by the quality of an advertiser's web site. To do this, Yahoo! created a score system to rank bidders' links.

Even though this rule change of Yahoo! sponsored search auction in 2007 is also very important and interesting, this paper keeps its focus on the Yahoo! sponsored search auction upgrade which happened in 2002. When the GSP auction was introduced, bidders could choose whether to submit their bid in the GFP auction system or in the original GSP auction system, making the choice of the GSP auction endogenous.

This new 2002 auction rule had a dramatic change on bidding behavior. Figure 3 shows the sharp jump in the number of bidders submitting their bids through the new GSP auction system. The y axis denotes the portion of the bidders who switched from the GFP auction

Figure 3: The Portion of bidders adopting the GSP auction



rule. The number jumped from zero to around 70% immediately after June 26, 2002. After that, it remained steady at around 70%. This jump is actually the identification source of the causal effect in the following regression discontinuity approach.

The probability function of whether a bidder receives a GSP auction treatment is endogenous, instead of a function with a probability equal to one. The literature (See Jansen and Mullen (2008)) does not distinguish this subtle difference, and actually no research has analyzed how bidders bid when the GSP auction is endogenously chosen. In this paper, we take this endogeneity into account and further details are addressed in section 5

3 Literature

Recent research on the sponsored search auction mainly focus on three perspectives. First, economists are interested in providing a theoretical game foundation for this new auction mechanism. Varian (2006) and Edelman et al. (2008) first introduced equilibrium concepts for the GSP auctions based on the idea of "envy-free", which assumes that in the equilibrium no bidder would like to place a bid that would cause retaliation. All authors suggest that the GSP auction can achieve efficient allocations. In a similar setup, Athey and Ellison (2007) further introduce consumer search behavior into the model and analyze the implications for reserve prices, product variety, etc.

Second, both economists and search engine developers are interested in the bidders' overall advertising campaign performances taking the sponsored search auction as given. Ghose and Yang (2007) propose a novel empirical model to quantify how different metrics affect bidders' advertising campaign performances. Rutz and Bucklin (2007) use hierarchical Bayes binary choice model to estimate the keyword conversion rate and, based on the model, propose better advertising campaign strategies.

Third, many other topics derived from the sponsored search auction are also attracting economists' attention. Goldfarb and Tucker (2008) investigate the relationship between matching difficulty and bidding prices. They found evidence showing that the more difficult it is to make a match between the firms and customers, the higher the bids in the sponsored search auction. Animesh, Ramachandran and Viswanathan (2005) study the relationship between an advertiser's quality and his bidding strategies and find evidence of significant adverse selection associated with product uncertainty.

This research is an empirical work, which is closely related to the second group of the literature. A bidder's advertising campaign mainly consists of two parts. The first part is how to place a bid to obtain a good placement, which is related to costs; the second part is how to increase purchases to generate more revenue. This paper mainly focuses on the cost side and asks the question: How will a specific type of sponsored search auction affect advertisers' bidding behaviors? Although studying the performance differences among different sponsored search auctions is an important question, from the perspectives of both the search engine developers and advertising bidders, all of the current empirical research analyzes economic behavior under one specific sponsored search auction. None has conducted any empirical comparisons among different sponsored auction mechanisms adopted in the industry. This paper, to my knowledge, is the first empirical paper comparing the performances of the GFP auction mechanism and the GSP auction mechanism, and providing evidence that contradicts the current prevailing beliefs.

These results also have important implications for the current sponsored search auction theory literature. The theory papers authored by Edelman et al. (2008), Varian (2006) and Athey and Ellison (2007) are based on a static game theory structure that analyzes the GSP auction. Edelman et al. (2008) and Varian (2006) argue that this game framework "describes the basic properties of the prices observed in Google's ad auction reasonably accurately." However, Google is not using a pure GSP auction analyzed by the theories; Borgers et al. (2007) suggest that this static GSP auction model actually may have a very poor explanation power on the real data collected from the Yahoo! sponsored search auction. This paper also draws similar conclusion from another angle. If the evidence shows that the bidding behaviors in the GSP auction are more volatile and more aggressive than those in the GFP auction, it may suggest that our theoretical understanding about the bidding behavior and equilibrium properties under the sponsored search auction from a stable framework, which also provided the guidance for the latter sponsored search auction upgrade, might not be well founded. The following subsections will introduce the current prevailing belief about the GSP auction and the GFP auction, which is the hypothesis this paper will test.

3.1 Conventional Wisdom on the GSP auction and the GFP auction

Currently theories mainly focus on the GSP auction in a static setting; in contrast, hardly any formal theoretical analysis has been done on the GFP auction. The conventional wisdom about the comparison of the two auctions was based mainly on concrete examples instead of formal game theory setup. Edelman et al. (2008) proposed a simple example, which the following literature frequently cited. (See Edelman and Ostrovsky (2006) and Jansen and Mullen (2008).) In this subsection we also follow this example to illustrate the current consensus and what it misses.

Example 1 (Edelman et al. 2008): *There are two slots for the links. The first slot receives 400 clicks per hour, and the second slot receives 100 clicks per hour. There are three advertisers bidding to place their product. The value per click for the bidders are \$5, \$4 and \$2. Call these three bidders A, B, C respectively.*

Edelman et al. (2008) use this example to illustrate the superiority of the GSP auction. They show that in the GSP auction, the equilibrium bids of A, B, C will be \$5, \$4 and \$2 and that with these bids, efficient allocation is achieved. But in the GFP auction, the equilibrium will not be stable. B will bid \$2.01 instead of \$4 and A will bid \$2.02 instead of \$5. B will outbid A at \$2.03 and the bids escalate until \$4. B will pull his bid back to \$2.01 and the bid escalation goes on again. These bidding behaviors will result in the sawtooth pattern of a bidding war, which is well documented in the literature. (See Edelman and Ostrovsky (2006) and Zhang (2005).) Based on this example, they argue that the GSP auction is more efficient at allocating resources and more stable when it comes to bids with the GFP auction.

The above argument ignored the dynamic bidding behavior in the GSP auction because of the nature of the "envy-free" equilibrium concept proposed by Edelman et al. (2008) and Varian (2006). The bid retaliation is assumed out of the equilibrium path, however, in reality, the GSP auction does display a dynamic bidding pattern.

One famous example is "bid jamming". Sponsored search auction experts frequently suggest the use of a strategy called "bid jamming". Bid jamming happens when advertiser B bids 1 cent below his competitor, A, in an effort to drain up A's budget. Of course, this behavior will cause a back fire and A may drop his bid 1 cent below B's. Then B might further drop his and another kind of bidding war starts. This strategy was not made up by economists. Indeed, it has already been programmed into auto bidding softwares and is a strategy that is "actually widely-used". (See Ganchev et al. (2007).)

If we still take the above example, but allow bidders to use a bid jamming strategy, the equilibrium picture will be totally changed. Suppose bidder C bids at \$2. Suppose A adopts the bid jamming strategy and bids at \$3.99. If Bidder B retaliates and submits his bid at \$3.98, then the bidding war starts. The bids will fall to \$2 and then rise back up to \$4 again.

Now let us examine whether A and B have incentive to engage in this bidding war when bid jamming strategy is available. If A and B just stick to the envy free equilibrium strategies, A will receive $400 * (5 - 4) = 400$ profit and B will receive $100 * (4 - 2) = 200$ profit. If A and B engage in the bidding war and they split the highest slot half and half, A will receive $0.5 * 400 * (5 - 3) + 0.5 * 100 * (5 - 2) = 550$ while B will receive $0.5 * 400 * (4 - 3) + 0.5 * 100 * (4 - 2) = 300$. Both of them will be better off.

The only loser will be the search engine. There will be efficiency loss because the high value bidder does not get the higher position all the time. The loss will be the profit bidder A should have received if A had been higher than B. Therefore, the total social loss will be $0.5 * 400 * (5 - 4) = 200$.

As illustrated in the above example, the bidding behavior in the GSP auction is not necessarily more stable without further theoretical analysis. Actually, the dynamic bidding is very complicated and dynamic equilibrium does not have to be unique. However, constructing a theory to formally compare the two auctions is not the purpose of this paper. This research only attempts to examine the performance difference of the two auctions from an empirical perspective.

In the following research, I will try to test the above conventional wisdom by estimating how much more stable the bidding behaviors in the GSP auction are and how much more efficient the GSP auction is. To be specific, if the above belief holds, we should have the following hypothesis: Because the "second-price structure makes the market less susceptible to gaming" (See Edelman et al. (2008)), on average, an individual bidder in the GSP auction

Table 1: Bid Statistics of the top 10 Most Clicked Markets

Market	Observations	mean	std dev	min	Max
1	2,286,978	13.66	1.67	0.05	49
2	3,075,005	7.95	1.40	0.05	41.13
3	46,706	5.57	1.73	0.05	11
4	6,344	22.01	10.43	0.05	100
5	1,477,566	14.48	1.19	0.05	33
6	46,980	4.98	1.38	0.05	22
7	7,198	21.62	6.34	0.05	100
8	7,493	18.30	7.37	0.05	100
9	15,724	5.33	3.62	0.05	21.01
10	6,764	23.32	8.44	0.05	100

Note: there are 18,634,347 bids collected from 1000 markets in the sample.

will change his bid less frequently and his bid range will be smaller. In addition, because the GSP auction can more efficiently allocate the resources, the bidder with a higher value will obtain the better slots more often.

4 Data

Yahoo!'s research department provides a data set, which records all of the bids for the top 1000 keyword search by volume and all of the associated accounts for the time period from June 15, 2002 through June 14, 2003.

Each observation in the data has 5 variables: bidder ID, bidder's bid, the time when the bid was submitted, auction market and a dummy variable indicating whether the bid was placed under the GFP auction rule or under the GSP auction rule.

Table 1 shows the market statistics: the max bid, mean bid, minimum bid and the standard deviation for the top 10 most clicked markets. Five cents is the minimum requirement for bidding. One striking observation is the value of the maximum bid. According to this data set, some bidder is paying Yahoo! $100 for just one click through the sponsored search.

Because this paper is using RDD to estimate the local average causal effect of the GSP auction on the individual bidder, we also present the individual bidding statistics from June 15, 2002 through July 15, 2002 in Table 2. Table 2 provides the maximum value, mean value, minimum value and the standard deviation for the following daily statistics:

- Bid frequency: the number of times that an individual bidder changes his bid each day.
- Bid range: the difference of the maximum bid and the minimum bid of each bidder on each day.
- The Maximum bid, 75 percentile bid, mean bid, median bid, and 25 percentile bid of each bidder on each day.

Bid frequency and bid range measure the bidding stability of the auction system. The maximum bid, 75 percentile bid, median bid, mean bid and the 25 percentile bid measure the impact on the bid distribution of an individual bidder.

Table 2 shows how the mean values of the above statistics change after the launch of the new auction. The mean values of both the daily bid frequency and the daily bid range increase, which suggests that the new auction system is more unstable. The mean values of the max bid and 75 percentile bid increase while the mean value of the 25 Percentile bid decreases, which suggests that the bids are more dispersed.

It is impossible to plot the statistics because of the huge amount of data. To further show the big picture of how the change of auction systems affects bidding behaviors, we run a simple OLS regression first. In the OLS regression, we control for the market fixed effect and weekday effect. We also control for bidders' entry and exit by focusing on bidders who submit bids both before and after the auction rule change.

The OLS regression examines the percentage change of the variable $y_{i,m,t}$ in the following three months from July 2002 to September 2002.

$$log(y_{i,m,t}) = \mu_m + \sum_{j=1}^{3} \alpha_j \cdot I_j(t) + \sum_{d \in \{T,W,Th,F,Sa,Su\}} \beta_d \cdot I(t = d) + u_{i,m,t}$$

Table 2: Summary Statistics from June 15th to July 5th

Before June 25th	Mean	Stv	Min	Max
Bid Frequency	20.9	108	1	4,934
Bid Range	0.617	1.55	0	48.99
Max Bid	2.66	3.69	0.05	100
75 percentile bid	2.74	3.89	0.05	100
Median Bid	2.54	3.59	0.05	100
Mean Bid	2.52	3.55	0.05	100
25 Percentile bid	2.38	3.48	0.05	100
After June 25th	Mean	Stv	Min	Max
Bid Frequency	23.3	143	1	6,011
Bid Range	0.983	3.33	0	49.95
Max Bid	3.02	4.64	0.05	100
75 percentile bid	2.82	4.10	0.05	50
Median Bid	2.59	3.75	0.05	50
Mean Bid	2.56	3.62	0.05	42.8
25 Percentile bid	2.30	3.42	0.05	50

Note: there are 1,099,781 bids collected from 812 markets.

$y_{i,m,t}$ is the statistic of interest for bidder i, market m at time t. u_m is the market fixed effect. $I_j(t)$ is an indicator function, showing whether the time t is in the jth month after the policy change. Therefore, α_1 measures the impact of the new auction system on the market in the first month after the policy change. α_2 measures the impact in the second month after the policy change, and so on. β_d is the weekday dummy for Tuesday through Sunday.

Table 3 provides the estimation results, showing the impact of the new auction on bidding behaviors after June 26, 2002. First, both the daily bidding frequency and the daily bid range increase after the auction rule change. This result shows that the bidding behavior is more unstable under the GSP auction in contrast to the prevalent belief which suggests the opposite result.

Second, the individual daily bid distribution expends as the percentage changes of the daily maximum bid, mean bid and median bid are bigger than that of the daily 25 percentile

Table 3: The Change of the Statistics in Three Months

	α_1	α_2	α_3
Bid Frequency	14.5%	28%	18.8%
	(0.008)	(0.009)	(0.010)
Bid Range	14.0%	14.4%	15.2%
	(0.013)	(0.013)	(0.014)
Max Bid	10.3%	12.7%	15.1%
	(0.007)	(0.007)	(0.008)
75 percentile Bid	9.7%	12.1%	14.6%
	(0.007)	(0.007)	(0.008)
Mean Bid	9.1%	10.9%	13.6%
	(.007)	(.007)	(.008)
Median Bid	8.8%	10.8%	13.7%
	(.007)	(.007)	(0.008)
25 percentile Bid	6.7%	8.2%	11.1%
	(0.007)	(0.008)	(0.008)

Note: there are 5,877,945 bids collected from 833 markets.

bid.

However, the unobserved heterogeneities will make these OLS estimations biased, even misleading. First, the OLS regression does not control for the competition brought by bidders' entry and exit in each market, which is also impossible to do. Less amount of bidders might make the bidding behavior less aggressive. This may bring bias to the estimation of the average daily bidding frequency and the daily bidding range.

Second, the OLS regression cannot control for the competition among search engines markets. During that time period, around 2002, Google's popularity was beginning to rise, becoming more and more popular and grabbing more and more sponsored search market shares. Bidders often had accounts in both search engines. The implication, thus, would be that bidders might have been transferring more resources to Google accounts and paying less attention to their Yahoo! searches. This might also have caused bidders to bid less aggressively, which would affect our OLS estimation.

Moreover, the OLS regression cannot control for many unobserved variables, which play an important role in the bidding strategies, such as bidders' budget, the conversion rate of purchases, and so on.

Finally, the OLS regression does not control for the portion of bidders adopting the GSP auction, which might result in a lower estimation of the effects.

The bottom line here is that although the above regressions present the big picture of the bidding behavior change and suggest that the GSP auction does not bring stability to the system, the story is not convincing as there are so many unobserved variables which might plague the estimation. Therefore, to identify the true average effect, in the following section we use a RDD approach to avoid the above impacts of the unobserved heterogeneities.

5 Model

The basic idea behind the RDD model is to exploit the sharp jump shown in Figure 3 to identify the treatment effect of the GSP auction, which is the performance difference between the two auction systems.

Let $y_i(x_i, t)$ denote the statistics of interest for individual i at time t. x_i is a variable vector including all other characteristics such as market dummy, and weekday dummy.

After June 26, 2002, under Yahoo!'s new auction rules, each bidder could choose either the GFP auction or the GSP auction to submit his bids. Let $y_i(S, x_i, t)$ denote the statistics when the bidder is submitting the bid through the GSP auction; if the bid is submitted through the GFP auction, the statistics will be denoted by $y_i(F, x_i, t)$.

Then the statistics y_i can further be rewritten as

$$\begin{aligned} y_i(x_i, t) &= y_i(S, x_i, t)I_i(\text{GSP}) + y_i(F, x_i, t)I_i(\text{GFP}) \\ &= \beta(x_i, t) + \alpha_i(x_i, t) \cdot I_i(\text{GSP}) + u_i(x_i, t) \qquad (1) \end{aligned}$$

where

$$\begin{aligned} y_i(F, x_i, t) &= E[y_i(F, x_i, t)] + u_i(x_i, t) \\ &= \beta(x_i, t) + u_i(x_i, t) \end{aligned}$$

and $\alpha_i(x_i, t) = y_i(S, x_i, t) - y_i(F, x_i, t)$. $I(\text{GSP})$ is the indicator function of whether the bidder is choosing the GSP auction and $I(\text{GFP})$ is the indicator function of whether the bidder is choosing the GFP auction. $E[\alpha_i(x_i, t)]$ is the average treatment effect, which is what we want to estimate.

Because the choice of the GSP auction is endogenous, the new auction system will be a combination of both the GFP auction and the GSP auction. It is interesting to notice that choosing the GSP auction would dominate the choice of the GFP auction because of the lower payment while maintaining the same slot. However, from the data, we observe that a big portion of the bidders still choose the GFP auction, which suggests that unobserved heterogeneity was playing an important role. Besides, this new auction is a very complicated game system and no theory has been conducted on it yet. Therefore, to simply the problem, we make the following assumption about bidder behaviors before and after the auction rule change.

Assumption 1 *(**Behavioral Assumption**): Bidders submit bids either in a pure GFP auction system or a pure GSP auction system.*

The above assumption simplifies auction bidding behavior in this extreme complicated environment and enables us to identify the average treatment effect $[\alpha_i(x_i, t)]$. When the literature discusses this part of the history of the sponsored search auction, it usually ignores this endogeneity of the GSP auction choice. (See Jansen and Mullen (2008).) If we also maintain the same assumption that every bidder was submitting as if he was in a GSP action, the probability of receiving treatment of the GSP auction will jump to 1 for everyone. This assumption will even simplify the estimation. We will leave this no endogeneity case to the estimation section, and in the following, we will allow the endogeneity of the auction choices.

Assumption 2 ***(Continuity Assumption)***: *$E[y_i(S, x_i, t)|t]$ and $E[y_i(F, x_i, t)|t]$ are continuous in t at $\bar{t}$. $\bar{t}$ is the beginning time when the new auction was launched.*

This continuity assumption assumes that the bidding statistics near the critical value are continuous. In other words, bidders would have continued to behave as they would have before the auction rule change as if there had been no auction switch. Therefore, any bidding behavior change will be attributed to the treatment, or the launch of the GSP auction.

It is worth emphasizing that $I(\mathrm{GSP})$ is an endogenous variable which is affected by the unobserved heterogeneity. The GSP auction should dominate the GFP auction, because without changing the slot placement, choosing the GSP auction makes the bidder pay less than choosing the GFP auction. However, data shows that not all bidders chose the GSP auction after the auction upgrade. Whether bidders choose the GSP auction or the GFP auction was determined by their stochastic process, which is not observed by economists. However, the following proposition shows that we can still have the identification.

Proposition 1 *Under assumption 1 and 2,*

$$E[\alpha_i(x_i, t)|I(GSP) = 1] = \frac{\lim_{t\downarrow\bar{t}} E[y_i(x_i, t)|t] - \lim_{t\uparrow\bar{t}} E[y_i(x_i, t)|t]}{\lim_{t\downarrow\bar{t}} E[I_i(GSP)|t]}$$

Proof. The proof follows Imbens and Angrist (1994).

Pick two numbers $t_0 < \bar{t} < t_1$, we have

$$\begin{aligned}
& E[y_i(x_i,t)|t_1] - E[y_i(x_i,t)|t_0] \qquad (2)\\
= \; & E[y_i(S,x_i,t)I(\text{GSP},t) + y_i(F,x_i,t)I(\text{GFP},t)|t = t_1]\\
& -E[y_i(S,x_i,t)I(\text{GSP},t) + y_i(F,x_i,t)I(\text{GFP},t)|t = t_0]\\
= \; & E[y_i(S,x_i,t_1)I(\text{GSP},t_1) + y_i(F,x_i,t_1)(1 - I(\text{GSP},t_1))]\\
& -E[y_i(S,x_i,t_0)I(\text{GSP},t_0) + y_i(F,x_i,t_0)(1 - I(\text{GSP},t_0))]
\end{aligned}$$

By assumption 2 and equation 2, we will have

$$\begin{aligned}
& \lim_{t\downarrow\bar{t}} E[y_i(x_i,t)|t] - \lim_{t\uparrow\bar{t}} E[y_i(x_i,t)|t]\\
= \; & \lim_{t\downarrow\bar{t}} E[y_i(S,x_i,\bar{t})I(\text{GSP},t) + y_i(F,x_i,\bar{t})(1 - I(\text{GSP},t))] - y_i(F,x_i,\bar{t})\\
= \; & \lim_{t\downarrow\bar{t}} E[I_i(\text{GSP})(y_i(S,x_i,t) - y_i(F,x_i,t))]\\
= \; & \lim_{t\downarrow\bar{t}} \Pr[I_i(\text{GSP}) = 1] \cdot \lim_{t\downarrow\bar{t}} E[y_i(S,x_i,t) - y_i(F,x_i,t)|I(\text{GSP}) = 1]\\
= \; & \lim_{t\downarrow\bar{t}} \Pr[I_i(\text{GSP}) = 1] \cdot \lim_{t\downarrow\bar{t}} E[\alpha_i(x_i,t) + u_i(x_i,t)|I(GSP) = 1]
\end{aligned}$$

Therefore

$$E[\alpha_i(x_i,t)|I(GSP) = 1] = \frac{\lim_{t\downarrow\bar{t}} E[y_i(x_i,t)|t] - \lim_{t\uparrow\bar{t}} E[y_i(x_i,t)|t]}{\lim_{t\downarrow\bar{t}} E[I_i(GSP)|t]} \qquad (3)$$

■

This proposition provides the foundation for the identification strategy. Because the continuity assumption is addressing local properties, this proposition is also about the local properties. It is also worthwhile pointing out that equation 3 does not identify the average average causal effect $E[\alpha_i(x_i,t)]$ of the whole population. Instead, it identifies the local average $E[\alpha_i(x_i,t)]$ of the bidders who chose to submit bids through the GSP auction, making them a subgroup of the population.

Another thing worth pointing out is that in the traditional RDD theory, to have this identification result, we must have both the above continuity assumption and the local independence assumption. (See Hahn et al. (2001) and Van der Klaauw (2007).) This proposition, however, only requires the continuity assumption. This is because the treatment variable here is time, instead of a random variable which might be correlated with the decision of treatment I(GSP). Therefore, although I(GSP) is endogenously affected by unobserved heterogeneity, the identification assumption only requires the continuity assumption.

Lastly, if we go back to the assumption that the the auction system was affecting all the bidders in the markets in the same way, no matter what auction system they appeared to choose and that each bidder was bidding as if they were in a pure GSP auction system, then this assumption would rule out the endogeneity of the GSP auction and $\lim_{t \downarrow \bar{t}} E[I_i(GSP)|t] = 1$. Therefore, equation 3 will give an identification result for the whole population. In the estimation section, this No GSP Auction Endogeneity case is also estimated along with the case allowing bidders to endogenously choose the auction system.

6 RDD Estimation

The RDD estimation in this paper follows the standard nonparametric regressions. Imbens (2007) and Van der Klaauw (2007) have very good surveys for the literature of RDD, especially the estimation methods used in RDD.

The asymptotical boundary properties of the standard kernel estimator is not ideal because of the poor convergence rate, as pointed out by Hahn et al. (2001) and Porter (2003). Therefore, we consider the local linear regression method proposed by Fan and Gijbels (1996).

Let α_y and β_y solve the following minimization problems for the numerator:

$$\min_{\alpha_{yl},\beta_{yl},\alpha_{yr},\beta_{yr}} \sum_{i|\bar{t}-h<t_i<\bar{t}} (y_i - \alpha_{yl} - \beta_{yl}\cdot(t_i-\bar{t}) - \delta\cdot X_i)^2 + \sum_{i|\bar{t}<t_i<\bar{t}+h} (y_i - \alpha_{yr} - \beta_{yr}\cdot(t_i-\bar{t}) - \delta\cdot X_i)^2$$

Here h is the bandwidth on either side of the discontinuity point. X_i is the covariate vector, which in the estimation includes the market dummy variable and the weekday dummy variable.

For the denominator, let α_p and β_p solve

$$\min_{\alpha_p,\beta_p} \sum_{i|\bar{t}<t_i<\bar{t}+h} (I_i(GSP) - \alpha_p - \beta_p\cdot(t_i-\bar{t}))^2$$

then the estimator for the average causal effect will be $\hat{\tau} = \frac{\hat{\alpha}_{yr}-\hat{\alpha}_{yl}}{\hat{\alpha}_p}$.

To estimate the bidding behavior difference, we try 2 different bandwidths, 11 days and 6 days around the discontinuity point on June 26, 2002. We also consider the case assuming the GSP auction choice to be endogenous, and no auction endogeneity case assuming every bidder bidding under the GSP auction.

The bidding statistics examined here include the daily bid changing frequency, the daily bid range, the daily maximum bid, 75 percentile bid, mean bid, median bid, 25 percentile bid and the daily individual bidder payment. For the last statistics, because we can not observe the click-through-rate on each bidder's link, we simply assume every bidder received one unit of clicks in every 15 minutes.

Lastly, we estimate the bidding behavior change using both the absolute value and the log value of the statistics.

Table 4 shows the RDD estimation results, which are consistent with the OLS regression results shown in Table 3. Column 1 shows the individual daily bidding frequency increased 6.8 times which represents a 36% increase. For each individual bidder, his bid range also

Table 4: RDD Estimation Results

	h=11		h=6	
	Absolute	Relative	Absolute	Relative
Bid Frequency	6.80	35.9%	6.09	50.9%
	(10.62)	(0.116)	(11.78)	(0.125)
Bid Range	1.313	85.1%	1.78	125%
	(0.239)	(0.118)	(0.302)	(0.149)
Daily Payment	-21.76	-67.9%	-22.1	-64.38%
	(4.26)	(3.22)	(5.01)	(0.086)
Max Bid	0.551	6.1%	0.840	17.3%
	(0.281)	(0.052)	(0.329)	(0.072)
75 percentile bid	0.09	-0.002%	0.233	9.8%
	(0.256)	(0.040)	(0.340)	(0.112)
Median Bid	-0.209	-6.11%	-0.164	0.98%
	(0.282)	(0.057)	(0.307)	(0.104)
Mean Bid	-0.190	-4.03%	-0.158	3.7%
	(0.250)	(0.047)	(0.287)	(0.061)
25 Percentile bid	-0.537	-17.9%	-0.609	-13.9%
	(0.203)	(0.071)	(0.271)	(0.080)
Treatment Probability Jump	0.542		0.505	
	(0.005)		(0.006)	

Notes: There are 1,099,781 bids collected from 812 markets. Absolute measures the absolute value change; Relative measures the percentage change. h is the bandwidth taking value of 11 days and 6 days respectively. To analyze payment, we normalize the click on the first slot to one. The click declining rate follows the Brooks (2005). Therefore, the relative change for the daily payment is more meaningful.

increased by 131 cents, or 85% in relative value. All suggest that the GSP auction did not increase the stability of the system.

The estimation results also present how an individual bidder's payment changed each day after the new auction was launched. Table 4 shows that each bidder's average daily payment decreased by about 70%. One of the reasons that this number is so big might be that we cannot observe the actually click through rates on each slot and have to calibrate the numbers from Brooks (2005). This might bring bias to the magnitude of the estimation.

Table 4 also presents how an individual bidder's bid distribution changed after the launch of the GSP auction. Column (2) shows that the maximum bid and 75 percentile bid tended to increase and the mean bid, median bid and the 25 percentile bid tended to decrease, which

Table 5: RDD Estimation Results: No Auction Endogeneity Case

	h=11		h=6	
	Absolute	Relative	Absolute	Relative
Bid Frequency	3.671	19.4%	3.110	26%
	(2.031)	(0.021)	(2.284)	(0.024)
Bid Range	0.709	46%	0.91	64%
	(0.006)	(0.006)	(0.053)	(0.039)
Daily Payment	-11.40	-36.7%	-11.28	-32.9%
	(0.62)	(1.42)	(0.25)	(0.030)
Max Bid	0.298	3.3%	0.430	9.0%
	(0.051)	(0.017)	(0.066)	(0.022)
75 percentile bid	0.049	-0.001%	0.119	5.0%
	(0.046)	(0.017)	(0.061)	(0.019)
Median Bid	-0.113	-3.3%	-0.084	0.5%
	(0.042)	(0.017)	(0.048)	(0.020)
Mean Bid	-0.106	-2.2%	-0.078	1.9%
	(0.040)	(0.017)	(0.047)	(0.019)
25 Percentile bid	-0.287	-9.7%	-0.309	-7.1%
	(0.039)	(0.019)	(0.050)	(0.021)

Notes: There are 1,099,781 bids collected from 812 markets. Absolute measures the absolute value change; Relative measures the percentage change. h is the bandwidth taking value of 11 days and 6 days respectively. To analyze payment, we normalize the click on the first slot to one. The click declining rate follows the Brooks (2005). Therefore, the relative change for the daily payment is more meaningful.

is also consistent with the increase of bid range.

Lastly, Figure 4 plots how each of the statistics evolves before and after June 26, 2002. We fit the estimation results with a smooth function.

When Yahoo! launched the new auction system, they named the new bid "maximum willingness to pay", hoping bidders would simply bid their highest possible payment. They hoped this would reduce the instability of the system and increase Yahoo!'s revenue. But the above results suggest these goals were not achieved. Instead of reducing the strategic behaviors, bidders submitted their bids in a bigger range and changed their bids at a higher frequency. All this suggests that the bidders were actually taking advantage of the GSP auction system and were more likely to "game the system".

Figure 4: The Statistics Before and After June 26, 2002

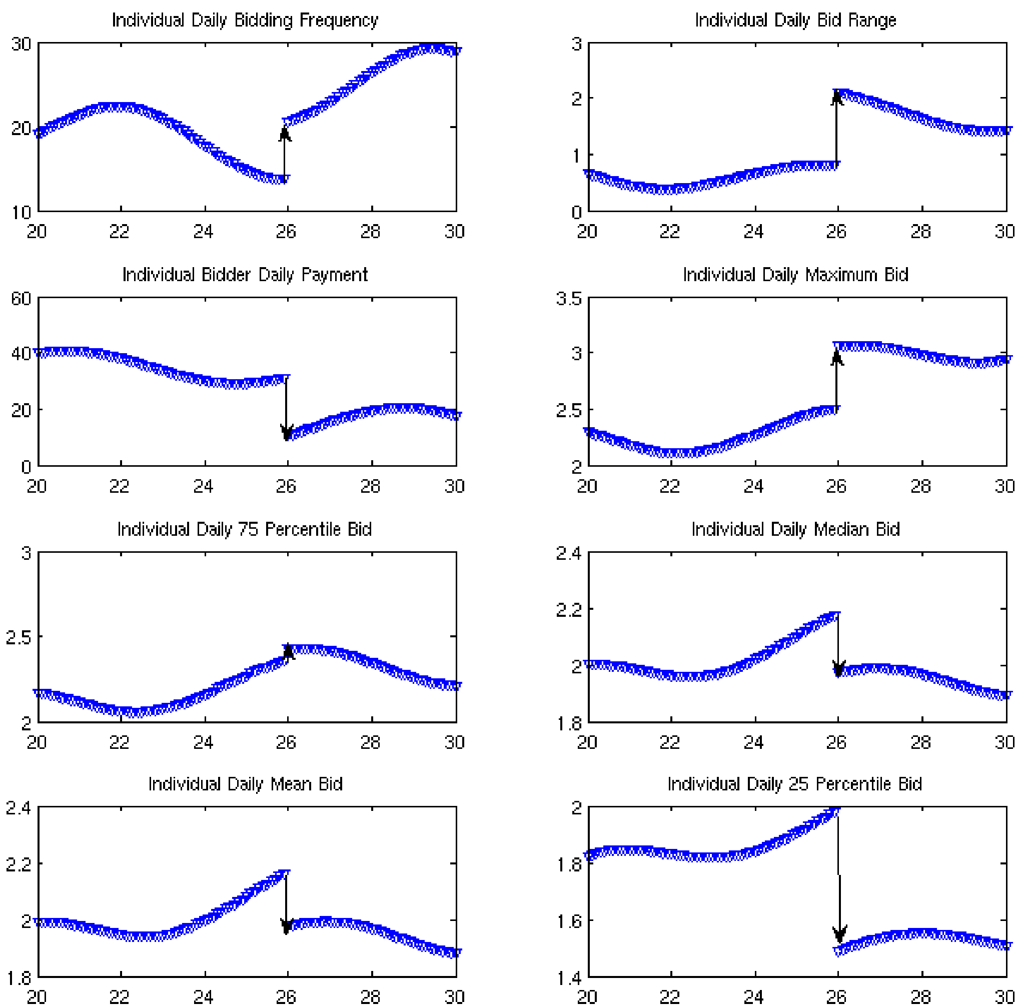
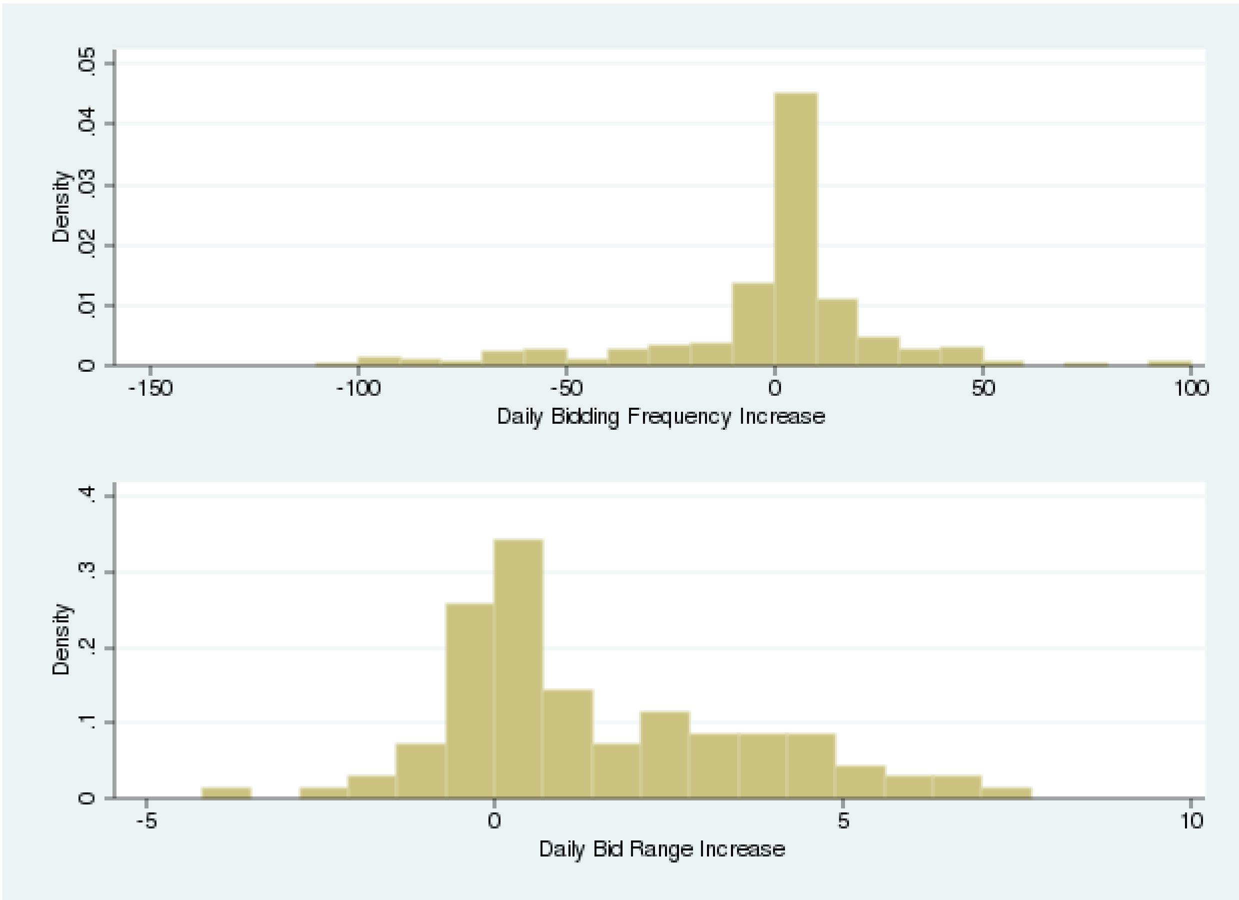


Figure 5: The Histograms of the Daily Biding Frequency Increase and the Daily Bid Range Increase across Markets



The above estimation provides the average effect across the markets. We also conduct the RDD estimation of the individual daily bidding frequency change and daily bid range for each market m. Figure 5 provides the histograms of the RDD estimation results. By looking at the graph, we can see most of daily bidding frequency increase reside between 0 and 10 times. The mean and median are also between 0 and 10. Meanwhile, the histogram of the daily bid range increase shows that in most of the markets bidders increase their daily bid ranges and the mean and median value are around \$1. All these results are consistent with the previous RDD estimation results and suggest the robustness of the above RDD estimation regarding the daily bidding frequency increase and the daily bid range increase.

We also estimate the average treatment effect of the no endogeneity case. Because in this case, we assume that all the bidders receive the treatment of the GSP auction, as discussed in section 5, the magnitude of the estimation must be smaller than in the case which assumes that bidders could endogenously choose the auction system. The estimation results are shown in the Table 5. It is worth noting, however, that the interpretation of the results in this case would be a little bit different: the estimation results here are the average effect of the whole population, instead of the local average effect of a subgroup of the population as in the previous case.

One possible factor, which might have an impact on the above estimation, is the learning. If bidders were testing and learning the new auction system, the estimation of the bidding frequency difference may have biases. However, before Yahoo! launched the GSP auction in June, Google had already started its GSP auction in April. (See Jansen and Mullen (2008).) Because serious advertisers would have accounts in both search engines' sponsored search auctions, it would, therefore be, reasonable to assume that learning was unnecessary for the bidders when they had the Yahoo! GSP auction and did not play a role in the bidding behavior.

Trying to interpret why the GSP auction is more unstable than the GFP auction is dangerous here as this paper only provides evidence showing that the second price auction structure does not make the bidders less susceptible to gaming and not why the bidding behavior in the GSP auction is more volatile and aggressive. However, our conjecture is that second price auction structure in the GSP auction makes the bidders pay less, and therefore, the bidders have more resources to engage in strategic bidding behaviors. This might contribute to the estimation results above.

7 Efficiency Comparison

7.1 Model Setup

In this section we want to answer the question of whether or not there was any efficiency improvement under the GSP auction system, as claimed by the literature. To measure efficiency, we first construct an index measure based on the ranking.

Suppose there are two bidders, A and B. A's value per click is V_A and B's value per click is V_B. If the system is efficient and higher ranks receive more clicks, then $\Pr[A$ higher than $B] = 1$. If the auction mechanism is less efficient, this probability will be smaller than 1; the less efficient the mechanism is, the smaller the probability should be.

Therefore, this relative ranking between two bidders can be used as an index to measure the efficiency of the auction mechanism. Based on this efficiency index, the idea behind the identification is the following: If the system improves the bidding efficiency, it should make the winner more likely to win and the loser more likely to lose. In other words, the probability index bigger than $\frac{1}{2}$ should be even bigger than $\frac{1}{2}$ in the new auction system, and the probability index smaller than $\frac{1}{2}$ be even smaller than $\frac{1}{2}$ in the new system.

Given a unit of time, define λ_{AB} to be the portion of time that A ranks higher than B.[6] If $V_A \geq V_B$, because of the inefficiency of the GFP auction design or measurement error, λ_{AB} should be smaller than 1. This difference will reflect the efficiency loss.

Assumption 3 *If $V_A \geq V_B$, then $\lambda_{AB} = 1 - \alpha + u_{AB}$ with $\alpha < \frac{1}{2}$.*

Here α captures the efficiency loss caused by the GFP auction design and u_{AB} can be taken the measurement error, or a random shock. Assumption 3 also implies that in the GFP auction, although the bidder with the low value might take advantage of the auction design and sometimes dominate his competitor, this should not happen over 50% the time. In other words, the bidder with the higher value should get the higher position more often.

[6] The unit of time can be an hour, a day, etc.

Assumption 4 *$Eu_{AB} = 0$ and u_{AB} is iid. Its distribution function is denoted by $F(u)$.*

The literature claims that the GSP auction improves efficiency. Therefore, α will decrease according to the prediction.

Assumption 5 *Under the GSP auction, the observed frequency is*

$$\lambda_{AB} = (1 - \alpha) + \beta + u_{AB}$$

Therefore, the new function is

$$\lambda_{AB} = (1 - \alpha) + \beta * I(GSP) + u_{AB} \tag{4}$$

We can not observe V_A or V_B, therefore we do not know which is bigger if we just randomly pick any bidders as A and B. The estimation of $\widehat{\beta}$ will be meaningless if we simply regress the equation 4.

Therefore, the empirical question is how to estimate β. The following propositions show the estimation strategy, which is discussed at the beginning of the section.

Proposition 2 *Let N be the number of the observations. Define $\eta_{AB} = \max\{\lambda_{AB}, 1-\lambda_{AB}\}$. Regress $\eta_{AB} = \gamma_\alpha + \gamma_\beta * I(GSP) + u_{AB}$. Then the OLS result provides a lower bound for β. That is $\lim_{N\to\infty} \hat{\gamma_\beta}_N = \beta_\infty < \beta$*

Proof. *: By OLS, it can be shown that*

$$\begin{aligned}
\beta_\infty &= E\{\max\{(1-\alpha)+\beta*I(GSP)+u_{AB}, 1-((1-\alpha)+\beta*I(GSP)+u_{AB})\}\} \\
&\quad -E\{\max\{(1-\alpha)+u_{AB}, 1-((1-\alpha)+u_{AB})\}\} \\
&= (1-\alpha)+\beta+2\int(-u-(\frac{1}{2}-\alpha+\beta))I(u<-(\frac{1}{2}-\alpha+\beta))dF(u) \\
&\quad -((1-\alpha)+2\int(-u-(\frac{1}{2}-\alpha))I(u<-(\frac{1}{2}-\alpha))dF(u)) \\
&= \beta+2\int(-u-(\frac{1}{2}-\alpha+\beta))I(u<-(\frac{1}{2}-\alpha+\beta))dF(u) \\
&\quad -2\int(-u-(\frac{1}{2}-\alpha+\beta))I(u<-(\frac{1}{2}-\alpha))dF(u) \\
&\quad -2\beta\int I(u<-(\frac{1}{2}-\alpha))dF(u) \\
&= \beta-2\beta\int I(u<-(\frac{1}{2}-\alpha))dF(u) \\
&+ \ 2\int(u+(\frac{1}{2}-\alpha+\beta))I(-(\frac{1}{2}-\alpha+\beta)<u<-(\frac{1}{2}-\alpha))dF(u)
\end{aligned}$$

Because $I(-(\frac{1}{2}-\alpha+\beta)<u<-(\frac{1}{2}-\alpha)) \leq I(u<-(\frac{1}{2}-\alpha))$ *and* $u+(\frac{1}{2}-\alpha+\beta)\leq\beta$ *when* $u<-(\frac{1}{2}-\alpha)$

Therefore $2\beta\int I(u<-(\frac{1}{2}-\alpha))dF(u) > 2\int(u+(\frac{1}{2}-\alpha+\beta))I(-(\frac{1}{2}-\alpha+\beta)<u<-(\frac{1}{2}-\alpha))dF(u)$

Therefore $\beta_\infty<\beta$ ■

Proposition 3 *Let* $N=N_1*N_2$. *Define* $\eta_{AB,\ N_1}=\max\{\frac{\Sigma\lambda_{AB}}{N_1}, 1-\frac{\Sigma\lambda_{AB}}{N_1}\}$. *Regress* $\eta_{AB,\ N_1}=\gamma_\alpha+\gamma_\beta*I(GSP)+u_{AB}$. *Then the OLS result provides a consistent estimate. That is* $\lim_{N_1,N_2\to\infty}\hat{\gamma_\beta}_{N_1,N_2}=\beta_\infty=\beta$

Proof. *: By OLS, it can be shown that*

$$\begin{aligned}
\beta_\infty &= \lim_{N_1\to\infty}\beta+2\int(-\frac{\Sigma u_{AB}}{N_1}-(\frac{1}{2}-\alpha+\beta))I(\frac{\Sigma u_{AB}}{N_1}<-(\frac{1}{2}-\alpha+\beta))\Pi(dF(u)) \\
&- \int(-\frac{\Sigma u_{AB}}{N_1}-(\frac{1}{2}-\alpha))I(\frac{\Sigma u_{AB}}{N_1}<-(\frac{1}{2}-\alpha))dF(u)
\end{aligned}$$

Table 6: Estimation Result of Relative Ranking Change

	Active Bidders	All Bidders
β	0.038	0.037
	(0.0027)	(0.0028)
N of Obs	15,316	15,343

As $Eu_{AB} = 0$ *and* u_{AB} *is iid,* $\frac{\Sigma u_{AB}}{N_1} \to 0$*. And* $|-\frac{\Sigma u_{AB}}{N_1} - (\frac{1}{2} - \alpha + \beta) < 2$*, therefore*

$$
\begin{aligned}
|\int(-\frac{\Sigma u_{AB}}{N_1} - (\frac{1}{2} - \alpha + \beta))I(\frac{\Sigma u_{AB}}{N_1} &< -(\frac{1}{2} - \alpha + \beta))\Pi(dF(u))| \\
&\leq 2\int I(\frac{\Sigma u_{AB}}{N_1} < -(\frac{1}{2} - \alpha + \beta))\Pi(dF(u)) \\
&\to 0
\end{aligned}
$$

Therefore $\beta_\infty = \beta$. ■

7.2 Estimation

We first randomly pick an auction market and then select two bidders as A and B in this market. From June 15, 2002 to July 21, 2002, we randomly choose 500 pairs. Second, we calculate λ_{AB} for each day. Next, we define $\eta_{AB} = max\{\lambda_{AB}, 1 - \lambda_{AB}\}$. Then by the above Propositions, the following regression will provide a lower bound for the efficiency improvement:

$$\eta_{AB} = \alpha_{AB} + \beta * I(\text{GSP}) + \gamma_{day} * D_{day} + u_{AB}$$

Here we control for the pair fixed effect α_{AB}, weekday effect γ_{day}.

We estimate the efficiency improvement for two cases. The first case includes all the bidders, and the second case only includes active bidders who change their their bids at least 400 times everyday. Table 6 shows the efficiency improvement brought by the launch of the new auction. The value of $\hat{\beta}$ suggests that after the new auction launched, the bidder with the higher value was more likely to dominate the lower-value bidder and that this probability increased by around 4%. $\hat{\beta}$ is positive, therefore, it is consistent with the literature that the GSP auction is more efficient than the GFP auction. But the magnitude is not significantly large.

For active bidders, the estimation result is smaller, which means there is not much change in the relative rankings after the launch of the GSP auction. This suggests that the active bidders might still engage in strategic bidding behavior, which is consistent with the results in the RDD section.

8 Conclusion

The evolution of sponsored search auctions is an important and interesting phenomenon. Having a deep understanding about different sponsored search auctions, especially the performance differences, can help us design superior auctions in the future.

When Yahoo! launched the GSP auction, their purpose was to bring a more stable, more profitable and more efficient auction. People in the industry and academia did expect that the bidders would be less likely to "game the system" and that the new auction system would bring Yahoo! more revenue. One important factor to note was that at that time, Google, which was rising in popularity, was adopting the GSP auction, and that Yahoo! wanted to copy Google's success.

However, this paper provides solid evidence suggesting that under the new system, instead of being more stable, bidders tended to update their bids more frequently, their individual bid range tended to be bigger, and that Yahoo!'s revenue shrunk after the launch of the GSP auction. This is in contrast with Google's success. One of the key differences between these two is that Google was not only using the GSP auction, but also a different slot allocation rule. Google was using a score system created by itself, which depended on both bidders' bids and their web link qualities, to allocate the web link placements instead of just their bids. This score system made the bid manipulation play a less important role in determining a bidder's rank. In other words, it was more difficult for an individual bidder to manipulate his slot allocation just by frequently changing bids. Instead, in the Google sponsored search auction the incentive for a bidder was to improve his product's quality in order to obtain higher position by improving his score. We conjecture this is the key difference which makes Goolge more successful. In 2007, Yahoo! also adopted this score system. This also suggests that the GSP auction may not be as superior as most of the conventional wisdom believes, and that the score system probably plays a crucial role in improving the GFP auction. These conjectures will be left to the future research.

References

Angrist, Joshua and Victor Lavy, "Using Maimonides' Rule to Estimate the Effect of Class Size on Scholastic Achievement," *The Quarterly Journal of Economics*, 1999, *114*, 533–575.

Animesh, Animesh, Vandana Ramachandran, and Siva Viswanathan, "Quality Uncertainty And Adverse Selection in Sponsored Search Markets," 2005. Working Paper.

Athey, Susan and Glenn Ellison, "Position Auctions with Consumer Search," 2007. Working Paper.

Borgers, Tilman, Ingemar Cox, Martin Pesendorfer, and Vaclav Petricek, "Equilibrium Bids in Sponsored Search Auctions: Theory and Evidence," 2007. Working Paper.

Brooks, Nico, "The Atlas Rank Report: How Search Engine Rank Impacts Traffic," *Digital Marketing Insights*, 2005.

Chen, Susan and Wilbert Van der Klaauw, "The work disincentive effects of the disability insurance program in the 1990s," *Journal of Econometrics*, 2008, *142*, 757–784.

der Klaauw, Wilber Van, "Estimating the Effect of Financial Aid Offers on College Enrollment: A Regression-Discontinuity Approach," *International Economic Review*, 2002, *43* (4), 1249–1287.

——, "Regression-Discontinuity Analsis," *The New Palgrave Dictionary of Economics*, 2007.

Edelman, Benjamin and Michael Ostrovsky, "Strategic Bidder Behavior in Sponsored Search Auctions," *Decision Support Systems*, 2006.

——, ——, **and Michael Schwarz**, "Internet Advertising and the Generalized Second-Price Auction: Selling Billions of Dollars Worth of Keywords," *American Economic Review*, 2008.

Fan, Jianqing and Irene Gijbels, "Local Polynomial Modelling and Its Applications," 1996. Chapman and Hall: London.

Ghose, Anindya and Sha Yang, "An Empirical Analysis of Search Engine Advertising: Sponsored Search Auction and Cross-Selling in Electronic Markets," September 2007. NET Institute Working Paper.

Goldfarb, Avi and Catherine Tucker, "Search Engine Advertising: Pricing ads to Context," April 2008. Working Paper.

Hahn, Jinyong, Petra Todd, and Wilber Van der Klaauw, "Identification and Estimation of Treatment Effects With A Regression-Discontinuity Design," *Econometrica*, 2001, *69* (1), 201–209.

Imbens, Guido, "Regression Discountinuity Designs: A Guide to Proctice," April 2007. NBER Working Paper 13039.

—— **and Joshua Angrist**, "Identification and Estimation of Local Average Treatment Effects," *Econometrica*, 1994, *62* (2), 467–475.

Jansen, Bernard J. and Tracy Mullen, "Sponsored search: an overview of the concept, history, and technology," *Int. J. Electronic Business*, 2008, *6* (2), 114–131.

Lee, David, "Randomized experiments from non-random selectionin US House election," *Journal of Econometrics*, 2008, *142*, 675–697.

Porter, Jack, "Estimation in the Regression Discontinuity Model," 2003. Working Paper, Department of Economics, University of Wisconsin.

Rutz, Oliver and Randolph Bucklin, "A Model of Individual Keyword Performance in Paid Search Advertising," June 2007. Working Paper.

Varian, Hal R, "Position Auctions," 2006. Working Paper.

Zhang, Xiaoquan, "Finding Edgeworth Cycles in Online Advertising Auctions," 2005. Working Paper.