LEARNING FROM TURTLES[i]:
AN AGENT-BASED MODEL OF A GENERALIZED SECOND-PRICE AUCTION

By

Wenjuan Ma

A DISSERTATION

Submitted to
Michigan State University
in partial fulfillment of the requirements
for the degree of

Information and Media—Doctor of Philosophy

2016

ABSTRACT

LEARNING FROM TURTLES:
AN AGENT-BASED MODEL OF A GENERALIZED SECOND-PRICE AUCTION

By

Wenjuan Ma

A generalized second-price (GSP) auction is the standard way to allocate search advertising slots on a search results page. In order to acquire slots, advertisers submit bids for a search keyword. Each time a search user sends a search query, an auction begins. The search engine sorts the bids and concludes the auction with the slot allocation. Currently, advertisers do not have access to the bids submitted by their opponents. In this dissertation, I used agent-based modeling to simulate auctions under different information disclosure policies. I investigated three information disclosure policies: no information disclosure, partial information disclosure, and perfect information disclosure. Under the no information disclosure policy, a search engine does not disclose bid information. Under partial and perfect information disclosure policies, a search engine announces bid statistics and bids from the prior round respectively. The simulated auctions ran in different scenarios, which were formed by varying values of several parameters. My goal was to learn about the effects of bid information disclosure policies on search engine revenue and surplus generation from these simulations.

Through the simulations and analyses, I illustrated that a search engine can generate higher levels of revenue under the partial information disclosure policy than under the other two information disclosure policies. I also found that GSP auctions were relatively robust in terms of surplus generation.

Copyright by
WENJUAN MA
2016

This work is dedicated to my parents, Jinbao Ma and Yonglian Hu and to my husband Xinren Li.

ACKNOWLEDGEMETNS

I want to thank Steve Wildman for guiding me through my PhD program. Steve motivated me, shaped my views and my methods for solving research problems. The researcher I am today I owe to him. During the past few years, I cannot recall a time when Steve was not busy. However, he always managed to find time to talk with me over many questions – research related and more. Above all, I thank him for guiding me by setting examples. I am sure that his influence will reach beyond my PhD program.

I consider myself lucky to have attended such a wonderful place as Michigan State University, where I encountered many great mentors, who provided guidance and inspiration. I'd like to thank Rick Wash for pointing me in the right direction. Specifically, his big data class filled my mind with lots of interesting research questions. Many thanks to Johannes Bauer. He is one of the most knowledgeable scholars that I know of and I can never thank him enough for stepping in and taking over responsibilities as my advisor while Steve worked as the chief economist of FCC. Many thanks to Bob LaRose; I was fortunate to have the opportunity to collaborate on research projects with him. Bob taught me multiple ways of managing research projects and showed me his ingenious ways of problem solving. I also want to thank Emilee Rader for her insights and support. As many other graduate students I not only learned knowledge but also gained insights about academia from her. To many other faculty members that I had pleasure to be in contact with during my PhD years, many thanks.

Thanks are also due to all my CSTAT colleagues, including Brain Maurer, Steven Pierce, Sarah Hession, Frank Lawrence, and Dhruv Sharma, who provided me with suggestions about

statistical analyses, as well as warm friendships. I cannot count how many times they helped and inspired me with research, career and life. They showed me how to conduct myself as a statistician and researcher.

I also want to express my appreciations to a few fellow Phd, students who worked on their dissertations around the same time. They are CK, Yumi, Kanni Huang, Kang Li, Chen Lou, Dawn Chang, and Keyin Wang. Because of the shared experience, we were able to encourage one another, and provide hope to each other. As we are moving on to the new chapters of our lives I sincerely believe the comradery will carry on.

On a personal note, I want to thank my husband for his devotion and sacrifice. Our daughter, Naomi, was born when I was working on my dissertation. He became her major caregiver so I had time to work on my dissertation. As we move towards the next stage of our life together, I am sure I will find more ways to express my love and gratitude to him. I also want to express special gratitude and appreciation to my mom, who flew thousands of miles to help me with child care. She has always seen my potential and encouraged me to chase after my dreams.

TABLE OF CONTENTS

LIST OF TABLES ix

LIST OF FIGURES xi

INTRODUCTION 1

LITERATURE REVIEW 13
2.1 General Settings and Assumptions in Keyword Auction Research 14
2.2 The VCG Mechanism 16
2.3 Equilibria Identification and Refinement 17
2.4 Efficiency Loss of GSP auctions 20
2.5 Revenue Levels of GSP Auctions 22
2.6 Auction Instruments 23
2.6.1 Reserve Price 23
2.6.2 Ranking Rules 24
2.7 Value Estimation 24
2.8 Bidding Strategies 26
2.9 The Effects of Bid Information Disclosure Policy in Single-Item Auctions 27
2.10 Summary 28

RESEARCH DESIGN 31
3.1 The Goals for this Study and Research Questions 31
3.2 Research Method Comparison 31
3.3 Assumptions, Settings, and Scenarios for the Simulated GSP Auctions 35
3.4 Metrics for Measuring Performance 45
3.5 Analysis Plan 48

RESULTS 49
4.1 Revenue Ratio 49
4.2 Yield 56
4.3 Surplus Ratio 62

CONCLUSION AND DISCUSSION 71
5.1 Conclusion and Discussion 71
5.2 Limitations and Future Research Directions 76

APPENDICES 80
APPENDIX I 81
DENSITY, MEAN AND RELATIONSHIPS AMONG THE 10 VARIABLES FOLLOW THE VALUE DISTRIBUTIONS OF THE ADVERTISERS 81

APPENDIX II 90
TESTS FOR AUCTION CONVERGENCE OVER THE ROUNDS 90
APPENDIX III 93
TESTS FOR AUCTION CONVERGENCE OVER THE ROUNDS GIVEN DIFFERENT SIZES OF ADVERTISER POOL 93
APPENDIX IV 95
SENSITIVITY TEST FOR NECESSARY NUMBER OF AUCTIONS REQUIRED TO REACH EQUILIBRIUM IN EACH SCENARIO 95
APPENDIX V 97
COMPARISONS OF DIFFERENT METHODS FOR INITIATING THE AUCTIONS 97
APPENDIX VI 101
PROGRAM TESTING PROCEDURE 101
APPENDIX VII 104
THE PROGRAMS 104
APPENDIX VIII 106
THE REGRESSION MODELS TO ESTIMATE THE EFECTS OF BID INFORMATION DISCLOSURE POLICY AND OTHER AUCTION PARAMETERS 106

BIBILIOGRAPHY 109

LIST OF TABLES

Table 1: PoA Bounds for GSP Auctions. 22

Table 2: Revenue ratio comparison across different information disclosure policies and mix of participating advertisers 52

Table 3: Revenue ratio comparison across different information disclosure policies and value distributions..... 52

Table 4: Revenue ratio comparison across different information disclosure policies and statistical character of relationships among value distributions..... 53

Table 5: Revenue ratio comparison across different information disclosure policies and learning algorithms 54

Table 6: Revenue ratio comparison across different information disclosure policies for common and different learning algorithms..... 54

Table 7: Revenue ratio comparison across different information disclosure policies for different click decay rates 55

Table 8: Yield comparison across different information disclosure policies and mix of participating advertisers 58

Table 9: Yield comparison across different combinations of information disclosure policies and value distributions 59

Table 10: Yield comparison across different information disclosure policies for four different relationships among value distributions..... 60

Table 11: Yield comparison across different information disclosure policies and learning strategies 60

Table 12: Yield comparison across different information disclosure policies and common and different learning algorithms..... 61

Table 13: Yield comparison across different information disclosure policies and different click decay rates..... 62

Table 14: Surplus ratio comparison across different information disclosure policies and mix of participating advertisers 65

Table 15: Surplus ratio comparisons across different information disclosure policies and for the Beta and Gamma value distributions 66

Table 16: Surplus ratio comparisons across different information disclosure policies for four different statistical relationships among advertisers' value distributions 66

Table 17: Surplus ratio comparison across different information disclosure policies for the Cournot learning and Fictitious play as learning strategies 67

Table 18: Surplus ratio comparison across different information disclosure policies when competing bidders use common and different learning algorithms 68

Table 19: Surplus ratio comparison across different information disclosure policies and common and different click decay rates. 69

Table 20: t-test table between adjacent rounds. 91

Table 21: Mann-Whitney test table between adjacent rounds. 92

Table 22: Comparisons of convergence rounds given different sizes of advertiser pools..... 94

Table 23: Sensitivity tests for numbers of auctions in each scenario. 96

Table 24: Changing over round with different initial bids. 98

Table 25: Comparisons with different initial bids and under different information disclosure policy..... 99

Table 26: Regressions to estimate effects of bid information disclosure policies, and the auction parameters 107

LIST OF FIGURES

Figure 1. A search result page for “SUV for sale.” 2

Figure 2: Histogram for revenue ratio under no information disclosure policy 49

Figure 3: Histogram for revenue ratio under partial information disclosure policy 50

Figure 4: Histogram for revenue ratio under perfect information disclosure policy 51

Figure 5: Histogram for yield under no information disclosure policy 56

Figure 6: Histogram for yield under partial information disclosure policy 57

Figure 7: Histogram for yield under perfect information disclosure policy 58

Figure 8: Histogram for surplus ratio under no information disclosure policy 63

Figure 9: Histogram for the surplus ratio under partial information policy 64

Figure 10: Histogram for the surplus ratio under perfect information policy..... 65

Figure 11: The density, mean, and relationships among the 10 identically and independently distributed random variables following Beta distribution. The diagonal graphs demonstrate the density of the 10 variables. The upper right parts are the correlations among the 10 variables. And the lower left parts are the means of the 10 variables. The figure is generated by using 1000 random draws from the distributions. 82

Figure 12: The density, mean, and relationships among the 10 identically and correlated distributed random variables following Beta distribution. The diagonal graphs demonstrate the density of the 10 variables. The upper right parts are the correlations among the 10 variables. And the lower left parts are the means of the 10 variables. The figure is generated by using 1000 random draws from the distributions. 83

Figure 13: The density, mean, and relationships among the 10 independently distributed random variables which are different and follow Beta distribution. The diagonal graphs demonstrate the density of the 10 variables. The upper right parts are the correlations among the 10 variables. And the lower left parts are the means of the 10 variables. The figure is generated by using 1000 random draws from the distributions. 84

Figure 14: The density, mean, and relationships among the 10 different and correlated random variables following Beta distribution. The diagonal graphs demonstrate the density of the 10 variables. The upper right parts are the correlations among the 10 variables. And the lower left parts are the means of the 10 variables. The figure is generated by using 1000 random draws from the distributions. 85

Figure 15: The density, mean, and relationships among the 10 identically and independently distributed random variables following Gamma distribution. The diagonal graphs demonstrate

the density of the 10 variables. The upper right parts are the correlations among the 10 variables. And the lower left parts are the means of the 10 variables. The figure is generated by using 1000 random draws from the distributions. 86

Figure 16: The density, mean, and relationships among the 10 different and correlated distributed random variables following Gamma distribution. The diagonal graphs demonstrate the density of the 10 variables. The upper right parts are the correlations among the 10 variables. And the lower left parts are the means of the 10 variables. The figure is generated by using 1000 random draws from the distributions. 87

Figure 17: The density, mean, and relationships among the 10 independently distributed random variables which are different and follow Gamma distribution. The diagonal graphs demonstrate the density of the 10 variables. The upper right parts are the correlations among the 10 variables. And the lower left parts are the means of the 10 variables. The figure is generated by using 1000 random draws from the distributions. 88

Figure 18: The density, mean, and relationships among the 10 identical and correlated random variables following Gamma distribution. The diagonal graphs demonstrate the density of the 10 variables. The upper right parts are the correlations among the 10 variables. And the lower left parts are the means of the 10 variables. The figure is generated by using 1000 random draws from the distributions. 89

INTRODUCTION

Search engines have revolutionized the business of advertising. Serving as a platform, a search engine facilitates interaction between advertisers and search engine users. In traditional media, such as newspapers or magazines, advertisers and users find each other by inferring from the content of a media vehicle who its audience will be and whose ads it will carry. For example, skin care product advertisers often find customers in fashion magazines, and interested consumers can acquire samples of these products through such magazines. However, even though not all readers of a fashion magazine are interested in skin care products, these advertisements are nevertheless displayed to all readers. In contrast, a search engine is a much more precisely targeted channel. Rather than display an advertisement to all users, a search engine only shows advertisement to users who have searched the relevant keywords. Moreover, advertisers only have to pay when their advertisements are clicked. One of the outcomes of this revolution is increased efficiency because of the improved match between search engine users and advertisers.

Today, search engines comprise a fast-growing advertising channel. In the United States, search advertising revenue was $20.5 billion in 2015, an increase of 8% from the preceding year (Interactive Advertising Bureau, 2016).

To be successful in the business of search advertising, a search engine needs to solve a two-part problem: 1) attract user attention and 2) sell it to advertisers who place different values on the resource. Currently, search engines provide free search services to attract users' attention and use keyword auctions to allocate this resource among advertisers. This dissertation focuses on the second part of the problem: allocating user attention among advertisers.

Figure 1 is an example of a search results page.

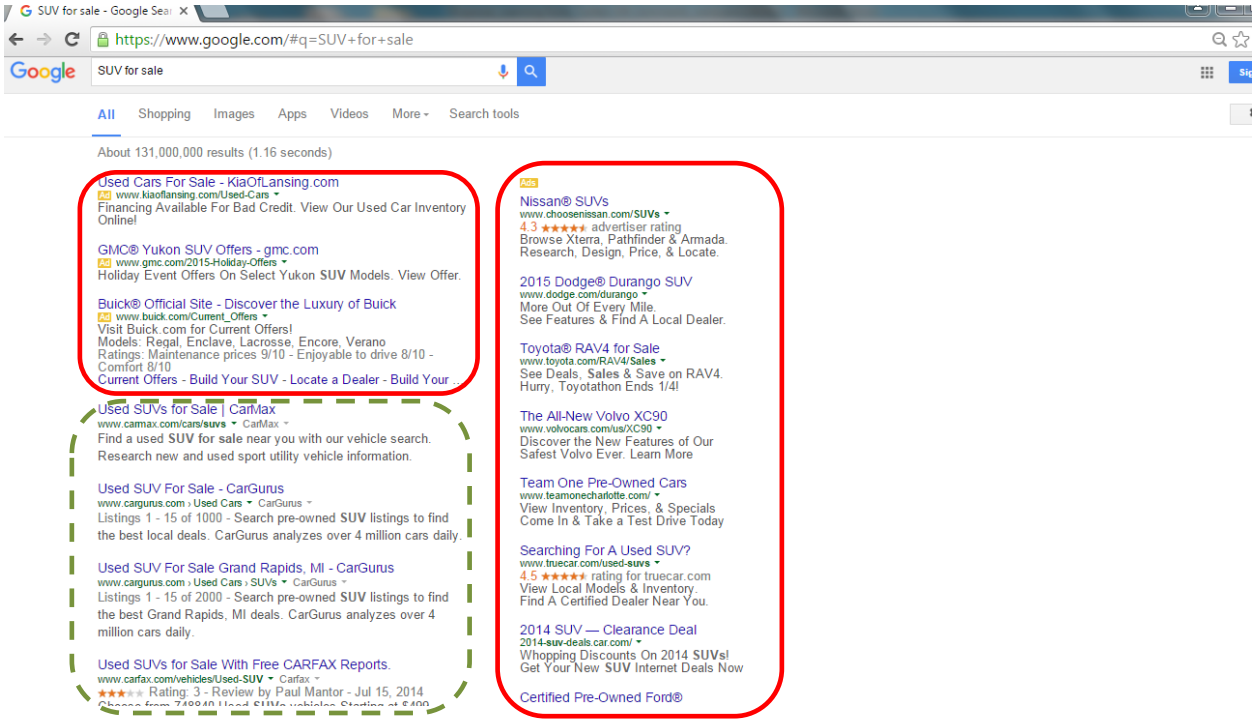

Figure 1. A search result page for "SUV for sale."

The top and right listings inside the red boxes in Figure 1 are advertisements. The slots in these boxes are sold to advertisers through auctions. The listings inside the green dashed box are organic search results. This content is generated through search algorithms and is free of charge. To obtain ad slots on search result pages, advertisers participate in keyword auctions held by the search engine. First, advertisers select the search keywords that their targeted consumers might use. Then, the advertisers specify the maximum amount they are willing to pay for each click generated by their advertisements as bids. The search engine stores these bids as standing bids. Each time a user searches for a keyword selected by any advertiser, an auction begins. The search engine retrieves relevant bids and sorts the bids, possibly weighted, according to the ranking rule in effect. The weights on bids are quality scores. A quality score is the search

engine's estimate of the probability that a user will click on an advertisers' ad. An auction concludes with the determination of how advertising slots on the search results page will be allocated.

The auction format that most search engines currently employ is a generalization of the second-price or Vickrey auction (1961), known as the generalized second-price (GSP) auction. In a GSP auction, when there is only one item for sale, the winner's payment is equal to the second-ranking bid. When there are multiple items for sale, the payment rule applies to every winner. That is, every winner of a slot, j, pays an amount that is equal to the winner's next highest rival's bid, b_{j+1}, or a weighted value of that bid, $\frac{q_{j+1}}{q_j} b_{j+1}$, where q_j is the quality score for the advertiser who occupies slot j and q_{j+1} is the quality score for the next highest rival, who occupies slot $j + 1$.

A commercial search engine's primary goal is to maximize its revenue from the sale of its ad slots. Because the revenue generated by an ad slot is the product of the number of clicks generated by ads in that slot and the price per click advertisers pay, researchers in this field often use the click-through rates (CTRs) CTR_j and CTR_{j+1} as approximations of q_j and q_{j+1}, where an ad's CTR is the percentage of exposures to an ad that generate clicks. However, in practice a search engine takes many factors into consideration when calculating an advertiser's quality score, such as the quality of the advertiser's landing page. Researchers use the CTR approximation because the factors search engines consider in quality score generation are proprietary information and because independent researchers generally believe that quality scores are dominated by CTRs (Maillé, Markakis, Maurizio, Stamoulis, & Tuffin, 2012).

Auctions are one of the most commonly used mechanisms for allocating resources (Chandrashekar, Narahari, Rosa, Kulkarni, Tew, & Dayama, 2007). The goals for a keyword auction are to maximize the auctioneer's revenue and to maximize social welfare. In order to achieve its goals, a search engine needs a mechanism that can assign ad slots in accordance with advertisers' contributions to the objectives of the search engine, which could be revenue maximization, surplus maximization, or a weighted combination of search engine revenue and surplus. An advertiser's contribution is the product of two factors: 1) the click generation potential of their advertisement and 2) the value of a click to the advertiser. Search engines use quality scores as indicators of advertisers' click generation potentials. And bids are the amounts advertisers offer to pay for clicks, which do not have to be their values for clicks. Search engines invest substantially in quality score estimation. There is a line of research focused on how to best respond to the trade-offs between bid levels and CTRs by assigning weights to CTRs (Feng, Bhargava, & Pennock, 2007; Lahaie & Pennock, 2007; Vorobeychik, 2009).

One major finding from previous research (e.g., Aggarwal, Goel, & Motwani, 2006; Edelman, Ostrovsky, & Schwarz, 2007; Edelman & Schwarz, 2010; Varian, 2007) is that bids in GSP auctions do not necessarily equal advertisers' click values, which is called nontruthfulness. Truthfulness in an auction means that the auction mechanism compels bidders to submit their true values for items as their bids. When bidding in a truthful auction, it is in bidders' best interest to submit their true values as their bids. Truthfulness is desirable for an auction mechanism because it guarantees that the auction's surplus is maximized at the equilibrium, which means that the total payoff for the players, including the auctioneer and bidders, reaches the highest possible level. Furthermore, bidders who play against each other in a truthful auction

do not have to engage in costly strategic activities, such as constantly monitoring other bidders and trying to undercut other bidders by frequently changing bids.

The nontruthfulness raises problems for search engines, advertisers and researchers who are trying to better understand the economics of search advertising. Nontruthfulness makes the dynamics of auctions much more complicated. The range of strategically plausible bids is dramatically increased and the number of equilibria become infinite. Thus, nontruthfulness makes it harder to predict the outcome of a GSP auction and to evaluate and improve the performance of the auction mechanism. Researchers have dealt with the nontruthfulness property of GSP auctions in three ways. First, many studies (e.g., Aggarwal, Goel, & Motwani, 2006; Edelman, Ostrovsky, & Schwarz, 2007; Edelman & Schwarz, 2010; Varian, 2007) have concentrated on static complete-information models in which a number of advertisers compete for a smaller number of ad slots in a single-shot auction. Advertisers' values for clicks are common knowledge in these models. These studies have found that although a static complete-information GSP model has infinitely many equilibria, it can nevertheless generate efficient equilibria, one of which is the lowest revenue envy-free Nash equilibrium[1], under certain conditions. Interestingly, this efficient equilibrium coincides with the equilibrium outcome of the Vickrey-Clarke-Groves mechanism (VCG; Vickrey, 1961; Clarke, 1971; Groves, 1973), which is a truthful mechanism. Second, researchers have studied all equilibrium outcomes of static complete-information and incomplete-information GSP models and quantified their maximum efficiency losses relative to the efficiency optimum (Caragiannis, Kaklamanis, Kanellopoulous,

[1] In GSP auctions, when a set of bids reaches a rest point, where no bidder has incentive to swap ad slot with the bidder just above it we call such set of bids envy-free. We call the outcomes produced by the sets of envy-free bids as envy-free Nash equilibria (Edelman, Ostrovsky, & Schwarz, 2007). In Varian's study (2007) these equilibria were called symmetric Nash equilibria. Among all the envy-free Nash equilibria there is one in which search engine receive lowest level of revenue and the payoff of bidders reaches highest level. This equilibrium is called lowest revenue envy-free Nash equilibrium.

& Kyropoulou, 2012; Caragiannis, Kaklamanis, Kanellopoulos, Kyropoulou, Lucier., 2012; Lahaie, 2006; Lucier & Paes Leme, 2011; Paes Leme & Tardos, 2011a). This line of research has found that the surplus and revenue levels of GSP auctions have positive lower bounds, which means that even for a non-efficient equilibrium, there is a limit of how far the surplus and revenue levels can depart from the optimal outcomes. Third, researchers (Varian, 2007; Edelman & Ostrovsky, 2007; Börgers et al., 2013; Athey & Nekipelov, 2011; Duong & Lahaie, 2011) have used bids observed from keyword auctions to estimate advertisers' click values. The estimated click values have been used to test hypotheses (Börgers, Cox, Pesenorfer, & Petricek., 2013) and for counterfactual experiments (Athey & Nekipelov, 2011).

The findings generated from the investigations of the revenue and efficiency implication of nontruthfulness shed light on the stability and desirability of the GSP mechanism. These findings are often used to explain why the GSP mechanism, a nontruthful mechanism, is the de facto choice in the search advertising industry. But the literature on GSP auctions is silent about whether a search engine can generate higher levels of surplus and revenue by using one or more auction instruments that can mitigate the negative effects on surplus and revenue of the nontruthfulness property. While a search engine controls a number of auction instruments (or potential instruments) that might affect performance, only a few have been examined for GSP auctions, such as the reserve price. Prior research on single item auctions has shown that the amount of information an auctioneer reveals about bids from prior rounds can impact performance, but to date no one has asked whether information about prior round bids is an instrument that search engines might employ to increase the amount of revenue and/or surplus their auctions generate. Furthermore, if a search engine could employ bid information from prior rounds as an instrument, then *how* should such instrument be used in the GSP auctions?

The knowledge from previous research, including studies of GSP auctions and single item auctions, is not sufficient to provide immediate answers to the questions raised in the previous paragraph. In studies that model the GSP auctions, researchers often use a static complete-information model to approximate a real world GSP auction. The argument for this modeling method is that keyword auctions are recurring events, in which auctions for the same keywords are held as long as search queries continue to come in. Therefore, advertisers could gain opportunities to experiment with different bids, learn about other advertisers' private information, such as their values for clicks and their bidding strategies, and find out their payoffs associated with obtaining different slots. Possessing this kind of knowledge could help advertisers to improve their payoffs and ultimately the auctions would converge to the efficient Nash equilibrium. However, this argument has not been generally accepted and as a result considerable effort has been devoted to models with less than perfectly informed bidders. Additionally, findings from previous single item auctions studies are inconclusive regarding the effects of disclosing bid information on auctioneer revenue and surplus generation for GSP auctions. Some studies (e.g., Katzman & Rhodes-Krop, 2008) have found that disclosing bid information has a positive effect on auction surplus, but might be detrimental to auctioneer revenue under certain conditions. Other studies (e.g., Elmaghraby & Keskinocak, 2000) have suggested that releasing bid information induces aggressive bidding behaviors, and thereby increases auctioneer's revenue or decreases their costs. Thus, it is theoretically and practically meaningful to study the effects of an information disclosure policy on surplus and search engine revenue of GSP auction.

In my dissertation study, I employed two different learning algorithms. For these learning algorithms I were able to ask what information bidders acquire by participating repeatedly in

auctions. This allowed me to compare the effects of alternative information disclosure policies on auction performance and address the following research question: will disclosing bid information, such as statistics or actual bids from prior rounds, help the search engines to improve the performance of GSP auctions?

Modern information technologies make it possible for an auctioneer to easily alter the information disclosure policy for an auction, which for this study is the extent to which information about earlier rounds' bids is disclosed. A search engine has a range of policy choices. At one end of the spectrum, an auctioneer can choose not to disclose any bid information. On the other end of the spectrum, all bids can be disclosed to the bidders in an auction after an auction round concludes. I studied the following information disclosure policies for GSP auctions: (i) no information disclosure, in which a search engine does not disclose bid information from the prior round; (ii) partial information disclosure, in which a search engine discloses the mean and standard deviation of bids from the prior round; and (iii) perfect information disclosure, in which bids from the prior round are disclosed.

Previous researchers often faced certain difficulties when studying GSP auctions. First, findings from analytical models indicated that a GSP auction can have infinite many equilibria. Therefore, the differences in auction performance could be the result of auctions converging to different equilibria rather than the effects of auctions utilizing different instruments. Second, search engines do not provide bid information to advertisers. Therefore, researchers cannot directly compare the performance of auction that employ different information disclosure policies. Third, GSP auctions operate in a dynamic, complex environment. For example, advertisers enter and exit the auctions stochastically. These difficulties make it challenging to draw conclusions about the effects of providing bid information to advertisers in GSP auctions.

From the perspective of researchers in this field, the predictive power of a Nash equilibrium concept is at best unclear. From the perspective of search engines, it is hard to assess the effects of auction instruments on surplus and search engine revenue and generalize the findings.

In previous research, researchers who built analytic and empirical models for GSP auctions have often put certain assumptions in place and refined the equilibria in order to deal with the difficulties GSP auctions pose for researchers. The most common assumptions are that advertisers' values for clicks are common knowledge (e.g., Edelman et al., 2007; Varian, 2007) and that advertisers' click values are random draws from identical and independent distributions (e.g., Gomes & Sweeney, 2014). The most frequently used refinement concept is the lowest revenue envy-free Nash equilibrium (e.g., Edelman & Schwarz, 2010). Without these assumptions and this equilibrium refinement concept, models can be analytically intractable. But imposing these assumptions and equilibrium refinement can result in findings that are unrealistic and lack practical value, which highlights the need for research method innovation. Agent-based models, through the manipulation of information disclosure policies and use of controlled auction parameters, can overcome the difficulties faced by GSP auction researchers identified in the previous paragraph. Therefore, to address the questions raised above regarding the effects of bid information disclosure policies, I built agent-based models to simulate GSP auctions subject to different information disclosure policies, ran them for different scenarios, and compared the results. These scenarios were constructed by varying and combing the values of several parameters, including whether the mix of advertisers changes over rounds, click value distributions, the statistical character of relationships among bidders' value distributions,

learning algorithms, whether advertisers use a common or different learning algorithms and the click decay rate for ad slots[2].

I obtained the following results:

The information disclosure policy selected can make a difference in search engine revenue, but the effect of the policy on search engine revenue was not monotone with respect to the amount of information revealed. Providing bid statistics, rather than bids, helps search engines generate higher levels of revenue and extract a larger portion of auction surplus than providing no bid information or providing all bids, although the market transparency level is not at its highest level when a search engine provides bid statistics.

All auctions generated at least 90% of maximum attainable surplus, with only very small differences among auctions that differed considerably in auction design features, assumptions about the nature of entry, assumptions about the numbers of clicks generated by ads in different slots, and assumptions about how advertisers learn and formulate bids. This indicates that the GSP auction is robust in terms of surplus generation. Even though these GSP auctions had different characteristics they nevertheless generated similar amounts of surplus.

This dissertation contributes to research on GSP auctions in the following ways:

1. It is the first attempt to explore the effect of the information disclosure policy as an auction instrument in GSP auctions. The effects of other auction instruments, such as the reserve price and ranking rules, have already been studied.

[2] This parameter controls how many clicks the ad slots generate. I followed the conventional wisdom for keyword auctions and assumed that the numbers of clicks will not increase as the rankings of ad slots go down. This means that the number of clicks from $slot_{j+1}$ can only be smaller than or equal to the number of clicks from $slot_j$.

2. It provides new insights into the properties of GSP auctions. I found that the GSP mechanism was robust in terms of surplus generation and that it can potentially generate higher level of search engine revenue than the VCG mechanism. This finding helps to explain why the GSP mechanism, rather than the VCG mechanism, is widely adopted.

3. It provides a way to answer "what-if" questions that otherwise cannot be effectively examined by researchers. There are many empirical studies of GSP auctions that have used real bidding data. Although these studies have described what was happening, they cannot answer what-if questions. Providing bid information has the potential to improve keyword auction outcomes for search engines, but empirical analysis of real bidding data cannot reveal this because keyword auctions are generated under a regime in which search engines do not provide bid information.

4. By running the auctions under different scenarios, I found that relaxing the assumption that click value distributions of advertisers are identical and independent did not qualitatively change the outcomes of auctions.

5. From the standpoint of research methodology for GSP auctions, this dissertation provides an alternative; namely, agent-based modeling. This method has two important advantages. First, it can handle more complexity than static complete information models. Second, it can be used to answer what-if questions that cannot be addressed with other analytical approaches. These advantages make agent-based modeling a valuable complement to other research methods in this field.

The remainder of this dissertation is organized as follows: in chapter 2, I provide an overview of the relevant literature. I discuss the literature on equilibria, efficiency, and search

engine revenue in GSP auctions, as well as the literature on auction instruments. In chapter 3, I specify the agent-based models and the scenarios for which the simulated GSP auctions were run. In chapter 4, I present the results from the simulations. Finally, in chapter 5, I conclude by discussing the meaning and uses of the findings from this study and discuss the limitations of this study, along with future research directions.

LITERATURE REVIEW

This chapter examines the literature related to keyword auctions, summarizes the findings, and identifies gaps in the literature. Although a relatively new research area, search advertising has attracted substantial attention among researchers (e.g., Aggarwal et al., 2006; Edelman et al., 2007; Edelman & Schwarz, 2010; Varian, 2007). The keyword auction is attractive to researchers for two reasons. First, the commercial success of keyword auctions justifies research interest. Second, distinctive features of search advertising auctions set them apart from traditional auctions, such as the English auction and Dutch auction. These new features present intriguing problems to solve. For example, will a search advertising auction produce an efficient outcome? And, are the auction instruments used in single-item auctions, such as a reserve price, useful for generating higher levels of auctioneer revenue and/or surplus in search ad auctions? One feature in particular sets search ad auctions apart from traditional auctions. That is, there are multiple vertically differentiated items to be auctioned off simultaneously; namely, the ad slots on a search results page.

The seminal works in the field (Edelman et al., 2007; Edelman & Schwarz, 2010; Varian, 2007) have shown that GSP auctions are not truthful and advertisers often engage in strategic bidding. The fact that GSP auctions are not truthful and that advertisers engage in strategic bidding (Edelman & Ostrovsky, 2007) means that advertisers may have an incentive to misreport their true values for clicks and that the surplus generated by auctions and search engines' revenue may not be maximized. These findings have invited more research attention because it is counter-intuitive that a seemingly inferior auction design is commonly used by search engines. Researchers (e.g., Lahaie, 2006; Lucier & Paes Leme, 2011; Paes Leme & Tardos, 2011a) who have investigated this mystery have found that although GSP auctions are not truthful, they work

reasonably well in terms of surplus and search engine revenue generation. Furthermore, the payment calculation method for GSP auctions is simpler than that of VCG auctions. To further advance knowledge regarding auctions for multiple vertically differentiated items, researchers have also investigated several auction instruments. The most frequently studied instruments include ranking rules (Lahaie & Pennock, 2007) and the reserve price (Edelman & Schwarz, 2006; Ostrovsky & Schwarz, 2011). Bid information disclosure, a potential instrument that has yet to be explored in the context of search engines, is examined in this dissertation.

2.1 General Settings and Assumptions in Keyword Auction Research

The general setting in keyword auction research is that multiple vertically differentiated items are auctioned off. In keyword auctions, advertising slots are the items to be auctioned off. Slots at higher positions are deemed higher quality because they tend to attract more search user attention than those at lower positions. Certain assumptions are often employed to simplify the problem of modeling and studying the revenue and surplus performance of search engines. The most frequently employed assumptions are the following:

- Advertisers are fully rational. This is a strong assumption, which implies that advertisers have well-defined utility functions and access to the information needed to maximize their utility.
- Advertisers do not have budget constraints (e.g., Gummadi, Key, & Proutiere, 2013).
- The CTRs for different advertisements are independent of each other and are common knowledge. The majority of the research in this area (e.g., Varian, 2007, 2009) also assumes that the CTR for an advertisement is the product of two independent components, an advertiser-specific factor and a bid ranking-specific factor.

- The same group of advertisers or advertisers drawn from the same population of advertisers bid against each other repeatedly over multiple rounds of bidding (e.g., Gomes & Sweeney, 2014).
- Advertisers' values for clicks are independent of the ranks of the ad slots their advertisements are placed in. In other words, for advertiser *i*, a click is a click regardless of the ad slot from which it is generated, which means that individual users who click on ads in different slots have the same value to advertisers (e.g., Edelman et al., 2007; Gomes & Sweeney, 2014; Varian, 2007, 2009).

In the real world, keyword auctions are much messier than most studies assume and the aforementioned assumptions are often violated in nontrivial ways. Advertisers participating in keyword auctions are heterogeneous in terms of their resources and capabilities for maximizing their payoffs. Additionally, because the system is complex and dynamic, full rationality is hardly a reasonable assumption. Furthermore, advertisers often bid under some kind of budget constraint. Moreover, it is impossible to precisely predict the CTRs of the advertisements on a search result page because the rank and quality of the advertisements and the identities of the advertisers on the same search results page influence one another's CTRs (Kempe & Mahdian, 2008). To capture the dynamic nature of search advertising, theoretical researchers sometimes model keyword auctions as repeated games in which the same groups of advertisers play against each other over and over. Therefore, advertisers' values for clicks will not remain private information for long. Unfortunately, this setting hardly describes real keyword auctions. The identities of competitors for the same keyword can vary drastically across rounds. Incumbent advertisers might drop out of a keyword auction at some time point due to budget depletion or for other reasons, and new advertisers can enter at any time. Moreover, empirical findings

suggest that all clicks are not created equal. Advertisers value clicks originating from higher ad slots more than those originating from lower ad slots (Börgers, et al, 2013). Although search advertising research began with these simplifying assumptions, researchers have invested substantial effort in moving this line of research forward by relaxing some of these assumptions.

2.2 The VCG Mechanism

A GSP mechanism is an auction mechanism for allocating vertically differentiated items, such as the paid ad slots on a search results page. A VCG mechanism is an alternative mechanism for selling such items. The equilibrium outcome of a VCG auction is often used as the reference point when evaluating the performance of the corresponding GSP auction because the lowest revenue envy-free Nash equilibrium of the GSP auction coincides with the equilibrium for the VCG auction and because the outcome of a VCG auction is surplus maximizing.

Prior research has investigated the VCG mechanism at length. Under the VCG mechanism, bidders are ranked based on their bids or weighted bids. The weights are the estimates of the click generating potentials of advertisers. Each winner's payment equals the externality that the winner imposes on other bidders, which is the difference in surplus levels between the scenario where the bidder participates in the auction and that where the bidder does not participate in the auction. For instance, supposing that the surplus of a VCG auction when advertiser *i* is in it is $surplus_i$ and the surplus of the VCG auction when advertiser *i* is not in the it is $surplus_{-i}$. Then the payment of advertiser *i* is $surplus_i - surplus_{-i}$. In the context of a search advertising auction, advertiser *i*'s payment is calculated as follows:

$$p_i = \sum_{i=i}^{k} \frac{q_{i+1}}{q_i} b_{i+1}$$

where , q_{i+1} and q_i are the estimates of the click generating potentials for advertiser i + 1 and advertiser i, and b_{i+1} is the bid of the advertiser i + 1 whose bid ranks just below advertiser i's.

The VCG mechanism is truthful. Therefore, bidding one's true value is a weakly dominant strategy, which means that advertiser i cannot do better by submitting something other than its true value as a bid. At equilibrium, where every advertiser submits a bid equal to its true value, the surplus of the auction is maximized (Maillé et al., 2012).

2.3 Equilibria Identification and Refinement

Although a GSP auction is a generalization of a second-price or Vickery auction, it has very different properties. Even modeled as a static, complete information game, a GSP auction can have infinitely many equilibria. These equilibria are different in terms of their slot allocations, search engine revenue, and efficiency levels. Out of all the equilibria, the lowest revenue envy-free Nash equilibrium (Edelman et al., 2007) has received the most attention because the surplus of the auction is maximized at this equilibrium and it coincides with the equilibrium of the corresponding VCG auction. Thus, researchers have often focused on the lowest revenue envy-free Nash equilibrium. However, researchers (Gomes & Sweeney, 2014; Hashimoto; 2013; Giotis & Karlin, 2008) have also found that this equilibrium can cease to exist unless certain assumptions are satisfied. The studies of Gomes and Sweeney (2014) and Giotis and Karlin (2008) found that the difference among the CTRs of ad slots have to be large enough to ensure that bidders don't see neighboring slots as very close substitutes. If neighboring ad slots are sufficiently close substitutes for each other, the advertiser who obtains the higher slot

will find it profitable to lower its bid and settle for the lower slot instead. And, the advertiser who obtains the lower slot will also lower its bid and try to retain the lower slot. The outcome of this dynamic is that the bids of all advertisers are the same and ad slots are assigned randomly. In addition, Hashimoto (2013) found that if there is an advertiser who bids irrationally the lowest revenue envy-free Nash equilibrium can also be eliminated.

Edelman and colleagues (2007), Varian (2007, 2009), Edelman and Schwarz (2010), and Li, Zeng, and Zhao (2012) identified the properties that characterize all pure strategy Nash equilibria. If the following inequalities hold for an allocation, then the bid profile supports a pure strategy Nash equilibrium:

$$CTR_s(v_s - p_s) \geq CTR_j\left(v_s - p_j\right) \ \forall j > s,$$

$$CTR_s(v_s - p_s) \geq CTR_j\left(v_s - p_{j-1}\right) \ \forall j < s,$$

where p_s and p_j are the payments of the advertisers occupying slots *s* and *j*, CTR_s and CTR_j are the click-through rates for slots *s* and *j*, and v_s is the value of a click for advertiser *s*. At a pure strategy equilibrium, each advertiser finds it is unprofitable to unilaterally change its bid to win another ad slot.

Edelman et al. (2007) and Varian (2007) have shown that among all the Nash equilibria, there exists a unique Nash equilibrium that leads to the lowest search engine revenue and highest advertisers' payoffs and this equilibrium coincides with the equilibrium outcome of the VCG mechanism. This equilibrium satisfies the following inequality:

$$CTR_s v_s \geq CTR_{s+1} v_{s+1}.$$

Gomes and Sweeney (2014) extended the study of GSP auctions to an incomplete information or Bayesian setting. In their model, an advertiser's value for a click is private information. But the click value is drawn from an independent, identical distribution with the support of [0, v], which is common knowledge. A bidding strategy, $b(v_i)$, produces a Bayesian Nash equilibrium if the following inequality is satisfied, where $b_i'(v_i)$ is a bidding strategy that is different from $b(v_i)$:

$$E_{v_{-i},b}[u_i(b_i(v_i), b_{-i}(v_{-i}))|v_i] \geq E_{v_{-i},b}[u_i(b_i'(v_i), b_{-i}(v_{-i}))|v_i].$$

where E is the expected value of a variable. This inequality says that the Bayesian Nash equilibrium strategy maximizes advertiser *i*'s expected utility given *i*'s click value and the opponents' bids according to strategies $b_{-i}(i)$.

On a search results page, search advertisements are displayed next to each other. When studied purely as auctions, the possibility that one ad's CTR might be influenced by other ads on the same search results page was assumed away in many models of the GSP auctions (e.g., Edelman, Ostrovsky, & Schwarz, 2007; Varian, 2007; Thompson & Leyton-Brown, 2009; 2013; Bu, Deng, & Qi, 2008). Giotis and Karlin (2008) incorporated such influence in their study of GSP auctions. The authors found that a GSP auction always has a pure Nash equilibrium. However, the Nash equilibrium is not necessarily an efficient one. The authors also found that even for the worst case scenario, the surplus level of a GSP auction is $\frac{1}{k}$ of that for the corresponding VCG auction, where *k* is the number of slots the search engine decides to sell, provided advertisers do not bid above their values for clicks.

Researchers have proposed several equilibrium refinement concepts. For example, Edelman and colleagues (2007), Varian (2007) and Li, Zeng, and Zhao (2012) refined Nash

equilibria by eliminating the equilibria with lower efficiency levels. Edelman and colleagues (2007) and Varian (2007) identified the subset of Nash equilibria called *symmetric Nash equilibria* or *locally envy-free Nash equilibria*. Li, Zeng, and Zhao (2012) restricted the set of equilibria considered to *stable Nash equilibria* (STNE). A STNE is either the same as or a proper subset of the set of symmetric Nash equilibria. In these equilibria, an advertiser will not have an incentive to exchange its position and payment for those of the neighbor immediately below or above it. Hashimoto (2013) proposed a different strategy for equilibrium refinement. Specifically, Hashimoto introduced a noisy bidder who submits a bid that is a random draw from a uniform distribution and participates in the auction stochastically. Hashimoto found that the VCG-equivalent equilibrium is eliminated when a noisy bidder participates in an auction.

2.4 Efficiency Loss of GSP auctions

GSP auctions do not guarantee that all outcomes will constitute efficient equilibria (Edelman et al., 2007; Gomes & Sweeney, 2014; Varian, 2007, 2009). Yet, GSP auctions are still widely used in the search ad industry. A number of studies have focused on quantifying the extent to which the surplus of a GSP auction outcome departs from the efficient outcome (e.g., Caragiannis, Kaklamanis, Kanellopoulous, & Kyropoulou, 2012; Caragiannis, Kaklamanis, Kanellopoulos, Kyropoulou, Lucier et al., 2012; Lahaie, 2006; Lucier & Paes Leme, 2011; Paes Leme & Tardos, 2011). The main goal of this line of research is to calculate the minimum efficiency level achieved with a GSP auction.

To measure the efficiency loss of a nontruthful auction mechanism, such as a GSP auction, the price of anarchy (PoA) is often used, which is defined as follows:

$$PoA = \sup_{v,b} \frac{OPT(v)}{SW(b,v)}.$$

where, OPT(v) is the maximum attainable surplus given the click values, (v) of the advertisers and SW (b, v) is the realized surplus given the bids, b, and the click values, v, of the advertisers. An auction's PoA is the lowest upper bound of the ratio between its maximum attainable surplus level and its realized surplus level of a GSP auctions and can range from 1 to positive infinity. When the PoA equals 1, the surplus of the GSP auction is at its optimum level. When the PoA is greater than 1, the surplus of the GSP auction is less than the maximum attainable level.

This line of research started with the analysis of a special case where the CTRs of the slots decay exponentially. The exponent is $\frac{1}{\delta}$, which is greater than zero and smaller than 1. Lahaie (2006) proved that the PoA for GSP auctions is the larger of δ and $\frac{\delta}{\delta-1}$. Paes, Leme and Tardos (2010a) extended this analysis to include situations where the period-to-period decay of slots' CTRs can follow any monotonically declining pattern and calculated the PoAs for pure, mixed, and Bayesian-Nash equilibria as 1.618, 4, and 8, respectively. According to the calculations of Caragiannis and colleagues (2011) the PoA is only 1.282 at the pure strategy Nash equilibrium. Lucier and Paes Leme (2011) and Caragiannis, Kaklamanis, Kanellopoulous, and Kyropoulou (2012) improved the calculation of the mixed strategy PoA in Paes Leme and Tardos's (2010) study and lowered the PoA from 4 to 2.310. Furthermore, Lucier and Paes Leme (2011) and Paes, Leme and Tardos (2010b) carried out the PoA calculation in a Bayesian setting. and found that when the bidders' value distributions are independent the PoA is $2(1-\frac{1}{e})^{-1} \approx 3.16$, but when value distributions are identical and correlated the PoA is 4.

Caragiannis and colleagues (2011) found that the PoA is 2.927 for Bayes-Nash equilibria when advertisers face uncertainties about quality scores and their opponents' values for clicks.

This result is robust with respect to variation in the number of advertisers, the number of ad slots, and the distributions of advertisers' values for clicks.

The following table summarizes the findings of this line of research:

Table 1: PoA Bounds for GSP Auctions.

	Pure Nash equilibrium	Mixed Nash equilibrium	Bayesian Nash equilibrium	Nonequilibrium	
				Independent values for advertisers	Correlated values for advertisers
PoA	1.282	2.310	2.927	3.16	4

2.5 Revenue Levels of GSP Auctions

Similar to the surplus levels, the revenue levels of GSP auctions vary among equilibria. In order to make comparisons, researchers have often chosen the revenue level of the corresponding VCG auction as the reference point. Edelman and colleagues (2007) and Varian (2007) found that for arbitrarily chosen locally envy-free Nash equilibrium or symmetric Nash equilibria in a complete-information setting, the revenue levels of GSP auctions are at least equal to those for the corresponding VCG auctions. Lucier and colleagues (2012) extended this work to incomplete-information settings and proved that, combined with a Myerson's reserve price, a GSP auction will generate revenue at least equal to one sixth of the VCG auction revenue, assuming that advertisers' click values are drawn from identical and independent regular distributions. When advertisers' click values are drawn from identical and independent monotone hazard rate (MHR) distributions, then the revenue is at least $\frac{1}{3.46}$ of the maximum achievable revenue (Caragiannis, Kaklamanis, Kanellopoulous & Kyropoulou, 2012).

2.6 Auction Instruments

The auctioneer can impose create rules for an auction to increase the surplus it generates or auctioneer's revenue. These rules are called auction instruments. There is a long tradition in auctions of using an auction instrument to improve auction performance. For example, a reserve price is often used to increase auctioneer revenue. For GSP auctions, reserve prices and ranking rules are the most frequently studied instruments.

2.6.1 Reserve Price

Early work (e.g., Myerson, 1981; Riley & Samuelson, 1981) regarding reserve prices provided insights into calculating reserve prices for single-item auctions. Edelman and Schwarz (2006) conducted the first study regarding reserve prices in GSP auctions. One goal for this study was to investigate the effects of reserve prices in auctions that sold simultaneously multiple vertically differentiated items. Edelman and Schwarz found that, in general, a reserve price increases search engine revenue. This finding was confirmed by a subsequent empirical study (Ostrovsky & Schwarz, 2011). Edelman and Schwarz (2010) further argued that an optimal reserve price can increase revenue significantly with relatively little loss of efficiency.

In a GSP auction, the reserve price can be either universal or customized for each advertiser. When using a universal reserve price, all advertisers are subject to the same reserve price. When using a customized reserve price, reserve prices are calculated for individual advertisers based on their quality scores, which are the search engine's estimates of advertisers' ad's click generating potential. Studies have found that (Thompson & Leyton-Brown, 2013; Liu, Chen, & Whinston, 2010) a universal reserve price works better for improving search engine revenue than customized reserve prices.

2.6.2 Ranking Rules

A ranking rule is an instrument that search engines can use to increase the revenue and surplus generated by auctions. Ranking rules examined in the literature typically involve a weighted combination of advertisers' per click bids and their ads' click generating potential. At two extremes, a search engine can either rank advertisements based on bid size alone or click generating potential alone. A search engine can also choose a ranking rule that strikes a balance between bid size and click generating potential. Lahaie and Pennock (2007) proposed a class of ranking rules generated by squashing, which is a technique that varies the weight of the quality score in the bid ranking. For this class of ranking rules, advertisers are ranked based on the score $q_i^s b_i$, where q_i, is the quality score of advertiser *i*. and b_i is advertiser *i*'s bid. The power term, *s*, ranges from 0 to 1[3]. When *s* = 0, advertisements are ranked by bids only. When s = 1, advertisements are ranked by expected revenues alone. Researchers (Feng, Bhargava, & Pennock, 2007; Lahaie & Pennock, 2007; Vorobeychik, 2009) have often used computational experiments to find revenue-optimal values for *s*. The general finding is that a search engine must fine tune the value of *s* based on the relationship between advertisers' bids and click generating potential. When there is a positive correlation between *s* and *b*, a smaller *s* is preferable. When the correlation is negative, the search engine should set *s* at a value close to 1.

2.7 Value Estimation

The GSP mechanism does not elicit advertisers' true values for clicks. But click value information is important because it can be used for evaluating auction performance (e.g.,

[3] While there is no necessary upper bound for *s*, quality scores are generally assumed to be no larger than 1.

Edelman & Ostrovsky, 2007) and testing auction instruments (e.g., Ostrovsky & Schwarz, 2011). One way to deal with this problem is to estimate click values from available bid data.

In complete-information settings, researchers (Varian, 2007; Edelman & Ostrovsky, 2007; Börgers et al., 2013) have often estimated an advertiser's click value by looking into all bids over a fixed period of time. The assumption is that advertisers' bids are no higher than their true click values. Therefore, the highest observed bid is sometimes used as an estimate for an advertiser's click value.

In a Bayesian setting, Athey and Nekipelov (2011) developed a method for estimating click values that works for advertisers who rarely change their bids. The authors used all bids by advertisers competing for a search keyword to model the distribution of click values. Based on the location of an advertiser's bid in the distribution, the authors derived an incremental cost-per-click curve. Assuming that an advertiser is rational, an advertiser will submit a bid that equates the marginal click value and marginal cost of a click. Assuming this is in fact the case, the authors were able to estimate an advertiser's click value. Pin and Key (2011) adopted and simplified Athey and Nekipelov's (2011) method and obtained similar predictions. The best response of advertisers in Pin and Key's (2011) study was derived by assuming the existence of a Bayes-Nash equilibrium. When applied to advertisers who change their bids frequently, the method produces multiple click value estimates for each advertiser. Therefore, an advertiser might have different values for clicks in different rounds of a keyword auction.

Duong and Lahaie (2011) bypassed bids and instead used the observed ranks of ads to estimate advertisers' values for clicks. The authors argued that from the perspective of advertisers, a keyword auction can be viewed as a problem of choosing the best slot to maximize

advertiser utility. Therefore, the authors proposed a discrete-choice model to estimate the click values that the advertisers assign to slots and modeled the bidding behavior. By using data from Yahoo!, the authors found that utility increases with the CTR and decreases with price. On average, advertisers shade their bids relative to their click values by 20%.

2.8 Bidding Strategies

These studies of equilibrium and equilibrium refinement provide important insights into the stability of the GSP mechanism, but they leave open the question as to which equilibrium a set of GSP auctions converges. Studies of bidding strategies (e.g., Bu, Deng, & Qi, 2008; Cary et al., 2014) have shed light on the nature of this convergence.

Best response strategies are the most often studied bidding strategies (e.g., Bu et al., 2008; Cary et al., 2014). The idea behind the best response strategy for advertiser *i* is to find the bids that make the advertiser above advertiser *i* pay more while advertiser *i* stays at slot *i*. All bids that are lower than the bid of advertiser *i* -1 and larger than the bid of advertiser *i* +1 result in the same slot allocation and form the best-response range. Some researchers have analyzed equilibrium when the bids are in the middle of the best response range (Bu et al., 2008). Others (Cary et al., 2007; Zhou & Lukose, 2007; Liang & Qi, 2007) have studied the equilibria that emerge when advertisers submit bids that are just below the upper bound or just above the lower bound. These researchers found that if advertisers submit bids in the middle of their best-response ranges, auctions will converge to the lowest revenue envy-free Nash equilibrium. When advertisers submit bids that are close to the upper or lower bounds of their best-response ranges, the auctions do not always converge to an equilibrium. Yao, Chen, and Liu (2012) investigated a weighted-joint fictitious-play bidding strategy. Under this strategy, advertisers form their bids by calculating their best responses given their beliefs about their rivals' bid distributions.

Advertisers update their beliefs by summarizing the bidding histories of their rivals. The authors proved that when there are only two slots, this strategy produces a Nash equilibrium. However, they did not investigate more complicated situations in which more slots are available to auction off and left this for future research.

2.9 The Effects of Bid Information Disclosure Policy in Single-Item Auctions

The effects of bid information disclosure policy have often been studied in single-item auctions, in which there is only one item to auction off in one round of an auction. The choice of the information disclosure policy is one of the important problems in auction design (Elmaghraby & Keskinocak, 2000). There is a literature that has analyzed information disclosure policy in various single-item auctions.

Providing bid information can produce different outcomes in revenue and surplus. The difference in outcomes could be caused by auction format, how much information is disclosed and other factors. Researchers studied several of these factors, including an auction's pricing scheme (such as first price or second price), and whether an auction is a common value auction or a private value auction. For example, in a first price auction, if all the bids are disclosed after a round, the winner of the item could reduce its bid to the level that is just a little bit higher than the next highest rival. In the next round the winner can win the item and pay less for it. Therefore, the auctioneer's revenue will be lower. On the other hand, in a second price auction, if the bids are disclosed the second ranking bidder could raise her bid as high as possible so it can raise the payment of the winner. Thus, the auctioneer's revenue is increased (Bergemann, & Horner, 2010).

The effects of information disclosure are also influenced by how the items are valued by bidders. In common value auctions, the value of an auction item is the same for every bidder. In a private value auction, each bidder has her own value for an auction item and only the bidder knows how much she values it. Gneezy (2002) looked into three types of information disclosure policies for common value auctions. These information disclosure policies include no bid disclosed, winning bid disclosed, and all bids disclosed. They found that when all bids are disclosed the competition among bidders are fiercer and the auctioneer ’a revenue is higher than for the other information disclosure policies.

For private value auctions, Andreoni, Che, and Kim, (2007) used an experiment to test the effect of information disclosure policy on auctioneer revenue in first and second price auctions. They found that disclosing bid information induced bidders to submit bids that were higher than their values for items more often in first price auctions than in second price auctions. Katzman and Rhodes-Krop (2008) found that if the winner’s bid is not disclosed then other bidders might believe that the winner has a higher willingness to pay than it really is. Thus the bidding competition will be fiercer and the auctioneer’s revenue will be higher. Elmaghraby and Keskinocak (2000) suggested that bidders will bid more aggressively if the bids are disclosed and the auctioneer will receive more revenue.

2.10 Summary

The GSP mechanism is used in keyword auctions to allocate the ad slots on a search results page. Generally, ad slots that are higher on the search results page attract more attention from users than lower ones. One finding from static complete-information GSP auction models (e.g., Edelman et al., 2007; Varian, 2007) is that GPS auctions are not truthful. This property often results in suboptimal auction surplus and search engine revenue because the auctions often

do not converge or converge to one of the inefficient equilibria. Thus, a line of research exists that is aimed at quantifying the efficiency and revenue loss for GSP auctions. Researchers (e.g., Caragiannis, Kaklamanis, Kanellopoulous, & Kyropoulou, 2012; Caragiannis, Kaklamanis, Kanellopoulos, Kyropoulou, Lucier, 2012) have found that for inefficient equilibria there are limits of how far the surplus and revenue levels of GSP auctions differ from their optimal levels When assessing the outcomes of GSP auctions, or the effects of auction instruments, researchers have often assumed that the auctions converge to efficient Nash equilibria, one of which is the lowest revenue envy-free Nash equilibrium (Edelman et al., 2007). Alternatively, researchers (e.g. Athey & Nekipelov, 2011; Pin & Key, 2011) have also used observed bids to estimate advertisers' click values. Subsequently, researchers have used these estimated click values to evaluate the effects on search engine revenue and surplus generation for various auction instruments in counterfactual experiments (Athey & Nekipelov, 2011) or simulations (Thompson & Leyton-Brown, 2013). Reserve price (Edelman & Schwarz, 2006; Ostrovsky & Schwarz, 2011) and ranking rules (Lahaie & Pennock, 2007) are the most frequently studied instruments in GSP auctions. The general finding from prior research is that these instruments can be effective in terms of improving search engine revenue.

One gap in the literature on GSP auctions is The implications of research findings that suggest that bid information disclosure could be an important auction instrument for increasing search engine revenue and auction surplus has been largely overlooked in the literature on GSP auctions. As we can see from the findings of previous research, GSP auctions perform better in terms of search engine revenue and surplus when the bidders have more bid information. Caragiannis and colleagues (2011) found that the PoA can be as high as 1.282 when bidders have complete bid information. But when the bidders only have incomplete information the PoA can

be as high as 2.927. Similarly, search engine revenue at least equals that for the corresponding VCG auction when the bidders have complete bid information (Edelman et al, 2007 & Varian, 2007). However, when there is incomplete bid information, search engine revenue is one sixth its value for the corresponding VCG auction (Lucier, et al, 2012).

One possible reason for this gap in the literature on GSP auction is that some researchers (e.g., Edelman et al, 2007; Varian, 2007; Edelman & Schwarz, 2010) assume that advertisers can learn opponents' private click values through the bids revealed to them through the prices they pay for their own slots as they compete against each other repeatedly. However, this assumption is often criticized by researchers (e.g., Edelman & Schwarz, 2010 & Yao, Chen, & Liu, 2012; Athey & Nekipelov, 2011; Ostrovsky & Schwarz, 2011; Pin & Key, 2011). These researchers developed models for situations where advertisers start out less than complete bid information. This indicates that these researchers did not find the learning through experience argument for advertisers having all bid information persuasive. Nevertheless, these models can only take us so far due to the restrictive assumptions imposed to make the models analytically tractable. For instance, researchers (e.g., Gomes & Sweeney, 2014) have often assumed that the click values of the advertisers are random realizations of identical and independent distributions. My dissertation intends to fill this gap in the research by investigating the effects of information disclosure policies in GSP auctions. I used agent-based models to simulate GSP auctions under different information disclosure policies to explore the extent to which bidders learning from experience can produce outcomes close to the complete information outcomes and to ask whether search engines can benefit from disclosing bid information in varying amounts to advertisers.

RESEARCH DESIGN

3.1 The Goals for this Study and Research Questions

As identified in the literature review, no studies exist that look into the effects of bid information disclosure policies in GSP auctions, although such policies might be beneficial in terms of search engine revenue and efficiency. The goal for this study is to bridge this gap by addressing the following research questions:

Research Question 1. Will providing advertisers with bids or bid statistics from prior rounds produce higher levels of search engine revenue in GSP auctions than providing no bid information?

Research Question 2. Will providing advertisers with bids or bid statistics from prior rounds facilitate more efficient allocations of ad slots in GSP auctions than providing no bid information?

3.2 Research Method Comparison

Researchers (e.g., Gomes & Sweeney, 2014; Lucier & Paes Leme, 2011) focused on GSP auctions often use mathematical models. Four assumptions have frequently been imposed upon these highly stylized and simplified (relative to real world auctions) models in previous research. The first two assumptions are about advertisers' values. The first two assumptions are that advertisers' values for clicks are either common knowledge or are drawn from commonly known identical and independent distributions. The third and fourth assumptions are that the competitors in different rounds are the same or are generated through random draws from a single population. These assumptions are employed mainly for the sake of analytical tractability, so the practical and descriptive value of the findings may be limited (Aschenfelter, 1989; Rothkopf & Harstad,

1994). For example, findings under these assumptions can only be applied in scenarios where advertisers for the same product or service market compete against each other and the values they assign to clicks do not influence other advertisers' values. These scenarios obviously only encompass a small portion of reality. For instance, when advertisers come from the same product market, their values for clicks could be correlated with each other because they all employ a common set of inputs in producing their products. Thus, there is a need for a research method that can complement the mathematical approach and fill the void where previous research approaches and results have provided insufficient guidance. Three options are available: the first option is empirical analysis of field data (e.g., Yuan, 2011, 2012). The second option is to use a laboratory experiment (e.g., Goldman & Rao, 2014; Noti, Nisan, & Yaniv, 2014). The third option is to use agent-based modeling (e.g., Hailu, Rolfe, Windle, & Greiner, 2011; Hailu & Schilizzi, 2005). For the purposes of this research, I chose to use agent-based modeling.

Both empirical analysis and laboratory experiments have advantages and disadvantages. Empirical data reflect the behaviors of advertisers under all the vagaries of real keyword auctions, including budget constraints, quality scores, and so on. However, empirical data have three disadvantages. First, these data are often proprietary. Second, important information is often not available in such data sets. For example, quality scores, impressions, and click counts are often not present in such data sets. Third, because GSP auctions are nontruthful, advertisers' bids are not necessarily the same as their values for clicks, thereby forcing researchers to estimate these values. In order to estimate the values, researchers often must assume that the bidding strategies that advertisers use will produce the lowest revenue envy-free Nash equilibrium and this is likely an unrealistic assumption (Börgers et al., 2013).

As an alternative to analyzing real bidding data, researchers have employed laboratory experiments with human subjects (e.g., Noti, et al., 2014) to test theoretical predictions in keyword auction studies. In a sense, laboratory experiments are realistic because real people and real money are involved. Another advantage of laboratory experiments is that it is possible to control for confounding factors, so that causal inference is often plausible and valid. Furthermore, in the context of keyword auctions, a researcher can artificially assign values to bidders. In particular, a researcher can randomly draw a few numbers from several distributions and assign the numbers as their values for clicks to subjects who play advertisers in an experiment. Therefore, the researcher can map subjects' bidding strategies to their values for clicks.

However, a laboratory experiment is not necessarily a suitable choice for my research. My goal is to assess the effects of information disclosure policies on search engine revenue and auction efficiency. There are two reasons that laboratory experiments are unsuitable for this research. First, how advertisers come up with bids and how they learn opponents' bidding strategies are unobservable, and therefore, confounded with information disclosure policies in the experiments. Human subjects tend to use heuristics to come up with bids rather than follow researchers' instructions. For example, Noti and colleagues (2014) found that even after they explicitly told subjects that submitting the true value was the dominant strategy in VCG auctions, subjects' bids still frequently deviated from their true values, even after they had already played a few rounds of their auctions. Therefore, performance differences within auctions could reflect differences in the strategies that different advertisers use, rather than to differences in information disclosure policies. This creates a difficulty similar to that encountered in empirical analysis of bidding data, that is, critical information, such as bidding strategies, remains

inaccessible to the researcher. The second reason that a laboratory experiment is inappropriate for this research is that I am interested in the outcomes of the auction after advertisers have learned each other's bidding strategies. However, learning is a long-term process and human subjects learn at different rates. Some human subjects could be distracted or fatigued before learning can occur or solidify in an experiment.

In contrast, an agent-based model remedies the difficulties encountered with empirical analyses of field data and laboratory experiments. Agent-based modeling is an interdisciplinary approach that uses computers to analyze and solve scientific problems by running computational experiments (Damaceanu, 2010). I propose using agent-based modeling in my dissertation for the following reasons. First, by using an agent-based model, I can manipulate all parameters that might influence agents' behaviors in a simulated auction. For example, I can randomly draw a few numbers from certain distributions and assign them to computer agents as their values for clicks. Otherwise, these values are private information and not accessible to researchers. I can also determine the learning and bidding strategies the advertisers will use in the simulations, and the agents will not deviate from the prescribed learning and bidding strategies or use heuristics. This will allow me to make "apples to apples" comparisons to determine exactly how the outcomes of the simulated auctions differ under different information disclosure policies. Second, fewer assumptions are needed for agent-based models. The distributions for advertisers' values need not be identical and independent, as previous studies have assumed (e.g., Gomes & Sweeney, 2014; Lucier & Paes Leme, 2011). The value distributions can be more complex in statistical character than what has been assumed for mathematical models (e.g., Gomes & Sweeney, 2014), such as being different and correlated with each other. Third, agent-based models can handle more complicated dynamics among advertisers. For instance, models with

more than three advertisers using fictitious play as their learning algorithm are analytically intractable (Yao et al., 2012). However, researchers can use agent-based modeling to solve models with $n > 3$ advertisers. Fourth, compared to human agents, computer agents do not suffer from fatigue as the number of rounds played increases. The advantages of agent-based modeling over empirical analysis of field data and laboratory experiments make agent-based modeling better suited to analyzing the various forms of GSP auctions examined in this dissertation.

3.3 Assumptions, Settings, and Scenarios for the Simulated GSP Auctions

I simulated three variants of GSP auctions, each of which incorporates one of the following three disclosure policies: no information disclosure, partial information disclosure, and perfect information disclosure. One type of agent acted as an advertiser in these simulated GSP auctions, and there were four assumptions imposed on all the auctions.

The following assumptions were imposed on all auctions:

1. I assumed that advertisers can find the funds for all bids that make economic sense (Edelman et al, 2007; Gummadi, et al, 2013). The reason for this assumption is that my dissertation is concerned with in-auction bidding behaviors. Other studies of GSP auctions focusing on in-auction bidding behaviors (e.g., Börgers et al., 2013; Thompson & Leyton-Brown, 2013; Varian, 2009) have also employed this assumption. When considering in-auction bidding behavior, models that assume that advertisers can find the funds for bids that make economic sense are considered reasonable approximations for what happens in keyword auctions. In real world keyword auctions, advertisers often choose multiple keywords and group them together as ad groups. Keywords in the same ad group share a common set of ad creatives. The ad groups are then grouped into ad

campaigns. Budgets are typically set at the campaign level. Large numbers of keywords are often used in ad campaigns, and these numbers can range from one to two thousand (Skiera, Eckert, & Hinz, 2009). Compared to the size of the entire budget, the bid for a single keyword is small.

2. The reserve price was assumed to be 10 cents and it applied to every advertiser. This is the same as what is used in the industry (Ostrovsky & Schwarz, 2011).
3. An advertiser's value for a click does not depend on the rank of the slot the advertiser's ad occupies. This assumption says that for any individual advertiser all clicks have equal value. This is the same assumption that is imposed in many other studies (e.g., Edelman et al., 2007; Gomes & Sweeney, 2014; Varian, 2007, 2009).
4. Advertisers do not bid above their true values in any round of an auction, which means that the upper bound for the bids of any advertiser is that advertiser's value for a click. This was a reasonable assumption because it does not make economic sense for advertisers to overbid. If advertisers overbid and obtain ad slots, they will lose money (Paes Leme, & Tardos, 2010a; 2010b).

The parameters that I controlled and set up for each of the simulated GSP auctions were: the size of the advertiser pool, the number of advertisers participating in an auction, the number of ad slots they bid for, the number of rounds advertisers play for each auction, and how each auction is initiated, which determines the bids advertisers submit for the first two rounds.

I set the number of ad slots to four. Findings from previous studies suggested that setting the number of ad slots at four is a reasonable choice (Noti, et al, 2014; Ostrovsky, & Schwarz, 2011). Given the number of ad slots, I set the number of advertisers vying for the four slots at five, one more than the number of slots. There were two reasons for choosing five advertisers.

First, the number of advertisers had to be at least five. If there were four or fewer advertisers, every advertiser could pin down all bids from the prior round under the partial information disclosure policy. In this case, each advertiser would have all the information needed to calculate all bids because after each round, the advertiser knows its own bid and the next highest rival's bid and the search engine announces the mean and standard deviation for the bids from the prior round. This is sufficient information to calculate the remaining bids. Because the search engine announces all bids under a perfect information disclosure policy, auctions with four advertisers vying for four slots under the perfect information disclosure policy and auctions with partial information disclosure would effectively be the same. Second, five advertisers in the model is more parsimonious than more advertisers and I found no compelling theory or prior research suggesting that increasing the number of advertisers beyond five would qualitatively change the character of the bidding competition.

The advertisers participating in an auction were selected in the following way. Before a simulated auction began, 10 advertisers were generated. The number of advertisers in the pool influenced how many rounds were required before the auctions converged to a steady state, but did not affect the outcomes of the auctions, which I showed through a sensitivity test (see Appendix III). Each advertiser in the pool had its own value for a click, which was determined by a random draw from a Beta or a Gamma distribution. I explain the reasons for choosing these distributions in the following section. Subsequently, five advertisers were either randomly selected round by round throughout the course of an auction or five advertisers were randomly selected to compete in all of an auction's rounds.

The pricing scheme for all simulated auctions was generalized second pricing. An advertiser who obtained a slot paid an amount per click equal to the bid of the next highest rival

weighted by the click decay rate, which is assumed to be a constant and the same for all ad slots and common knowledge to all market participants. Academic models typically assume that a search engine applies a weight of $\frac{CTR_{j+1}}{CTR_j}$, rate at which clicks are predicted to decline when moving from ad slot *j* to the next lower slot *j* +1, to the bid that purchases slot *j* +1. While actual weighting schemes are more complicated, this formula is considered to be a close approximation to the actual weighting formulas, which are proprietary. This dissertation further simplifies the weighting scheme by assuming that $\frac{CTR_{j+1}}{CTR_j}$ is common knowledge and a constant within each auction.

Each auction ran for 30 rounds. The number of rounds was determined through a sensitivity test (see Appendix II). The purpose of the sensitivity test was to find out how many rounds its took for an auction to reach a steady state, where the surplus and revenue levels from one round to the next do not differ by statistically significant amounts. I ran *t* tests and Mann-Whitney tests to compare the surplus and revenue levels of the auctions between adjacent rounds. The finding was that starting from round 15, there were no significant differences between adjacent rounds.

I initiated the auctions by having the agents submit two random bids, one for each of the first two rounds. The bids were random draws from the uniform distribution with the support (0.1,v_i), where v_i was advertiser *i*'s value for a click (0.1 was the reserve price). With this strategy, bidders' bids do not directly reveal their true values, making this a non-monotone bidding strategy. An alternative auction initiation mechanism could have had bidders submit bids that revealed their true values or some fixed percentage of their true values. Other ways for forming bids for the first two rounds of auction include submitting click values, submitting

reserve prices and submitting bids from previous auctions. Findings from the sensitivity tests used to test the alternative ways to initiate the auctions (see Appendix V) showed that the findings were not sensitive to the ways the auctions were initiated. The reason for two rounds of random bids was that when an advertiser uses the fictitious play learning algorithm (which is a learning algorithm in my study and will be explained later), the advertiser needs at least two rounds of bids or bid estimates as inputs.

Advertisers used a forward-looking response function to form their bids (Bu et al., 2008; Edelman et al., 2007), beginning with the third round of an auction:

$$F^i\left(b_t^{-i}\right) = \begin{cases} v_i - \frac{click_{k+1}}{click_k}(v_i - \widehat{b_{t-1}^{k+1}}), & 2 \le k \le K \\ v_i, & k = 1 \; or \; k > K \end{cases}.$$

In words, when using this bidding function to calculate the bid that will be submitted in round *t*, advertiser *i* first sorted all bid estimates for round *t* - 1. If the rank of advertiser *i*'s bid was 1 or exceeded the number of available slots, then advertiser *i* submitted a bid equal to its own value at round *t*. Otherwise, advertiser *i* submitted a bid equal to $v_i - \frac{click_{k+1}}{click_k}(v_i - \widehat{b_{t-1}^{k+1}})$, where v_i is advertiser *i*'s value, $click_k$ and $click_{k+1}$ were the numbers of clicks for slots *k* and *k* + 1, and $\widehat{b_{t-1}^{k+1}}$ was the estimated value for the bid submitted by the advertiser who occupied slot *k* + 1 in round *t* - 1.

Under the perfect information disclosure policy, the search engine announced $\widehat{b_{t-1}^{k+1}}$ after round *t* - 1 was concluded. However, under the partial information disclosure policy and the imperfect information disclosure policy, advertiser *i* estimates the bids. The reason for using this bidding function is that it reflects the rational manipulative behaviors of utility maximizing

advertisers in GSP auctions (Bu et al., 2008; Tsung, Ho, & Lee, 2008). As Varian (2007) and Edelman and colleagues (2007) suggested that the strategy is a plausible choice because these bids ensure that advertisers will obtain the right ad slots according to their values, which means that the ordering of slots acquired by advertisers corresponds with the ordering of advertisers' click values.

The bid estimation method was as follows. Under the partial information disclosure policy, after a round of the auction, an advertiser knows its own bid, its next highest rival's bid, and is given the mean and standard deviation for the five bids. With this information, an advertiser calculates the mean for the three unknown bids and uses this as its estimate for the bid that ranks in the middle of the three unknown bids. Then, the advertiser has two unknown bids and uses a quadratic formula to calculate the remaining bids. Under the no information disclosure policy, after one round of an auction an advertiser knows only its own bid and its next highest rival's bid. The advertiser calculates the difference between its bid and the next highest rival's bid and add this distance to its bid to calculate the bid of the advertiser who ranks immediately above it, and also subtract this distance from the next highest rival's bid to calculate the bid of the advertiser who ranks immediately below the next highest rival.

In order to assess the effects of the information disclosure policy, I ran auctions with the above mentioned settings under different information disclosure polices in different scenarios. These scenarios were constructed by varying and combing the values of several parameters. These parameters included whether the mix of advertisers changed over rounds, click value distributions, the statistical character of relationships among bidders' value distributions, learning algorithm, whether advertisers used a common algorithm or different learning algorithms, and the click decay rate for ad slots.

The first parameter was whether the mix of advertisers changed over rounds. In real world auctions, search ad auctions can have the same advertisers or different advertisers competing against each other from one round to the next. In a small ad market, such as the local plumbing services market, there is only a small pool of competitors and the likelihood that they will encounter one another repeatedly is high. However, even though the population of advertisers remains the same in a larger market, the pool is larger and there will be more variation in the mix of competitors across rounds of auctions. Obviously, changing the mix of advertisers introduces uncertainty for both advertisers and the search engine. Thus, it is important to determine whether information disclosure policies could make a difference in both types of scenarios. The procedures were as follows for when the set of advertisers was fixed for all rounds of an auction: five advertisers were randomly drawn from the pool of advertisers before an auction began and these advertisers bid against each other throughout the 30 rounds of a simulated auction. When the set of advertisers varied round by round before the beginning of each round, five advertisers were randomly drawn from the 10-advertiser pool each round to bid against each other. Thus, the mix of advertisers was likely different from round to round.

Advertisers in search ad auctions can have different values. Variable factors such as cost of production and profitability determine how much an advertiser values a click. The common practice in the field when values vary among advertisers is to assume that the values for advertisers are random realizations from some distribution. These distributions often are monotone hazard rate distributions (Myerson, 1981; Golrezaei & Nazerzadeh, 2013; Deltas & Jeitschko, 2007; Edelman & Schwarz, 2010). The reason for this kind of distribution is that researchers using it could design a mechanism that assigns auction items in accordance with bidders' values. Additionally, the class of monotone hazard rate (MHR) distributions is wide

enough and includes many common distributions (such as uniform, and normal) (Barlowgom & Marshall, 1964). I chose Beta and Gamma distributions for my research. The reasons I chose these distributions are two-fold. First, these two types of distributions can have supports that are reasonable for keyword auctions. The Beta distribution employed has a support defined by two nonnegative values, $[v_l, v_u]$, where v_l is the lower bound of the support and v_u is the upper bound. Intuitively, this means that an advertiser values a click by at least v_l and at most by v_u. In all of the simulations, v_l was equal to or greater than the reserve price, 0.1, so that no advertiser was precluded from bidding by the reserve price. The Gamma distribution employed has a support range from the reserve price to positive infinity, $[v_l, \infty]$. Second, these two distributions are general. For example, the uniform distribution is a special case of the Beta distribution, and the exponential distribution is a special case of the Gamma distribution.

The statistical character of relationships among bidders' value distributions is another parameter that I used to construct the scenarios. There were four types of relationship. For the first type of relationship, the value distributions were identical and independent. For the second type of relationship, the value distributions were identical but correlated. For the third type of relationship, the value distributions were different and independent from each other. For the fourth type of relationship, the value distributions were different and correlated. Identical means that the distributions have the same mean and standard deviation. The mean of the distributions was 25 and the standard deviation was 11 when they were identical Beta distributions. The means ranged from 5 to 50, equally spaced, and the standard deviations ranged from 2 to 20, equally spaced, when they were different Beta distributions. The mean value of the distributions was 100 and the standard deviation was 70 when they were identical Gamma distributions. The means ranged from 20 to 200, equally spaced, and the standardized deviations ranged from 15 to

150, equally spaced, when they were different Gamma distributions. The correlation between these distributions was set to 0 when the distributions were independent, and was set to 0.5 when the distributions were correlated. The figures in appendix I show that I generated the intended distributions, as well as the statistical character of the relationships among the distributions (see figures in Appendix I for illustration.).

The computer agents used two learning algorithms for learning about competitors' bids: Cournot learning and fictitious play (Brown, 1951; Yao, Chen, & Liu, 2012). There are many algorithms that can model a learning process, such as Hart and Mas-Collel's (2000) regret matching strategy, or the multiplicative weight-updating strategy (Littlestone & Warmuth, 1994). Advertisers using these algorithms observe the bidding behaviors of their opponents and modify their own bids in such a way that they can maximize their total return up to that point. I chose Cournot learning and fictitious play as the learning algorithms for this dissertation because they are able to capture the intuition behind advertisers' behaviors in keyword auctions. In these auctions, advertisers attempt to learn about opponents' values and bidding strategies and modify their own bidding strategies to increase the difference between their click values and the prices they pay for a click. The underlying assumption is that the information bidders gain about opponents can be used to improve their payoffs in later stages of the game. For this reason, Cournot learning and fictitious play have also been studied in research on GSP auctions (e.g., Bu et al., 2008; Edelman et al., 2007; Yao, Chen, & Liu, 2012). Another advantage of these learning algorithms is that they do not have to produce an equilibrium, which reflects the reality of GSP auctions better than imposing an equilibrium refinement concept because real world GSP auctions often do not converge to an equilibrium (Börgers et al., 2013).

When Cournot learning is used as the learning algorithm, advertiser *i* calculates its bid using the forward-looking response function, assuming that the other advertisers will submit the same bids as in round *t* - 1.

When fictitious play is used as the learning algorithm, advertiser *i* devises bids given its beliefs about the distributions of opponents' bids and updates these beliefs by using the available bid information. To do this, advertiser *i* records the bid histories of its opponents and uses these bid histories to update its estimates of opponents' bids. More specifically, advertiser *i's* bid in round *t* (t > 2) is calculated using an exponential moving average: $b_t = \gamma(b_{t-1}) + (1-\gamma)b_{t-2}$, where b_{t-1} is the bid for round *t* - 1, b_{t-2} is the bid for round *t* - 2, and γ is the constant smoothing factor, which can range from 0 to 1. A larger γ discounts older observations faster. For the purposes of this dissertation, γ varied among the following values: 0, 0.25, 0.5, 0.75.

Most bidding strategy studies (Cary et al., 2007; Zhou & Lukose, 2007; Liang & Qi, 2007; Yao, Chen & Liu, 2012) assume that all advertisers employ the same learning algorithm. However, it is possible for advertisers to use different learning algorithms in real keyword auctions. Thus, I explored scenarios where advertisers used a common learning algorithm or different learning algorithms. In scenarios where advertisers used different learning algorithms, an advertiser was randomly assigned either the Cournot learning or the fictitious play learning algorithm. After each round was concluded, an advertiser looked at its payoff. If the payoff was the same or improved over the previous round, the advertiser used the same learning algorithm. Otherwise, it changed its learning algorithm to the one not used in the round just concluded.

For real world search results pages, the clicks generated from the slots decrease as slot rankings increase. The ratio between the clicks from adjacent slots *j* + 1 and *j* is defined as the

click decay rate and can range from 0 to 1. In reality, click decay rates vary wildly across keywords and slot rankings. “First Rate,” a Google adword partner, found that on average, the top slot received 17% of the clicks, while the second slot received 13% of the clicks (Sandberg, 2012). However, another study (Receptional, 2012) found that the top slot absorbed 53% of the clicks and the second slot received only 15% of the clicks. I explored scenarios for constant click decay rates with the following values: 0, 0.25, 0.5, 0.75, and 1. The first slot was set to produce 100 clicks, and the clicks from subsequent slots were calculated based on the click decay rates.

Three types of information disclosure policies were investigated: perfect information disclosure, partial information disclosure, and no information disclosure. With perfect information disclosure, the search engine provided the bidders with the bids from the prior round. With partial information disclosure, the search engine announced the mean and standard deviation for the bids from the prior round. For no information disclosure, the search engine did not reveal any information about bids from the prior round.

In total, after allowing the values of the various parameters to vary independent, there were 800 scenarios for which auctions were carried out. For each scenario, I ran 600 simulated auctions. A sensitivity test indicated that the observed ranges for auctions’ surplus and revenue of the auctions when the number changed very little when the number of auctions reached 600 (see Appendix IV). For each auction I ran 30 rounds. Thus, there were 480,000 simulated auctions and 1,440,000 rounds in total.

3.4 Metrics for Measuring Performance

For each round of each auction, I evaluated the performance from three perspectives: search engine revenue ratio, surplus ratio and yield.

I used the following formula to calculate search engine revenue:

$$REV(b, click) = \sum_{i=1}^{m} b_{i+1} click_i.$$

That is, the revenue for each round is the sum of the products of the per click prices paid and the corresponding numbers of clicks, where b_{i+1} is the bid of advertiser *i*'s next highest rival, which is also is advertiser *i*'s price, and $click_i$ is the number of clicks that advertiser *i* acquires by occupying slot *i*.

Directly comparing revenue levels is not sensible because a difference in revenue levels between auctions could be due to a difference in the participating advertisers' values. Furthermore, even if the absolute value of the revenue of an auction is higher, the search engine revenue could be small compared to the maximum revenue attainable in the auction. Therefore, I calculated the revenue ratio for a round by dividing the realized revenue by the maximum attainable revenue for the round. The maximum attainable revenue for a round was calculated by assuming that advertisers submit their click values as bids and the auction format is a GSP auction.

The formula for a bidding round's revenue ratio is:

$$Revenue\ ratio = \frac{REV(b,\ \ click)}{REV_{opt}(v,\ \ click)},$$

where *b* is the bid profile for advertisers who participated in that round of the auction, *v* is the value profile for the advertisers who participated in that round of auction, and *click* is the click profile for the ad slots.

Next, I calculated the surplus ratio, the ratio between the realized surplus and the maximum achievable surplus, for each round of each auction. I used the surplus ratio as a measure of efficiency and calculated it for each round using the following formula:

$$Surplus\ ratio = \frac{Surplus(v,p)}{OPT(v,p\ )},$$

where v is the value profile for the advertisers who participated in that round of the auction and p is the price profile of the advertisers who participated in that round. The numerator is the realized surplus for that round and the denominator is the maximum attainable surplus for that round of the auction. Because the maximum attainable surplus for a GSP auction equals that of the corresponding VCG auction (Edelman et al., 2007; Varian, 2007), the surplus of the corresponding VCG auction is used as the numerator in the ratio. The realized surplus of a round is the sum of the advertisers' payoffs and the search engine's revenue:

$$Surplus(v,p) = \sum_{i=1}^{k}(v_i - p_i)click_i + REV_j(b, click),$$

where v_i is advertiser i's value, p_i is advertiser i's price which is the next highest advertiser's bid, $click_i$ is the number of clicks for the slot occupied by advertiser i, b is the bid profile, and k is the number of ad slots.

I also calculated the yield of a round by dividing the search engine's revenue by the surplus for the round. The purpose of calculating this ratio was to measure the percentage of the entire surplus captured by the search engine as its revenue. The formula is as follows:

$$Yield = \frac{Revenue}{Surplus}$$

3.5 Analysis Plan

To answer my research questions, I compared the search engine revenue ratios, yields, and surplus ratios under different information disclosure policies across the different scenarios. Specifically, I ran 800 regression models for each performance metric using data generated by the auctions after excluding data from the first 20 rounds of each auction. The reason for using only use the last 10 rounds of auction data is that by this point, I could be confident that the stability of the auctions has been achieved, similar to what we would expected to find in real world auctions that have been running long enough to reach their equilibria assuming such equilibria exist.

RESULTS

4.1 Revenue Ratio

The average revenue ratio for auctions under the no information disclosure policy is 0.8009, and revenue ratios range from 0.6427 to 0.9364 (see Figure 2). These results indicate that when no bid information from a prior round is disclosed, the search engine's revenue reaches 80.09% of the maximum attainable level on average, 64.27% of the maximum attainable level in the worst case scenarios, and 93.64% of the maximum attainable level in the best case scenarios.

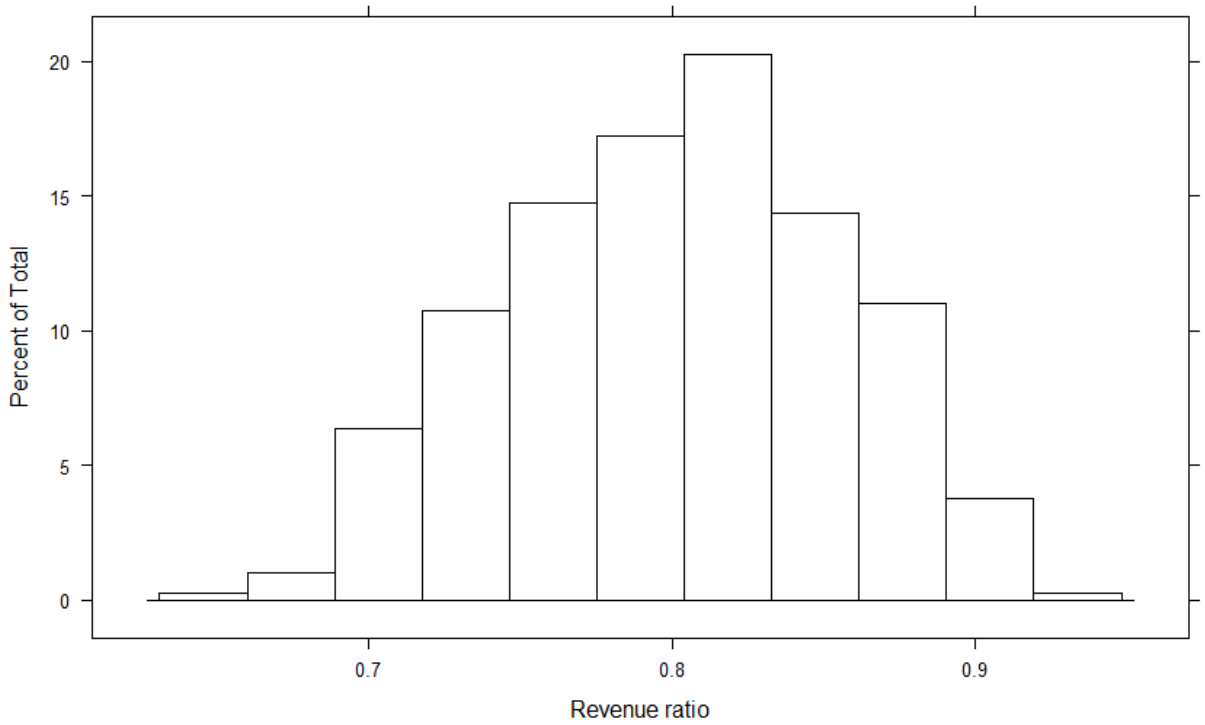

Figure 2: Histogram for revenue ratio under no information disclosure policy

The average revenue ratio for auctions under the partial information disclosure policy is 0.8315, and the ratios range from 0.7124 to 0.9541 (see Figure 3). These results indicate that when the mean and the standard deviation of the bids from the prior round are disclosed, the

search engine's revenue reaches 83.15% of the maximum attainable level on average, 71.24% of the maximum attainable level in the worst case scenarios, and 95.41% of the maximum attainable level in the best case scenarios.

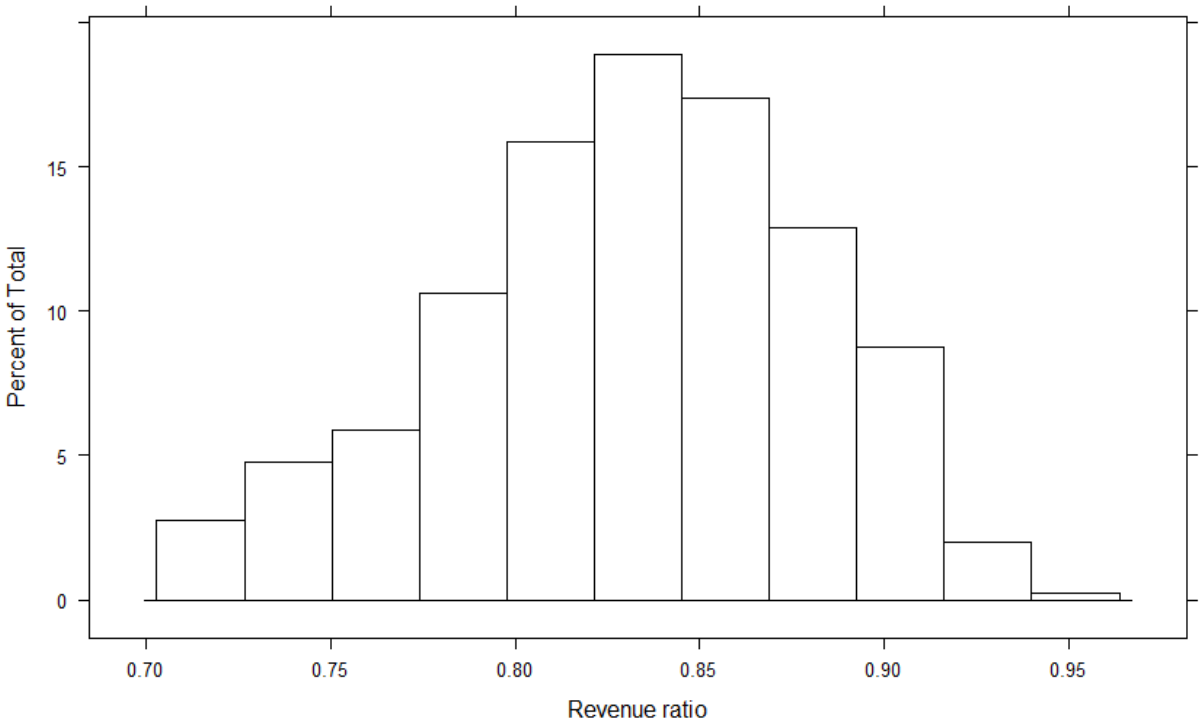

Figure 3: Histogram for revenue ratio under partial information disclosure policy

The average revenue ratio for auctions under the perfect information disclosure policy is 0.8058, and revenue ratios range from 0.6281 to 0.9388 (see Figure 4). These results indicate that when the bids from the prior round are disclosed, the search engine's revenue reaches 80.58% of the maximum attainable level on average, 62.81% of the maximum attainable level in the worst case scenarios, and 93.88% of the maximum attainable level in the best case scenarios.

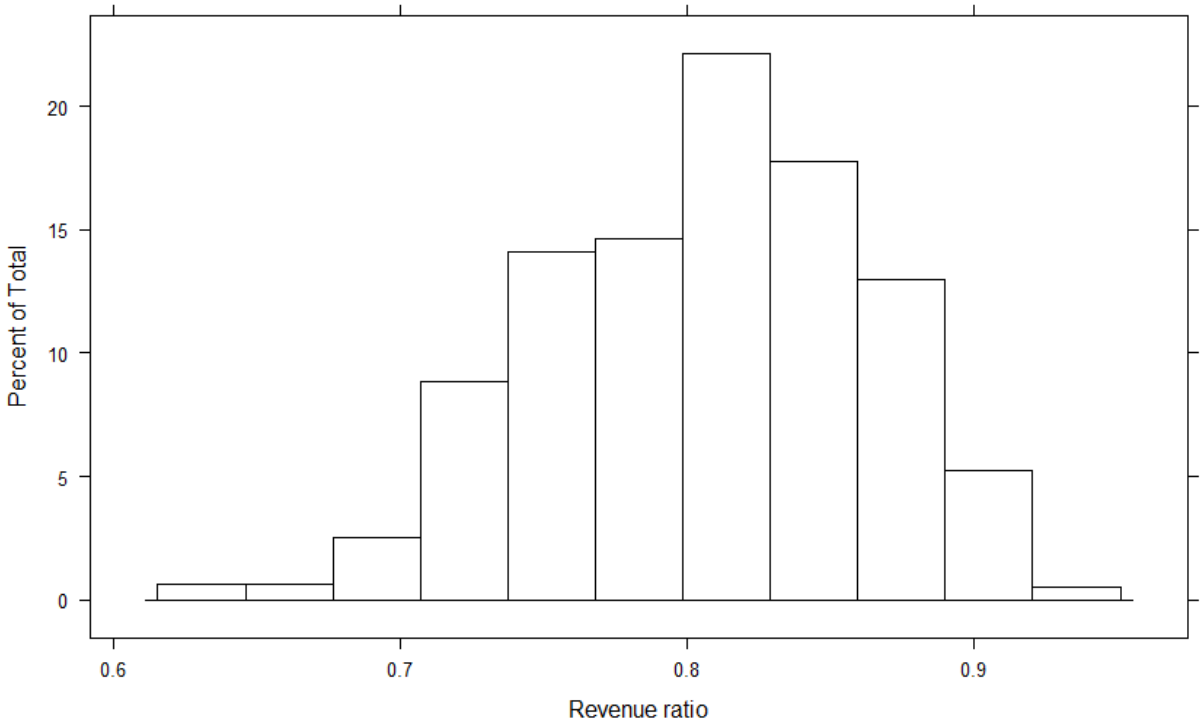

Figure 4: *Histogram for revenue ratio under perfect information disclosure policy*

These results show that under the partial information disclosure policy, the search engine revenue levels are generally closer to the maximum attainable levels than under the two other information disclosure policies. Although, on average, the revenue ratios under the perfect information disclosure policy are higher than those under the no information disclosure policy, the difference in the revenue ratio between the two information disclosure policies is very small.

Table 2 shows the revenue ratios across the different information disclosure policies for each of the two ways the mix of participating auction advertisers is determined. The results indicate that the patterns of revenue ratios across different information disclosure polices are similar for scenarios in which the mix of advertisers is fixed and for scenarios in which the mix of advertisers varies. Under the partial information disclosure policy, search engine revenue ratios are higher than those for the two other information disclosure policies.

Table 2: Revenue ratio comparison across different information disclosure policies and mix of participating advertisers

	No information disclosure policy	Partial information disclosure policy	Perfect information disclosure policy
Minimum			
Fixed mix of advertisers	0.6427	0.7139	0.6281
Varied mix of advertisers	0.6908	0.7124	0.7018
Maximum			
Fixed mix of advertisers	0.9364	0.9541	0.9388
Varied mix of advertisers	0.9181	0.9245	0.9177
Mean			
Fixed mix of advertisers	0.7839	0.8291	0.7873
Varied mix of advertisers	0.8178	0.8339	0.8243

Table 3 gives the revenue ratios for the various combinations of information disclosure policies and value distributions. The revenue ratios are higher under the partial information disclosure policy and the revenue ratios are higher when the distributions for advertisers' values are drawn from Beta distributions.

Table 3: Revenue ratio comparison across different information disclosure policies and value distributions

	No information disclosure policy	Partial information disclosure policy	Perfect information disclosure policy
Minimum			
Beta distribution	0.6555	0.7132	0.6338
Gamma distribution	0.6427	0.7124	0.6281
Maximum			
Beta distribution	0.9364	0.9541	0.9388
Gamma distribution	0.9141	0.9350	0.9116
Mean			
Beta distribution	0.8102	0.8355	0.8152
Gamma distribution	0.7915	0.8276	0.7964

Table 4 shows that revenue ratios are higher under the partial information disclosure policy for both value distributions considered. Furthermore, when the value distributions are identical and correlated with one another, the revenue ratios are higher than when the value distributions are identical and independent, different and independent, and different and correlated.

Table 4: Revenue ratio comparison across different information disclosure policies and statistical character of relationships among value distributions

	No information disclosure policy	Partial information disclosure policy	Perfect information disclosure policy
Minimum			
Identical and correlated	0.7334	0.7480	0.6872
Identical and independent	0.7046	0.7300	0.6742
Different and correlated	0.6755	0.7167	0.6454
Different and independent	0.6427	0.7124	0.6281
Maximum			
Identical and correlated	0.9364	0.9541	0.9388
Identical and independent	0.9141	0.9410	0.9146
Different and correlated	0.8795	0.9282	0.8792
Different and independent	0.8530	0.9331	0.8553
Mean			
Identical and correlated	0.8503	0.8695	0.8554
Identical and independent	0.8139	0.8378	0.8183
Different and correlated	0.7793	0.8150	0.7842
Different and independent	0.7614	0.8050	0.7668

In general, disclosing the means and standard deviations of bids from the prior rounds (partial information disclosure) results in higher revenue ratios than do the other bid information disclosure policies (no information disclosure and perfect information disclosure), regardless of which learning algorithm advertisers used (see Table 5).

Table 5: Revenue ratio comparison across different information disclosure policies and learning algorithms

	No information disclosure policy	Partial information disclosure policy	Perfect information disclosure policy
Minimum			
Cournot	0.6427	0.7378	0.7075
Fictitious ($\gamma = 0$)	0.6716	0.7708	0.6979
Fictitious ($\gamma = 0.25$)	0.7162	0.7630	0.6281
Fictitious ($\gamma = 0.5$)	0.6912	0.7809	0.6741
Fictitious ($\gamma = 0.75$)	0.6620	0.7124	0.6338
Maximum			
Cournot	0.9181	0.9245	0.9388
Fictitious ($\gamma = 0$)	0.9364	0.9297	0.9293
Fictitious ($\gamma = 0.25$)	0.9351	0.9173	0.9061
Fictitious ($\gamma = 0.5$)	0.9048	0.9541	0.9220
Fictitious ($\gamma = 0.75$)	0.9089	0.9131	0.9084
Mean			
Cournot	0.7964	0.8181	0.8097
Fictitious ($\gamma = 0$)	0.8131	0.8507	0.8207
Fictitious ($\gamma = 0.25$)	0.8156	0.8417	0.8117
Fictitious ($\gamma = 0.5$)	0.8087	0.8522	0.8156
Fictitious ($\gamma = 0.75$)	0.7706	0.7948	0.7713

Under the partial information disclosure policy, the auctions generated revenues that were closer to the maximum attainable levels than under the two other information disclosure policies (see Table 6).

Table 6: Revenue ratio comparison across different information disclosure policies for common and different learning algorithms

	No information disclosure policy	Partial information disclosure policy	Perfect information disclosure policy
Minimum			
Common algorithm	0.6427	0.7124	0.6281
Different algorithms	0.7234	0.7378	0.7178
Maximum			
Common algorithm	0.9364	0.9541	0.9388
Different algorithms	0.9089	0.9131	0.9084

Table 6 (cont'd)			
Mean			
Common algorithm	0.7935	0.8378	0.801
Different algorithms	0.8082	0.8252	0.8106

The results in Table 7 show that regardless of the click decay rates, disclosing mean and standard deviation of bids from prior round almost always enables search engine to generate higher revenue ratios than they generate with the other two information disclosure policies[4].

Table 7: Revenue ratio comparison across different information disclosure policies for different click decay rates

	No information disclosure policy	Partial information disclosure policy	Perfect information disclosure policy
Minimum			
decay = 0	0.6427	0.7183	0.6281
decay = 0.25	0.6782	0.7166	0.6822
decay = 0.5	0.6716	0.7124	0.7016
decay = 0.75	0.6620	0.7129	0.6741
decay = 1	0.6825	0.7134	0.6890
Maximum			
decay = 0	0.9181	0.9331	0.9177
decay = 0.25	0.9131	0.9201	0.9293
decay = 0.5	0.9351	0.9346	0.9388
decay = 0.75	0.9062	0.9344	0.9239
decay = 1	0.9364	0.9541	0.9041
Mean			
decay = 0	0.8011	0.8414	0.8023
decay = 0.25	0.8018	0.8334	0.8119
decay = 0.5	0.7991	0.8291	0.8113
decay = 0.75	0.7983	0.8273	0.8012
decay = 1	0.8040	0.8264	0.8023

[4] When the click decay rate is 0.25, the maximum value of the revenue ratio under perfect information disclosure is greater than that for partial information disclosure.

In summary, revenue ratios are nearly always higher under the partial information disclosure policy than for the two other information disclosure policies. But the differences between those under the no information disclosure policy and those under the perfect information disclosure policies are very small.

4.2 Yield

The mean of the yields for auctions under the no information disclosure policy is 0.5763, and the yields range from 0.3029 to 0.8401 (see Figure 5). These results indicate that when no bid information from the prior round is disclosed, the search engine extracts 57.63% of the surplus as its revenue on average, 30.29% of the maximum attainable level in the worst case scenarios, and 84.01% of the maximum attainable level in the best case scenarios.

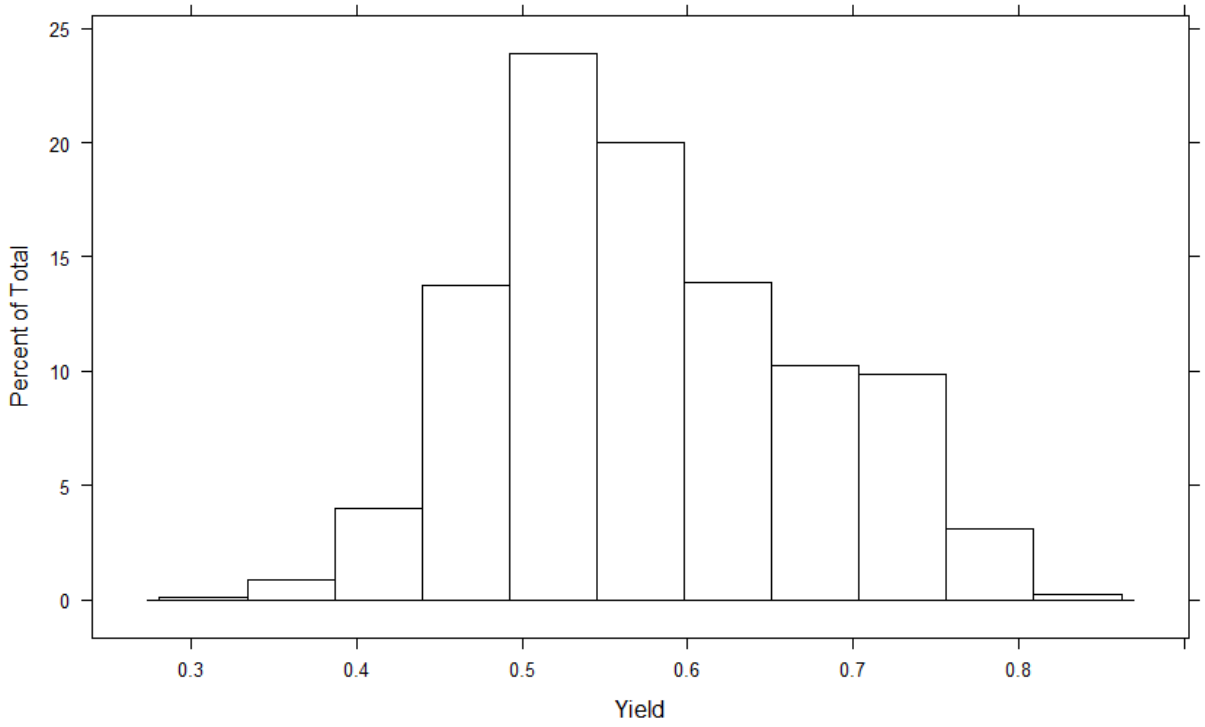

Figure 5: Histogram for yield under no information disclosure policy

The mean of the yields for auctions under the partial information disclosure policy is 0.5959, and the yields range from 0.3773 to 0.8448 (see Figure 6). These results indicate that when mean and standard deviation of bids are disclosed from the prior round, the search engine extracts 59.59% of the surplus as its revenue on average, 37.73% of the surplus as its revenue in the worst case scenarios, and 84.48% of the surplus as its revenue in the best case scenarios.

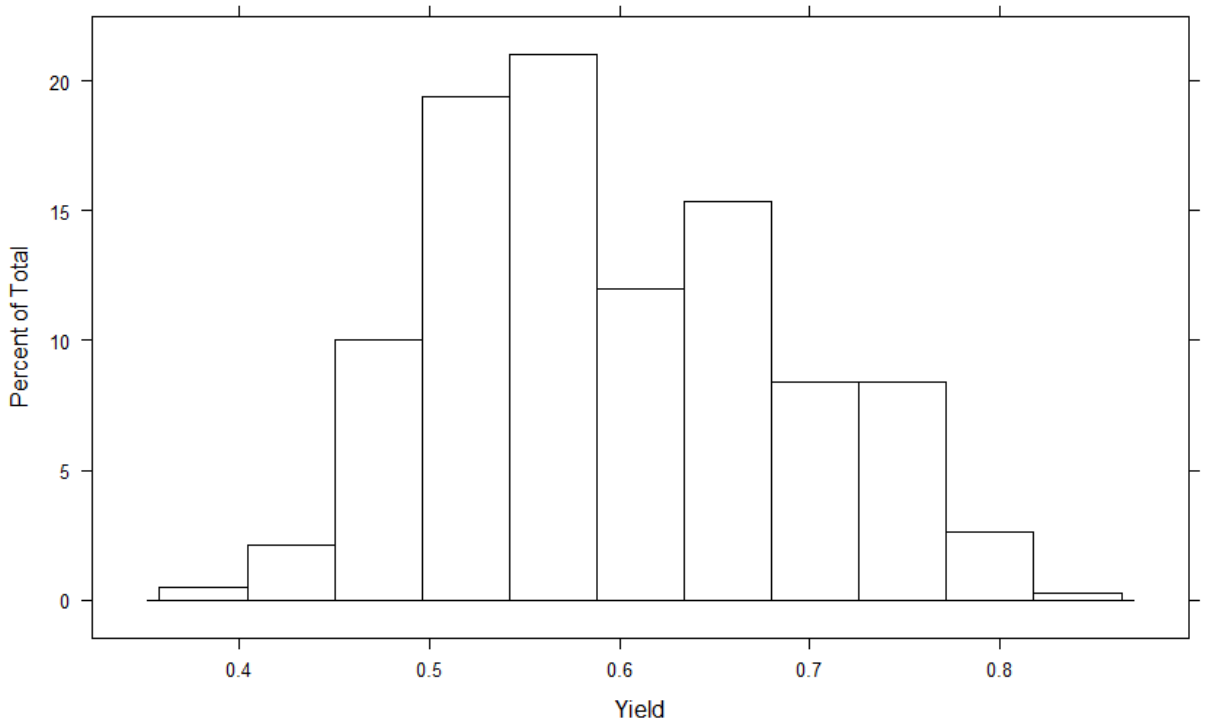

Figure 6: Histogram for yield under partial information disclosure policy

The mean yield for auctions under the perfect information disclosure policy is 0.5797, and the yields range from 0.2812 to 0.8145 (see Figure 7). These results indicate that when the bids from the prior round are disclosed, the search engine extracts 57.97% of the surplus as its revenue, 28.12% of the surplus as its revenue in the worst case scenarios, and 81.45% of surplus as its revenue in the best case scenarios.

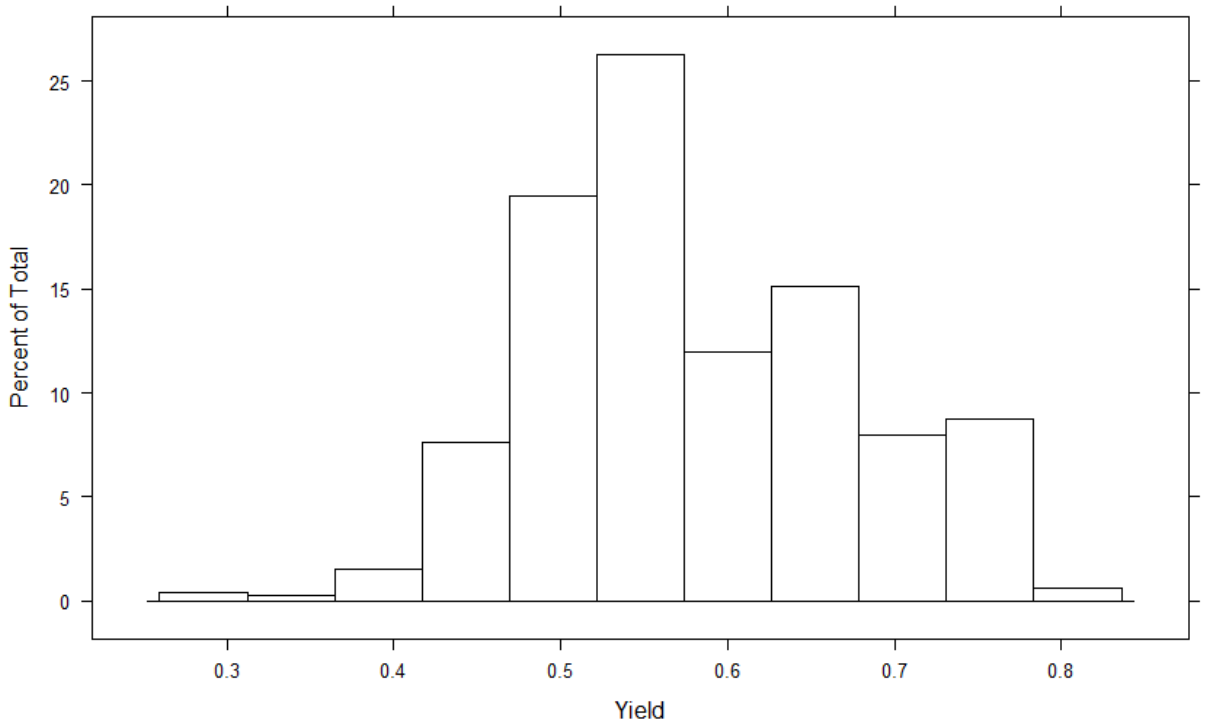

Figure 7: *Histogram for yield under perfect information disclosure policy*

Under the partial information disclosure policy, the search engine extracts larger portions of auction surplus as its revenue than under the two other information disclosure policies. This result is consistent across scenarios, regardless of whether the mix of advertisers is fixed or varies (see Table 8).

Table 8: Yield comparison across different information disclosure policies and mix of participating advertisers

	No information disclosure policy	Partial information disclosure policy	Perfect information disclosure policy
Minimum			
Fixed mix of advertisers	0.3029	0.3773	0.2812
Varied mix of advertisers	0.3997	0.4157	0.4083
Maximum			
Fixed mix of advertisers	0.8401	0.8448	0.8145
Varied mix of advertisers	0.7832	0.7894	0.7858

Table 8 (cont'd)			
Mean			
Fixed mix of advertisers	0.5657	0.5942	0.568
Varied mix of advertisers	0.5869	0.5975	0.5913

For both value distributions, the yields are always highest with the partial information disclosure policy while, holding information disclosure policy constant, yields are always higher for the Beta distribution than for the Gamma distribution (see Table 9).

Table 9: Yield comparison across different combinations of information disclosure policies and value distributions

	No information disclosure policy	Partial information disclosure policy	Perfect information disclosure policy
Minimum			
Beta distribution	0.3611	0.4047	0.3085
Gamma distribution	0.3029	0.3773	0.2812
Maximum			
Beta distribution	0.8401	0.8448	0.8145
Gamma distribution	0.7663	0.7450	0.6964
Mean			
Beta distribution	0.6134	0.6305	0.6166
Gamma distribution	0.5392	0.5612	0.5427

When the distributions are identical and correlated with each other, the yield levels are higher than those for the three other types of statistical relationships among the value distributions examined (see Table 10).

Table 10: Yield comparison across different information disclosure policies for four different relationships among value distributions

	No information disclosure policy	Partial information disclosure policy	Perfect information disclosure policy
Minimum			
Identical and correlated	0.4793	0.558	0.5105
Identical and independent	0.3907	0.4678	0.3882
Different and correlated	0.3666	0.4164	0.3349
Different and independent	0.3029	0.3773	0.2812
Maximum			
Identical and correlated	0.8401	0.8448	0.8145
Identical and independent	0.7857	0.8065	0.7744
Different and correlated	0.6866	0.7032	0.6781
Different and independent	0.6356	0.6588	0.646
Mean			
Identical and correlated	0.6814	0.6958	0.6852
Identical and independent	0.6099	0.6267	0.6132
Different and correlated	0.5312	0.5537	0.5348
Different and independent	0.4855	0.51	0.4883

On average, disclosing bid statistics helps a search engine capture larger portions of surplus (see Table 11).

Table 11: Yield comparison across different information disclosure policies and learning strategies

	No information disclosure policy	Partial information disclosure policy	Perfect information disclosure policy
Minimum			
Cournot	0.3029	0.4039	0.412
Fictitious ($\gamma = 0$)	0.3429	0.4385	0.2996
Fictitious ($\gamma = 0.25$)	0.4185	0.3902	0.2812
Fictitious ($\gamma = 0.5$)	0.3533	0.4067	0.3719
Fictitious ($\gamma = 0.75$)	0.3501	0.3773	0.4083

Table 11 (cont'd)			
Maximum			
Cournot	0.7832	0.7986	0.7946
Fictitious ($\gamma = 0$)	0.8401	0.8061	0.7858
Fictitious ($\gamma = 0.25$)	0.8226	0.8024	0.777
Fictitious ($\gamma = 0.5$)	0.7744	0.8448	0.8145
Fictitious ($\gamma = 0.75$)	0.779	0.7844	0.7757
Mean			
Cournot	0.5664	0.599	0.5821
Fictitious ($\gamma = 0$)	0.5876	0.6012	0.5841
Fictitious ($\gamma = 0.25$)	0.5908	0.5986	0.5751
Fictitious ($\gamma = 0.5$)	0.5742	0.6065	0.589
Fictitious ($\gamma = 0.75$)	0.5626	0.5741	0.5681

An advantage for the partial information disclosure policy also appears when comparing scenarios when the advertisers use common and different learning algorithms (see Table 12).

Table 12: Yield comparison across different information disclosure policies and common and different learning algorithms

	No information disclosure policy	Partial information disclosure policy	Perfect information disclosure policy
Minimum			
Common algorithm	0.3029	0.3773	0.2812
Different algorithm	0.4189	0.4325	0.4339
Maximum			
Common algorithm	0.5694	0.5984	0.5740
Different algorithm	0.5832	0.5933	0.5853
Mean			
Common algorithm	0.5694	0.5984	0.5740
Different algorithm	0.5832	0.5933	0.5853

On average, the yields for a search engine are larger when the information disclosure policy is partial information disclosure across all scenarios in which the click decay rates vary (see Table 13).

Table 13: Yield comparison across different information disclosure policies and different click decay rates

	No information disclosure policy	Partial information disclosure policy	Perfect information disclosure policy
Minimum			
decay = 0	0.3029	0.4067	0.4499
decay = 0.25	0.3659	0.3912	0.2812
decay = 0.5	0.3429	0.4362	0.4226
decay = 0.75	0.3501	0.4235	0.3667
decay = 1	0.3880	0.3773	0.2996
Maximum			
decay = 0	0.8001	0.7986	0.8145
decay = 0.25	0.7805	0.7854	0.8
decay = 0.5	0.8226	0.8024	0.7905
decay = 0.75	0.7548	0.8354	0.7667
decay = 1	0.8401	0.8448	0.7405
Mean			
decay = 0	0.6028	0.6201	0.6144
decay = 0.25	0.5844	0.5958	0.5878
decay = 0.5	0.5769	0.6029	0.5825
decay = 0.75	0.5595	0.5919	0.565
decay = 1	0.5578	0.5687	0.5487

4.3 Surplus Ratio

The average surplus ratios for auctions under the no information disclosure policy is 0.9881, and the surplus ratios range from 0.9124 to 1 (see Figure 8). These results indicate that when no bid information from a prior round is disclosed, the surplus from the auctions reached 98.81% of the maximum achievable level on average, 91.24% of the maximum attainable level in the worst case scenarios, and 100% of maximum attainable level in the best case scenarios.

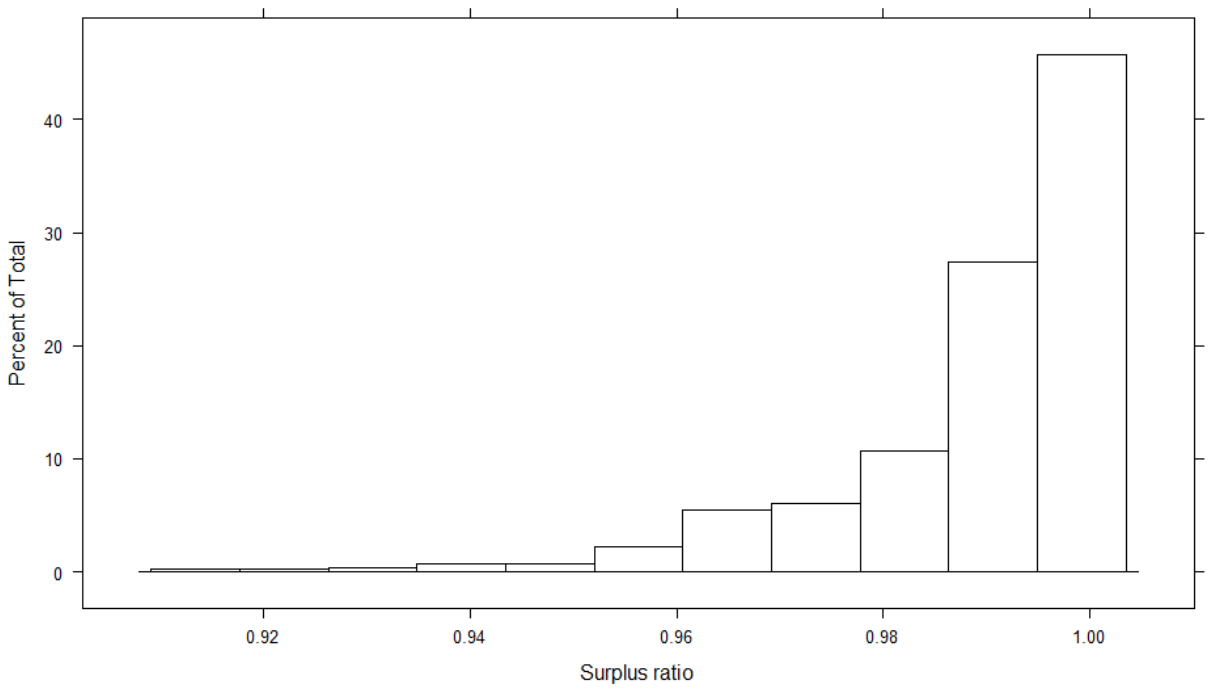

Figure 8: Histogram for surplus ratio under no information disclosure policy

The average surplus ratios for auctions under the partial information disclosure policy is 0.9881 and the surplus ratios range from 0.9058 to 1 (see Figure 9). These results indicate that when the mean and standard deviation for the bids from a prior round are disclosed, the surplus from the auctions reached 98.81% of the maximum attainable level on average, 90.58% of the maximum attainable level in the worst case scenarios, and 100% of the maximum attainable level in the best case scenarios.

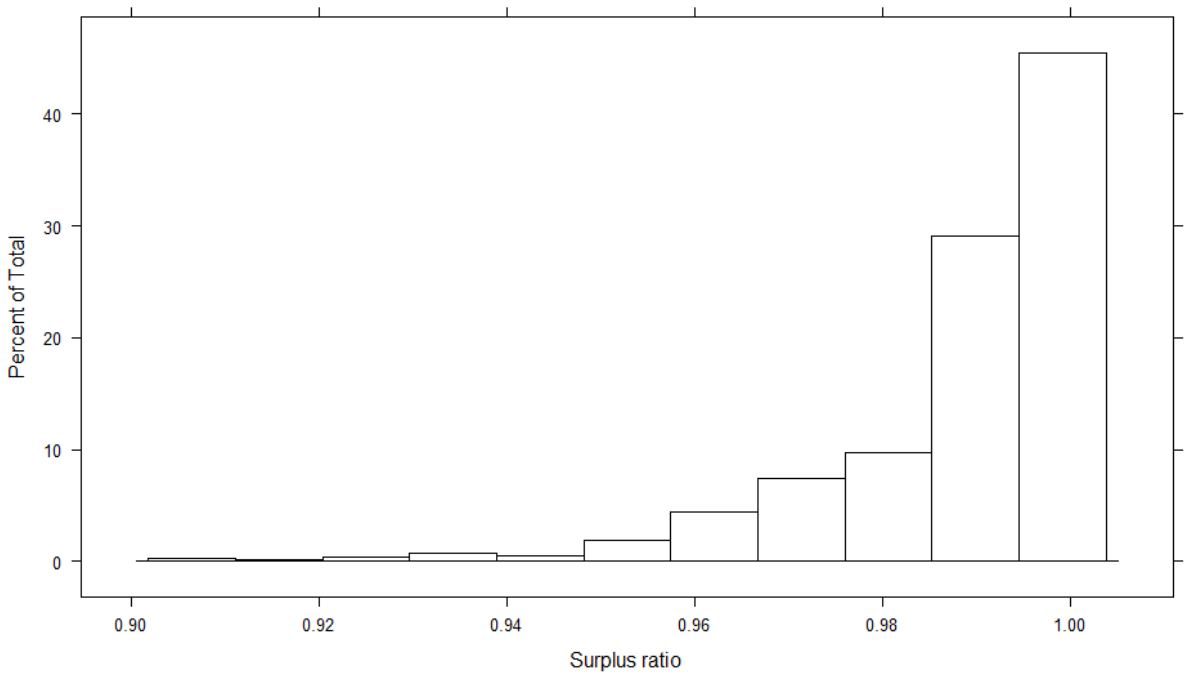

Figure 9: Histogram for the surplus ratio under partial information policy

The average surplus ratios for auctions under the perfect information disclosure policy is 0.9881 and the surplus ratios range from 0.9174 to 1 (see Figure 10). These results indicate that when all bid from a prior round are disclosed, the surplus from the auctions reaches 98.81% of the maximum attainable level on average, 91.74% of the maximum attainable level in the worst case scenarios, and 100% of maximum attainable level in the best case scenarios.

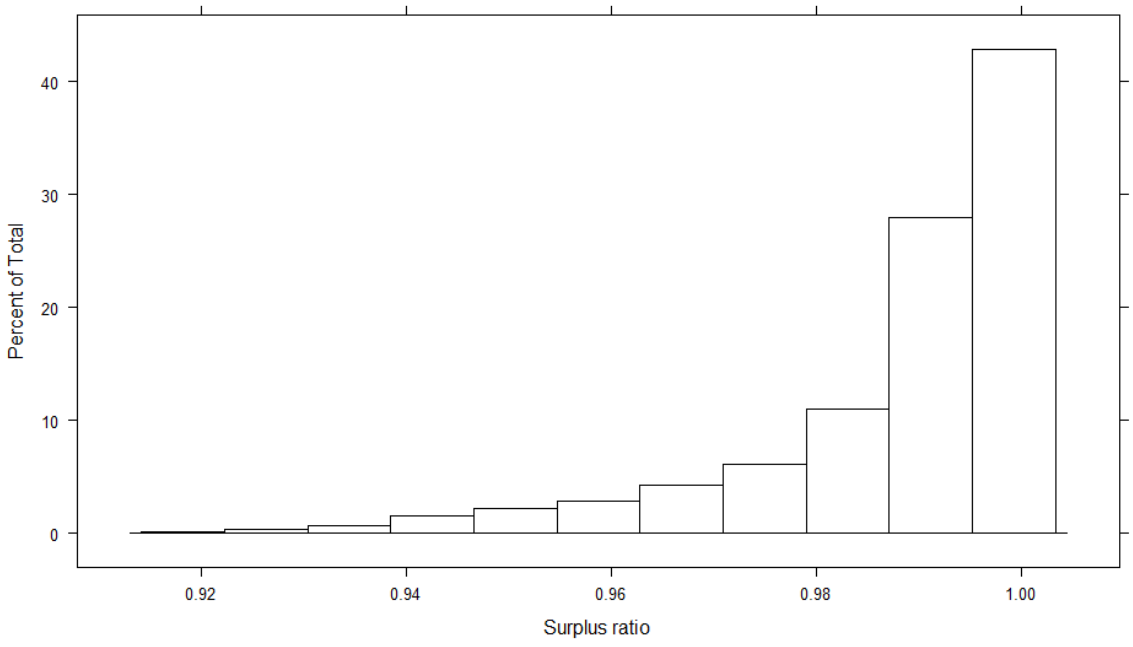

Figure 10: Histogram for the surplus ratio under perfect information policy

There are small differences in the surplus ratio across different information disclosure policies, regardless of whether the mix of participating advertisers is determined at the beginning of the auction or before each round (see Table 14).

Table 14: Surplus ratio comparison across different information disclosure policies and mix of participating advertisers

	No information disclosure policy	Partial information disclosure policy	Perfect information disclosure policy
Minimum			
Fixed mix of advertisers	0.9126	0.9056	0.9175
Varied mix of advertisers	0.9328	0.9370	0.9418
Maximum			
Fixed mix of advertisers	1.0000	1.0000	1.0000
Varied mix of advertisers	0.9997	0.9996	0.9998
Mean			
Fixed mix of advertisers	0.9845	0.9841	0.9831
Varied mix of advertisers	0.9927	0.9925	0.9932

The surplus ratios are similar across different information disclosure scenarios in which the values of advertisers are randomly drawn from Beta distributions and Gamma distributions (see Table 15).

Table 15: Surplus ratio comparisons across different information disclosure policies and for the Beta and Gamma value distributions

	No information disclosure policy	Partial information disclosure policy	Perfect information disclosure policy
Minimum			
Beta distribution	0.9126	0.9060	0.9175
Gamma distribution	0.9170	0.9056	0.9270
Maximum			
Beta distribution	1.0000	1.0000	1.0000
Gamma distribution	1.0000	1.0000	1.0000
Mean			
Beta distribution	0.9880	0.9878	0.9875
Gamma distribution	0.9891	0.9888	0.9887

There are negligible differences among the surplus ratios across different scenarios in which the statistical character of the relationships between advertisers' value distributions vary (see Table 16).

Table 16: Surplus ratio comparisons across different information disclosure policies for four different statistical relationships among advertisers' value distributions

	No information disclosure policy	Partial information disclosure policy	Perfect information disclosure policy
Minimum			
Identical and correlated	0.9328	0.9271	0.9352
Identical and independent	0.9170	0.9060	0.9270
Different and correlated	0.9126	0.9056	0.9175
Different and independent	0.9207	0.9254	0.9229

Table 16 (cont'd)			
Maximum			
Identical and correlated	1.0000	1.0000	1.0000
Identical and independent	1.0000	1.0000	1.0000
Different and correlated	1.0000	1.0000	1.0000
Different and independent	1.0000	1.0000	1.0000
Mean			
Identical and correlated	0.9883	0.9878	0.9879
Identical and independent	0.9884	0.9882	0.9883
Different and correlated	0.9884	0.9880	0.9875
Different and independent	0.9891	0.9889	0.9887

Surplus ratios are similar across the scenarios in which advertisers use Cournot learning and fictitious play as the learning algorithms (see Table 17).

Table 17: Surplus ratio comparison across different information disclosure policies for the Cournot learning and Fictitious play as learning strategies

	No information disclosure policy	Partial information disclosure policy	Perfect information disclosure policy
Minimum			
Cournot	0.936	0.9384	0.9429
Fictitious ($\gamma = 0$)	0.9568	0.9624	0.9622
Fictitious ($\gamma = 0.25$)	0.9624	0.9485	0.9399
Fictitious ($\gamma = 0.5$)	0.9635	0.9564	0.9349
Fictitious ($\gamma = 0.75$)	0.9126	0.9056	0.9175
Maximum			
Cournot	1	1	1
Fictitious ($\gamma = 0$)	1	1	1
Fictitious ($\gamma = 0.25$)	1	1	1
Fictitious ($\gamma = 0.5$)	1	1	1
Fictitious ($\gamma = 0.75$)	1	1	1
Mean			
Cournot	0.9852	0.9854	0.9867
Fictitious ($\gamma = 0$)	0.9921	0.9927	0.992
Fictitious ($\gamma = 0.25$)	0.9928	0.9916	0.9913
Fictitious ($\gamma = 0.5$)	0.9924	0.9922	0.9900
Fictitious ($\gamma = 0.75$)	0.9806	0.9796	0.9807

There is no clear pattern to the differences in surplus ratios across the different information disclosure scenarios depending on whether advertisers use a common learning algorithms or different learning algorithms except for the worst case scenarios where surplus ratios are always higher by a small but noticeable amount when a common learning algorithm is used (see Table 18).

Table 18: Surplus ratio comparison across different information disclosure policies when competing bidders use common and different learning algorithms

	No information disclosure policy	Partial information disclosure policy	Perfect information disclosure policy
Minimum			
Common algorithm	0.9513	0.9384	0.9429
Different algorithm	0.9126	0.9056	0.9175
Maximum			
Common algorithm	1.0000	1.0000	1.0000
Different algorithm	1.0000	1.0000	1.0000
Mean			
Common algorithm	0.9898	0.9895	0.9896
Different algorithm	0.9873	0.9870	0.9867

As the click decay rate increases the surplus ratios also increase. The increase with the decay rate is larger for minimum values than for mean values, and the increase is larger for mean values than for maximum values. But across the bid information disclosure policies the surplus ratios are at similar levels (see Table 19).

Table 19: Surplus ratio comparison across different information disclosure policies and common and different click decay rates.

	No information disclosure policy	Partial information disclosure policy	Perfect information disclosure policy
Minimum			
decay = 0	0.9126	0.9056	0.9175
decay = 0.25	0.9362	0.9326	0.9274
decay = 0.5	0.9603	0.9498	0.9477
decay = 0.75	0.9780	0.9683	0.9669
decay = 1	0.9892	0.9889	0.9887
Maximum			
decay = 0	0.9954	0.9967	0.9964
decay = 0.25	0.9968	0.9970	0.9975
decay = 0.5	0.9978	0.9981	0.9989
decay = 0.75	0.9987	0.9987	0.9990
decay = 1	1.0000	1.0000	1.0000
Mean			
decay = 0	0.9749	0.9748	0.9729
decay = 0.25	0.9842	0.9837	0.9846
decay = 0.5	0.9899	0.9896	0.9899
decay = 0.75	0.9951	0.9948	0.9946
decay = 1	0.9990	0.9987	0.9989

In summary, the information disclosure policy makes a difference in GSP auctions. The results show that the bid information disclosure policy can potentially serve as an auction instrument that improves search engine revenue. With the partial information disclosure policy, auctions generate revenue levels closer to their maximum attainable levels than auctions employing the other two information disclosure policies. The revenue ratios for the no information and perfect information disclosure policies are at similar levels. Additionally, for most scenarios' yields are higher for the partial information disclosure policy than for the no information and perfect information disclosure policies, which generally produce yields that are very close to each other.

Regressions using the same data underlying the results reported in this chapter provide further support for these results and, by extension, conclusions based on interpretations of these results. The results from the regression models are consistent with the comparisons of metrics in this chapter's tables. Changing the information disclosure policy has no appreciable effect on surplus ratios. But the revenue ratios and yields are higher with partial information disclosure than for the other two information disclosure policies after controlling for other factors. Furthermore, the effects on surplus ratios, revenue ratios and yields of all other parameters I used to generate the scenarios are identified using the comparisons presented in the tables. The regression results are reported in Appendix VIII.

CONCLUSION AND DISCUSSION

5.1 Conclusion and Discussion

This study investigated the effects of information disclosure polices on search engine revenue and efficiency in GSP auctions. The following research questions provided a framework for the investigation:

Research Question 1. Will providing advertisers with bids or bid statistics from prior rounds produce higher levels of search engine revenue in GSP auctions than providing no bid information?

Research Question 2. Will providing advertisers with bids or bid statistics from prior rounds facilitate more efficient allocations in GSP auctions than providing no bid information?

Three types of information disclosure policies were under investigation: no information disclosure, partial information disclosure, and perfect information disclosure. The market transparency level increased over the three types of information disclosure policies. Under the no information disclosure, a search engine does not provide bid information to advertisers. Under the partial information disclosure, the search engine announces the mean and standard deviation for the prior round's bids. Under the perfect information disclosure, a search engine provides advertisers with bids from the prior round.

In previous studies of GSP auction, researchers often employed mathematical models. In order to make the models analytically tractable, certain assumptions and equilibrium refinement concept were imposed. These assumptions and the equilibrium refinement concept often rendered the findings less valuable in practice. In order to overcome these difficulties and compensate for the limitations of mathematical models, I developed agent-based models to

simulate GSP auctions under different bid information disclosure policies. I ran the auctions for 800 different scenarios, summarized the results, and drew conclusions. The 800 scenarios were formed by varying the values of several parameters. These parameters included whether the mix of advertisers changed over rounds, click value distributions, the statistical character of relationships among bidders' click value distributions, learning algorithm, whether advertisers use a common learning algorithm or different learning algorithms, and the click decay rate for ad slots. These parameters were all deemed as important in prior research of GSP auctions (e.g., Edelman & Schwarz, 2010; Yao, Chen, & Liu, 2012; Bu et al., 2008; Paes Leme, & Tardos, 2010a).

The first research question asked: will providing advertisers with bids or bid statistics from prior rounds produce higher levels of search engine revenue in GSP auctions than providing no bid information? The answer was yes. I used revenue ratio and yield as metrics to measure search engine revenue. Revenue ratio is the revenue from a round as a percentage of its maximum attainable level and yield is the percentage of surplus captured by the search engine as its revenue. The results in chapter 4 demonstrated that different bid information disclosure policies produced different search engine revenue ratios and yields When there was bid information disclosed, that is, the market was more transparent, both revenue ratios and yields were higher than when those there was no bid information disclosed. However, the highest level of market transparency did not translate into the highest levels for the search engine revenue ratio and yield. On average, the revenue ratio for partial information disclosure was 3% higher than for no information disclosure, and was 2.6% higher than for perfect information disclosure. On average the yield under the partial information disclosure policy was 1.96% higher than that under the no information disclosure policy, and was 1.62% higher than that under the perfect

information disclosure policy. Given that search ad revenue in 2015 was 20.5 billion and search engines currently do not provide bid information to advertisers, changing the bid information disclosure policy from a no information to a partial information disclosure policy could mean that the revenue would increase by 0.615 billion dollars.

However, the finding that search engine revenue ratios were at their highest levels under partial information disclosure rather than under perfect information disclosure were somewhat counter to initial expectations. A few single item auction studies (e.g., Andreoni, Che, & Kim, 2007; Elmaghraby & Keskinocak, 2000) found that disclosing bid information increases auctioneer revenue. Therefore, the expectation was that the more bid information was disclosed the higher would be search engine revenue.

The second research question I analyzed in this dissertation involved the effects of information disclosure policies on surplus generation. My results showed that changing the information disclosure policy did not have noticeable effects on surplus generation in GSP auctions. Surplus generation was measured with the surplus ratio, which is a measure how close the surplus generated by an auction approaches its maximum attainable level. When other variables were held constant, mean surplus ratios varied little across the different information policies. This result indicates that GSP auctions generate similar amounts of surplus under different information disclosure policies holding other variables constant. Combined with the findings about revenue ratios and yields, these results suggest that providing bid statistics helped a search engine capture larger portions of surplus as revenue, at little or no cost to the efficiency of the auctions. The effect on search engine revenue of changes in information disclosure policy was reflected reallocations of part of the auctions' surplus from advertisers to the search engine.

One noticeable aspect of the findings for this dissertation is that the effects of changes in information disclosure policy on revenue ratios and yields varied with the changes in the parameters examined in this study, including whether advertisers used common or different learning algorithms, whether the mix of advertisers change over rounds, and the statistical relationship among click value distributions. The effect of disclosing bid statistics on the revenue ratio was relatively small in the following scenarios: when advertisers used different learning algorithms (The means of the revenue ratio were 0.8082, 0.8252, and 0.8106 for no information disclosure, partial information disclosure and perfect information disclosure, respectively.); when the mix of advertisers changed over rounds (The means of the revenue ratios were 0.8178, 0.8339, and 0.59, and 0.8243 for no information disclosure, partial information disclosure and perfect information disclosure, respectively.); and, when the click value distributions were identical and correlated (The means of revenue ratios were 0.8503, 0.8695, and 0.8554 for no information disclosure, partial information disclosure and perfect information disclosure, respectively.). Similarly, the effect of the information disclosure policy on yield was relatively small in the following scenarios: where advertisers used different learning algorithms (The means of yield were 0.5832, 0.5933, and 0.5853 for no information disclosure, partial information disclosure and perfect information disclosure, respectively.); when the same mix of advertisers competed against each other through all an auction's rounds (The yields were 0.5869, 0.5975, and 0.5913 for no information disclosure, partial information disclosure and perfect information disclosure, respectively.); and when the value distributions were identical and correlated (The means of yield were 0.6814, 0.6958, and 0.6852 for no information disclosure, partial information disclosure and perfect information disclosure, respectively.).

There is a long-lasting discussion in the search ad research literature concerning why the search ad industry uses the GSP mechanism rather than the VCG mechanism, since the GSP mechanism is not truthful. Findings from this dissertation contribute to this discussion. In the worst case scenario, under the no information disclosure policy, which coincides with the current GSP auctions, the surplus reached 91.24% of the optimal level. Additionally, variation of surplus ratios across the different information disclosure policies was small. These results suggest that the GSP mechanism is relatively robust in terms of surplus generation. When it comes to search engine revenue, a search engine could extract 57.63% of the surplus as its revenue on average using the GSP mechanism. On the other hand, a search engine always obtains half of the surplus as its revenue using the VCG mechanism. Thus, the findings suggest that one possible motivation for search engines to adopt the GSP mechanism is that this mechanism achieves a good trade-off between efficiency and revenue. In other words, search engines gain higher levels of revenue at little cost of efficiency by using the GSP mechanism for keyword auctions. On average a search engine can capture 52.58 % of surplus (.9124 x .5763 = .5258) as its revenue in GSP auctions. This is higher than that the 50% for VCG auctions.

This study represents a first step in exploring the effects of bid information disclosure policy on search engine revenue and surplus generation in GSP auctions. An extensive literature on auctions has analyzed the effect of information disclosure in various contexts (e.g., Goeree 2003, Das Varma 2003, Katzman and Rhodes-Krop 2008, de Silva Dunne, Kankanamge, & Kosmopoulou, 2008). To the best of my knowledge, my dissertation is the first to study information disclosure policies in GSP auctions in which there are multiple vertically differentiated items to auction off.

My dissertation is distinct from the existing literature in the field of search advertising with GPS auctions in several ways. First, it studied the effects of an auction instrument that has not been investigated in this line of research before. Moreover, the research method I used in the dissertation opened doors for incorporating more complexities in keyword auction analyses. For instance, it was possible to allow the value distributions of the advertisers to be related to each other in more complex ways than have been studied in previous research. Furthermore, my analysis offers an actionable managerial suggestion for improving search engines' yields from GSP auctions.

5.2 Limitations and Future Research Directions

There are several limitations that should be kept in mind when considering the applicability of this dissertation's findings to the real world keyword auctions. Expanding upon these limitations provides future research directions.

1. I ran the same number of simulated auctions for each scenario. The underlying assumption is that every scenario has an equal chance of occurring. However, in the real world keyword auctions different scenarios have different chances of occurring. The consequence of violating this assumption is that the effect of the bid disclosure policy might be manifest in a different way than was shown in the dissertation. For example, if it is more likely that the scenarios in which the information disclosure policy was less effective would occur in the real world keyword auction, then changing the information disclosure policy might be ineffective in actuality. Ultimately, it remains an empirical question as to what kinds of scenarios are more likely to occur in the real world.

2. Alternative learning algorithms can be explored in future research. The two learning algorithms I employed captured the learning behaviors in GSP auctions reasonably well. However, advertisers could use less or more sophisticated learning algorithms. Clearly, implementing other learning algorithms would be another way to broaden the scope of this research. There are two ways to move forward in this direction. The first would be to implement learning algorithms that already exist in the literature (e.g., Cary, et al, 2014). The second would be to use machine learning techniques to discover the learning algorithms that generate real bidding data or experimental data and then implement these algorithms. The advantage of the second approach is that the computer agents will behave more like human bidders or the agents employed by them. Thus, the predictions will be more accurate.

3. In this dissertation, I assumed that the goal of advertisers was to maximize their payoffs from clicks. However, search advertisers sometimes have different goals. For example, an advertiser might submit bids that are intended to exhaust the budget of one of its opponents. To incorporate this kind of goal is another way to move the study forward.

4. Another assumption of my study is that the advertisers are able to find the necessary funds for search advertising. In real world keyword auctions, advertisers have budget constraints. Budget constraints could influence advertisers' bid forming and auction entry decisions. Incorporating budget constraints in agent-based models could shed more light on the robustness of GSP auctions.

5. The results of this dissertation show that the surplus levels of GSP auctions are relatively high and are robust with respect to the bid information disclosure policy, whether the mix of advertisers varied over rounds, the characteristics of advertisers' click value distributions,

and changes in the click decay rates for the ad slots. But these findings all come from scenarios where advertisers do not have budget constraints. Future studies could investigate the effect of budget constraints on advertisers' bidding behaviors, search engine revenue, and the surplus of GSP auctions.

6. The advertisers in my dissertation used the same bidding strategy in the auctions and only differed in terms of their values for clicks (e.g., Edelman et al., 2007) and in their learning algorithms in half of the simulations. In reality, advertisers could be different in other respects, including bidding strategies, access to information, experience, and learning abilities. For example, inexperienced advertisers might be clueless about how a keyword auction works and might overbid their true values; these aspects can be incorporated in future research. The results could also shed more light on the robustness of GSP auctions and the findings about the effects of information disclosure policies.

7. In my agent-based models, an advertiser only has access to bid information only if it has participated in a prior auction round. In current keyword auctions, an advertiser can use auction tools to estimate which ad slot its advertisement will end up with based on its bid and how likely this and other slots will be acquired. This, in a sense, is a kind of bid information disclosure policy. It is approximately equal to providing bid estimates from all past rounds to all advertisers, regardless of their participation history. It would be interesting to see whether there are any discernable differences in effects on revenue ratios and yields between the kind of bid information disclosure policy currently employed by search engines and the partial information disclosure policy examined in my dissertation.

8. Another difference between real keyword auctions and my simulated auctions under partial information disclosure policy is that every advertiser in my simulation uses the bid information provided by the auctioneer but real advertisers do not always use the auction tool provided by search engines. One potential way to allow for such differences in future research would be to grant advertisers participating in the simulated actions different amounts of prior round bid information.

9. The models in my dissertation do not explicitly capture the behaviors of search engine users. Search engine users play a critical role in search advertising. A search engine only receives payment when a search user clicks on an advertisement. An advertiser will have a higher chance of making a sale if the search engine users notice or click on its ad. Adding parameters that can describe more complex clicking patterns would be one way to take search engine users into account. Another method of gaining insights would be to introduce search user agents in these models.

APPENDICES

APPENDIX I

DENSITY, MEAN AND RELATIONSHIPS AMONG THE 10 VARIABLES FOLLOW THE VALUE DISTRIBUTIONS OF THE ADVERTISERS

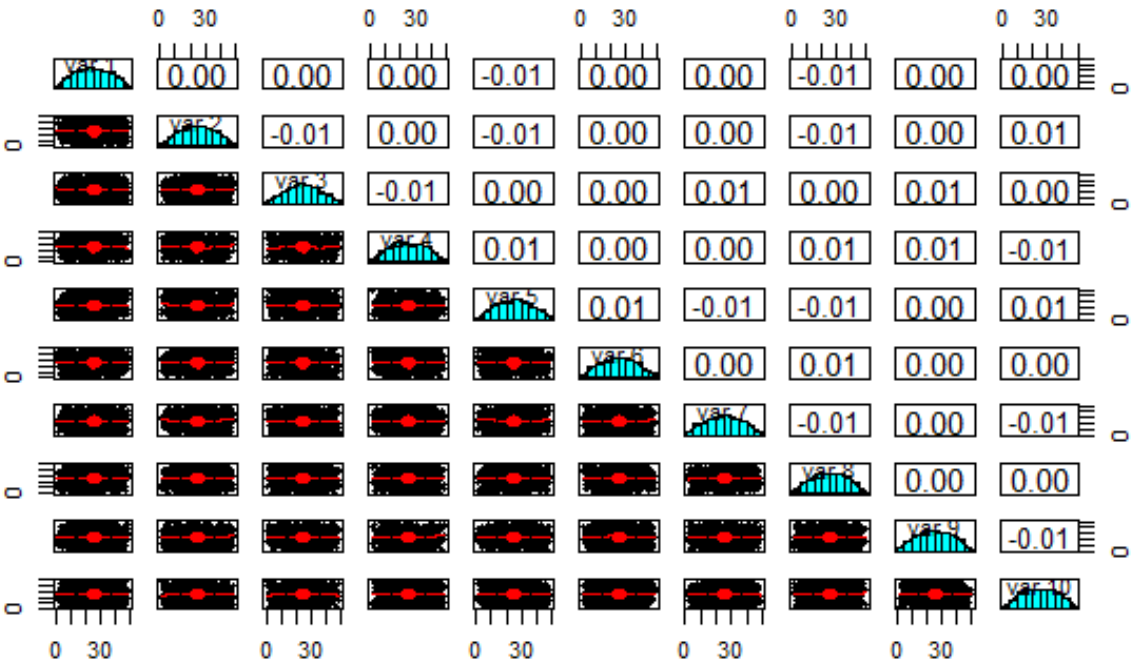

Figure 11: The density, mean, and relationships among the 10 identically and independently distributed random variables following Beta distribution. The diagonal graphs demonstrate the density of the 10 variables. The upper right parts are the correlations among the 10 variables. And the lower left parts are the means of the 10 variables. The figure is generated by using 1000 random draws from the distributions.

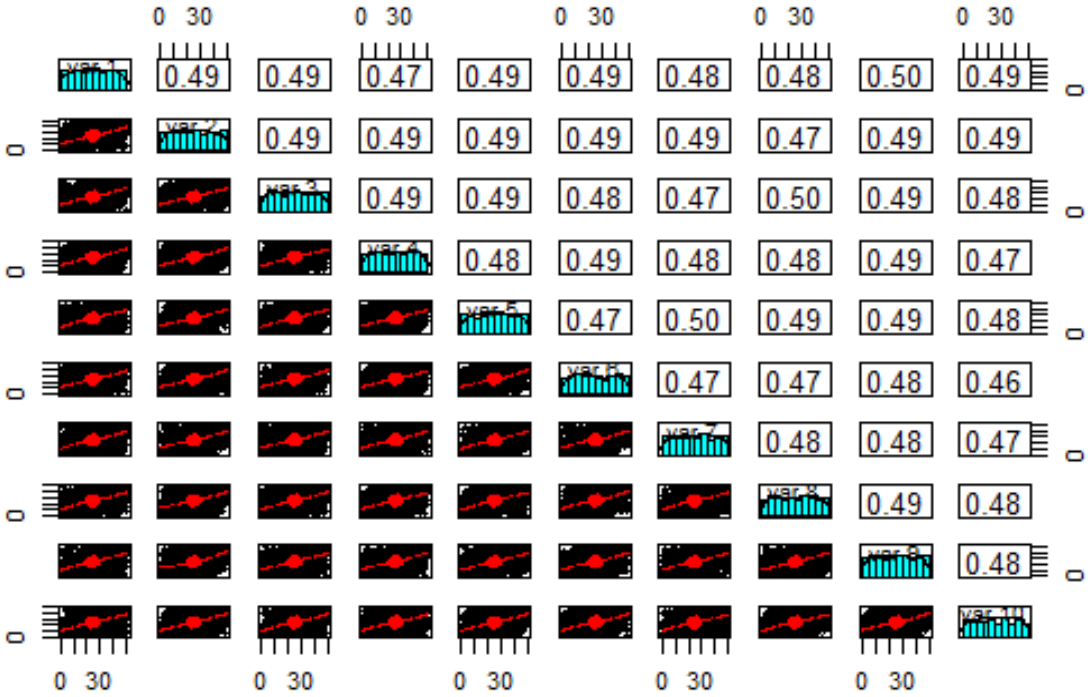

Figure 12: The density, mean, and relationships among the 10 identically and correlated distributed random variables following Beta distribution. The diagonal graphs demonstrate the density of the 10 variables. The upper right parts are the correlations among the 10 variables. And the lower left parts are the means of the 10 variables. The figure is generated by using 1000 random draws from the distributions.

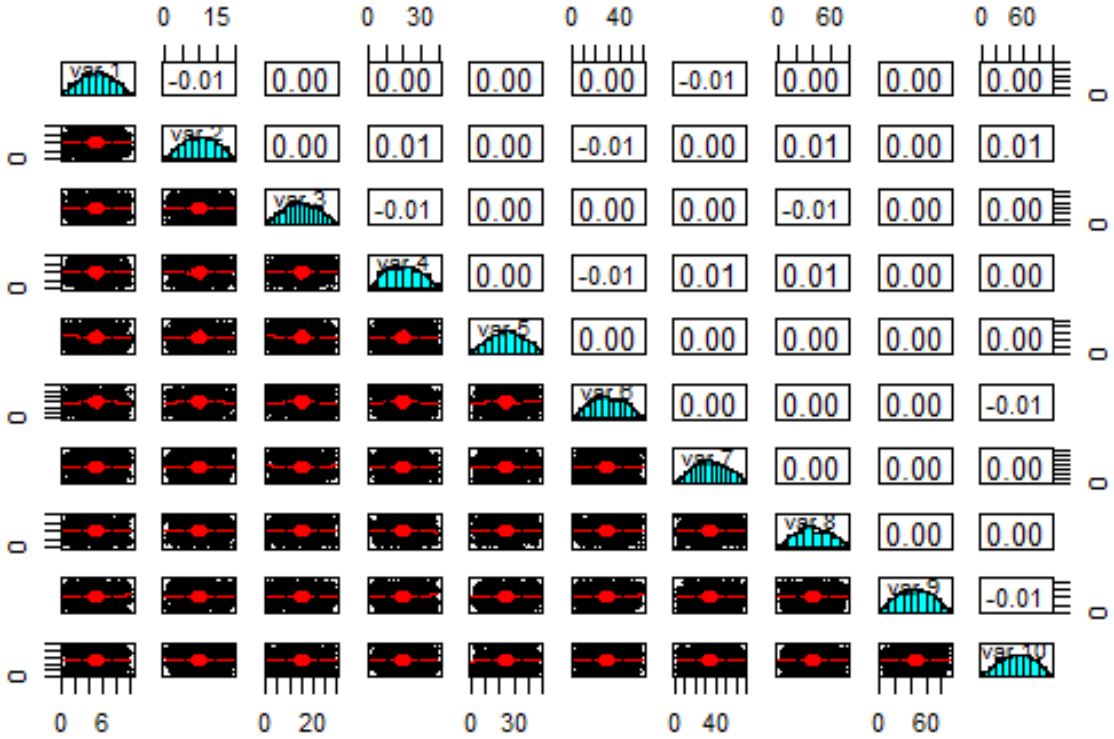

Figure 13: The density, mean, and relationships among the 10 independently distributed random variables which are different and follow Beta distribution. The diagonal graphs demonstrate the density of the 10 variables. The upper right parts are the correlations among the 10 variables. And the lower left parts are the means of the 10 variables. The figure is generated by using 1000 random draws from the distributions.

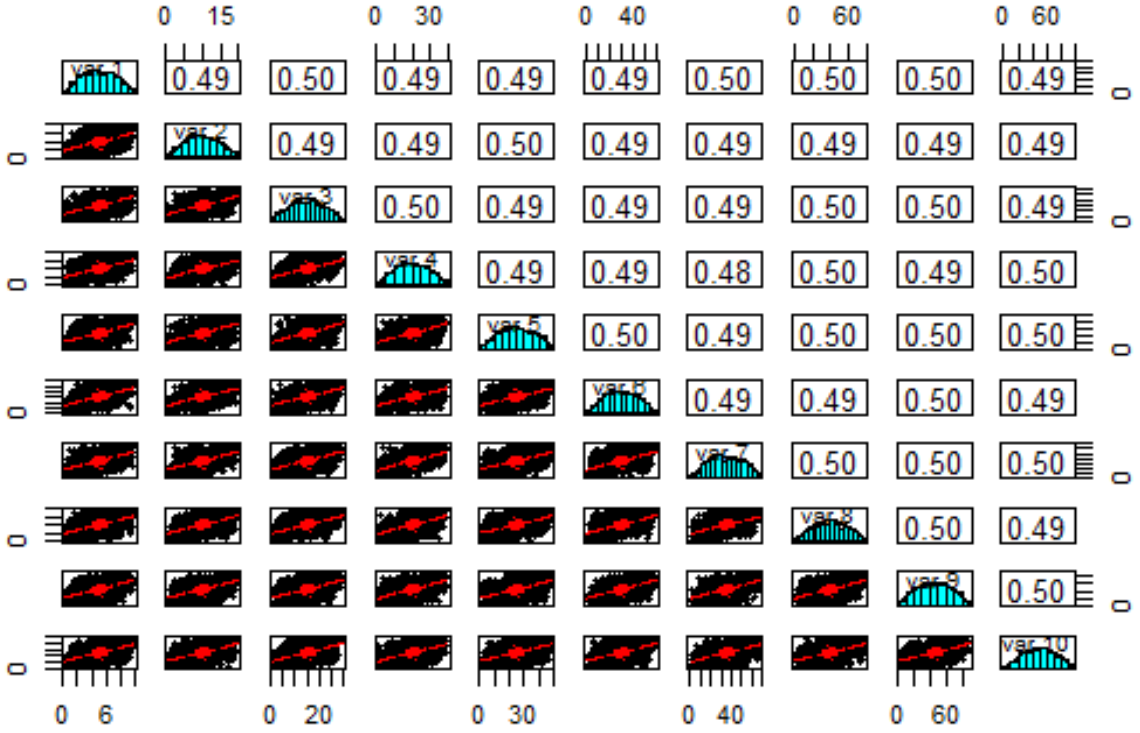

Figure 14: The density, mean, and relationships among the 10 different and correlated random variables following Beta distribution. The diagonal graphs demonstrate the density of the 10 variables. The upper right parts are the correlations among the 10 variables. And the lower left parts are the means of the 10 variables. The figure is generated by using 1000 random draws from the distributions.

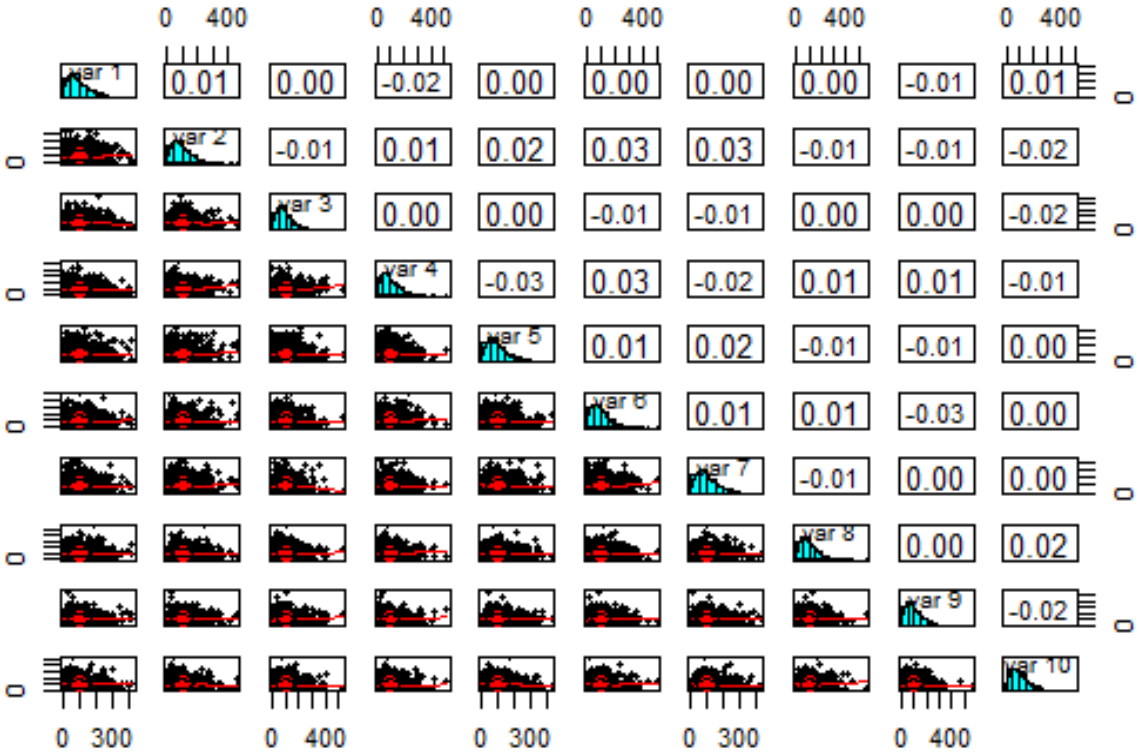

Figure 15: The density, mean, and relationships among the 10 identically and independently distributed random variables following Gamma distribution. The diagonal graphs demonstrate the density of the 10 variables. The upper right parts are the correlations among the 10 variables. And the lower left parts are the means of the 10 variables. The figure is generated by using 1000 random draws from the distributions.

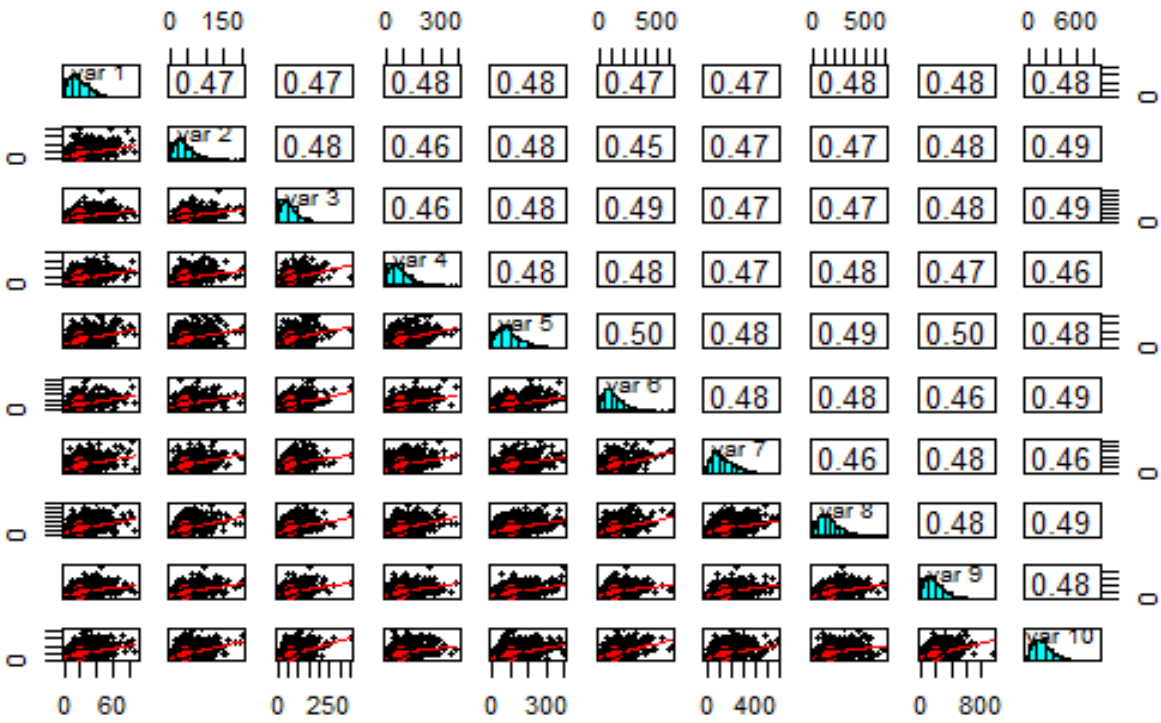

Figure 16: The density, mean, and relationships among the 10 different and correlated distributed random variables following Gamma distribution. The diagonal graphs demonstrate the density of the 10 variables. The upper right parts are the correlations among the 10 variables. And the lower left parts are the means of the 10 variables. The figure is generated by using 1000 random draws from the distributions.

Figure 17: The density, mean, and relationships among the 10 independently distributed random variables which are different and follow Gamma distribution. The diagonal graphs demonstrate the density of the 10 variables. The upper right parts are the correlations among the 10 variables. And the lower left parts are the means of the 10 variables. The figure is generated by using 1000 random draws from the distributions.

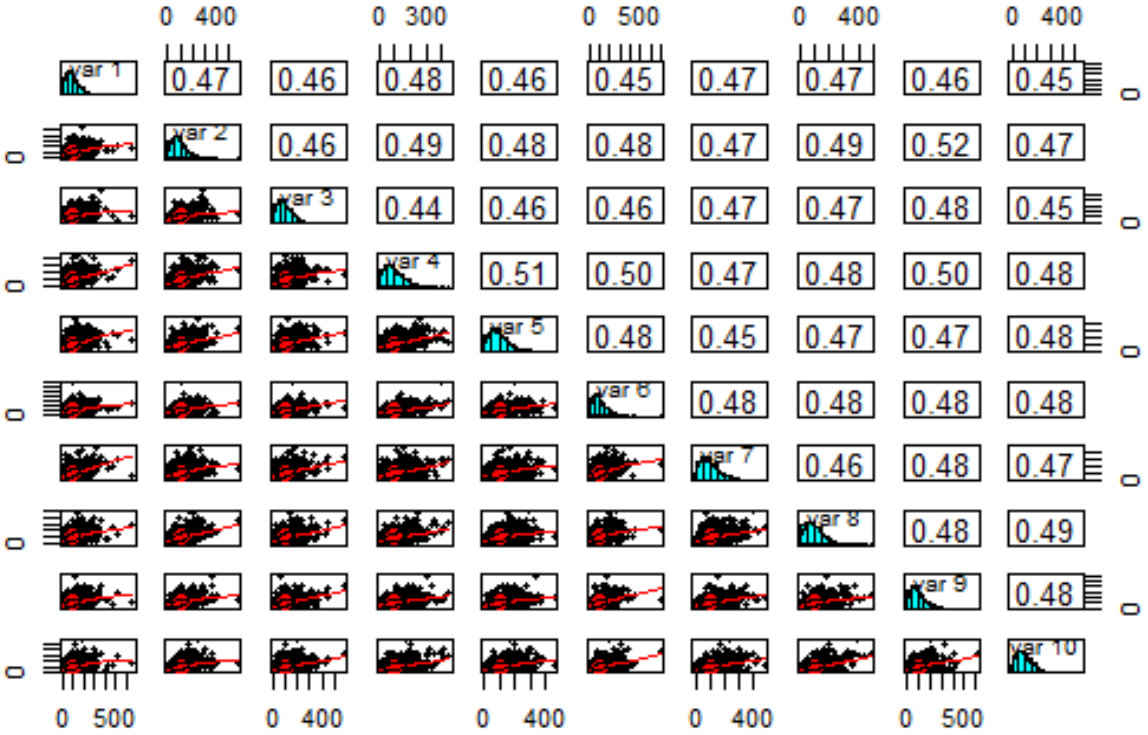

Figure 18: The density, mean, and relationships among the 10 identical and correlated random variables following Gamma distribution. The diagonal graphs demonstrate the density of the 10 variables. The upper right parts are the correlations among the 10 variables. And the lower left parts are the means of the 10 variables. The figure is generated by using 1000 random draws from the distributions.

APPENDIX II

TESTS FOR AUCTION CONVERGENCE OVER THE ROUNDS

Table 20: t-test table between adjacent rounds.

Rounds		t-value	p-value
Round 1	Round 2	-0.8695	0.3846
Round 2	Round 3	91.052	0.0000
Round 3	Round 4	34.533	0.0000
Round 4	Round 5	31.843	0.0000
Round 5	Round 6	28.790	0.0000
Round 6	Round 7	22.028	0.0000
Round 7	Round 8	18.320	0.0000
Round 8	Round 9	14.714	0.0000
Round 9	Round10	9.5448	0.0000
Round10	Round 11	7.7473	0.0000
Round 11	Round 12	5.2532	0.0000
Round 12	Round 13	3.8481	0.0001
Round 13	Round 14	2.3992	0.0164
Round 14	Round 15	2.2604	0.0238
Round 15	Round 16	1.0295	0.3032
Round 16	Round 17	1.4513	0.1467
Round 17	Round 18	0.4966	0.6195
Round 18	Round 19	1.0021	0.3163
Round 19	Round 20	0.8734	0.3825
Round 20	Round 21	0.1069	0.9149
Round 21	Round 22	0.8820	0.3778
Round 22	Round 23	0.0184	0.9854
Round 23	Round 24	0.8068	0.4198
Round 24	Round 25	-0.2799	0.7796
Round 25	Round 26	0.3454	0.7298
Round 26	Round 27	0.4892	0.6247
Round 27	Round 28	0.3574	0.7208
Round 28	Round 29	0.2999	0.7642
Round 29	Round 30	0.0339	0.9729

Table 21: Mann-Whitney test table between adjacent rounds.

Rounds		W	p-value
Round 1	Round 2	28767000000	0.4876
Round 2	Round 3	35196000000	0.0000
Round 3	Round 4	30975000000	0.0000
Round 4	Round 5	30980000000	0.0000
Round 5	Round 6	30484000000	0.0000
Round 6	Round 7	30032000000	0.0000
Round 7	Round 8	29725000000	0.0000
Round 8	Round 9	29495000000	0.0000
Round 9	Round10	29279000000	0.0000
Round10	Round 11	29173000000	0.0000
Round 11	Round 12	28800000000	0.0988
Round 12	Round 13	28760000000	0.0955
Round 13	Round 14	28660000000	0.1359
Round 14	Round 15	28500101100	0.1503
Round 15	Round 16	28490000000	0.1624
Round 16	Round 17	28351000000	0.2300
Round 17	Round 18	28236000000	0.2531
Round 18	Round 19	28139000000	0.2621
Round 19	Round 20	28034000000	0.2600
Round 20	Round 21	27965000000	0.2700
Round 21	Round 22	26360000000	0.2869
Round 22	Round 23	25136910000	0.3000
Round 23	Round 24	25003000000	0.3005
Round 24	Round 25	25000090000	0.3169
Round 25	Round 26	25000000900	0.3189
Round 26	Round 27	24995000000	0.3199
Round 27	Round 28	24800690000	0.3270
Round 28	Round 29	24726000000	0.5851
Round 29	Round 30	24622200000	0.6696

In order to assess the convergence of the auctions I ran t test and Mann-Whitney test to assess at which round the auctions started to converge to steady trend. I compared the surplus ratio between adjacent rounds for this purpose. If there is no significant difference between two adjacent round it means that the competition dynamics stayed the same across rounds.

The findings from the analyses was that the auctions converged to stable status around round 15. Therefore, the auctions should run at least 15 rounds.

APPENDIX III

TESTS FOR AUCTION CONVERGENCE OVER THE ROUNDS GIVEN DIFFERENT SIZES OF ADVERTISER POOL

Table 22: Comparisons of convergence rounds given different sizes of advertiser pools.

	t-test			Mann-Whitney test		
	Converged rounds	t	p	Converged rounds	W	p
9 advertisers in pool	Round 12 to Round 13	2.0122	0.0689	Round 10 to Round 11	28135000000	0.1023
10 advertiser in pool	Round 15 to Round 16	1.0295	0.3032	Round 11 to Round 12	28800000000	0.0988
11 advertiser in pool	Round 16 to Round 17	1.9982	0.06983	Round 13 to Round 14	28309000000	0.0931

The purpose of this sensitivity test is to find whether the size of advertiser pool will influence how fast the auctions will converge. I ran the simulations with 9, and 11 advertisers in the pool and compared the results from t-test and Mann-Whitney tests. The results indicated that the size of the advertiser pool had effect on how fast the auctions can converge to steady status. But the size of the advertiser pool did not influence the competition dynamics in the auctions.

APPENDIX IV

SENSITIVITY TEST FOR NECESSARY NUMBER OF AUCTIONS REQUIRED TO REACH EQUILIBRIUM IN EACH SCENARIO

Table 23: Sensitivity tests for numbers of auctions in each scenario.

Numbers of Auctions	Revenue ratio		Yield		Surplus ratio	
	Minimum	Maximum	Minimum	Maximum	Minimum	Maximum
1500	0.6132	0.9677	0.2803	0.8646	0.9068	1
1200	0.6269	0.9633	0.2833	0.8728	0.8992	1
900	0.6226	0.9526	0.2789	0.8607	0.9007	1
600	0.6281	0.9541	0.2812	0.8448	0.9058	1

The purpose of this sensitivity test was to determine the necessary amount of auctions to run in each scenario in order to cover the parameter spaces. I tested 1500, 1200, 900, and 600 auctions. As the results in Table 23 indicated when there were negligible differences when the numbers of auctions decreased from 1500 to 600. Therefore, I decided upon 600 auctions in each scenario.

APPENDIX V

COMPARISONS OF DIFFERENT METHODS FOR INITIATING THE AUCTIONS

Table 24: Changing over round with different initial bids.

	Submitting Values			Submitting Reserve Price			Submitting Bids from last auctions		
Round	Surplus ratio	Revenue ratio	Yield	Surplus ratio	Revenue ratio	Yield	Surplus ratio	Revenue ratio	Yield
1	0.5155	0.9984	0.9994	0.0050	0.0187	0.0070	0.8981	0.8478	0.7256
2	0.5704	1.0000	1.0000	0.0050	0.0129	0.0079	0.9187	0.8313	0.6897
3	0.9975	0.7332	0.5535	0.9323	0.6100	0.4524	0.8571	0.6164	0.5105
4	0.9960	0.7836	0.5398	0.9802	0.6939	0.5070	0.9240	0.7621	0.6169
5	0.9766	0.7763	0.5907	0.9993	0.7394	0.5190	0.9084	0.7447	0.6371
6	1.0000	0.7750	0.5557	0.9554	0.7091	0.5363	0.9022	0.7600	0.6524
7	1.0000	0.7628	0.5387	0.9906	0.7308	0.5136	0.9355	0.8012	0.6347
8	1.0000	0.7860	0.5843	0.9914	0.7414	0.5196	0.9271	0.7971	0.6911
9	0.9976	0.7982	0.5936	0.9993	0.7523	0.5040	0.9355	0.8145	0.6671
10	1.0000	0.7847	0.5782	1.0000	0.7612	0.5306	0.9105	0.8162	0.6892
11	1.0000	0.7900	0.5814	0.9940	0.7750	0.5760	0.9276	0.8095	0.6868
12	1.0000	0.7926	0.5610	1.0000	0.7655	0.5542	0.9160	0.7940	0.6719
13	0.9990	0.7829	0.5623	0.9959	0.7663	0.5537	0.9351	0.7872	0.6534
14	0.9989	0.8049	0.5496	1.0000	0.7805	0.5496	0.9351	0.8657	0.7048
15	1.0000	0.7989	0.5936	1.0000	0.7884	0.5386	0.8829	0.8640	0.7325
16	1.0000	0.8059	0.5999	1.0000	0.7867	0.5541	0.8829	0.8563	0.7225
17	0.9991	0.7943	0.5507	1.0000	0.7782	0.5578	0.9188	0.8356	0.7140
18	1.0000	0.8015	0.5915	0.9920	0.7743	0.5465	0.9283	0.8436	0.7116
19	1.0000	0.8024	0.6105	0.9921	0.7767	0.5619	0.9364	0.8593	0.7068
20	1.0000	0.7830	0.5598	1.0000	0.7797	0.5426	0.8826	0.8652	0.7480
21	1.0000	0.8026	0.5945	1.0000	0.7835	0.5604	0.8709	0.8504	0.7175
22	1.0000	0.7940	0.5691	0.9913	0.7829	0.5701	0.9278	0.8352	0.6798
23	0.9992	0.7953	0.5701	0.9957	0.7883	0.5756	0.9281	0.8362	0.6898
24	0.9992	0.7886	0.5949	0.9921	0.7748	0.5588	0.9188	0.8266	0.6997
25	0.9989	0.7944	0.5634	1.0000	0.7822	0.5308	0.8708	0.8221	0.7383
26	1.0000	0.8105	0.6141	1.0000	0.7885	0.5453	0.9368	0.8264	0.6661
27	1.0000	0.7986	0.5689	1.0000	0.7735	0.5463	0.9283	0.8192	0.6670
28	1.0000	0.7926	0.5764	1.0000	0.7842	0.5796	0.9262	0.8197	0.6954
29	1.0000	0.7950	0.5962	0.9959	0.7730	0.5558	0.8981	0.8478	0.7256
30	1.0000	0.7929	0.5600	1.0000	0.7889	0.5836	0.9187	0.8313	0.6897

Table 25: Comparisons with different initial bids and under different information disclosure policy.

	Submitting Values	Submitting Reserve Price	Submitting Bids from last auction
		Surplus ratio	
No information	0.9414	0.9495	0.9576
Partial information	0.9424	0.9848	1.0000
Perfect information	0.9562	0.9921	0.9315
		Revenue ratio	
No information	0.7466	0.6806	0.7978
Partial information	0.8535	0.7165	0.8207
Perfect information	0.8157	0.7311	0.8102
		Yield	
No information	0.5873	0.5010	0.7617
Partial information	0.6240	0.5159	0.7734
Perfect information	0.6267	0.5052	0.7766

There are multiple ways to initiate auctions: advertisers submitting their true values, advertiser submitting the reserve price and advertisers submit bids from last auction. In order to maintain the comparability between the sensitivity test and the models in the dissertation, I command the advertisers to initiate the auctions by submitting the same bids for first two rounds.

When the advertisers submitted their true values as initial bids, the search engine's revenue ratios and yields were almost maximized or maximized for the first two rounds. Later, the auctions converged to an efficient or near efficient state. The surplus ratios were 1 or close to 1. When advertisers submitted their reserve prices as the bids, the search engine's revenue ratios and yields were very low for the first two rounds. In later rounds, the auctions also converged to an efficient state or near efficient with surplus ratios that were close to 1 or equaled 1. When advertisers submitted bids from previous auctions that had converged to equilibrium, the patterns from the last auctions carried over. The revenue ratios, yields, and surplus ratios were stable

across all rounds and these auctions also converged to efficient or near efficient equilibria with the surplus ratios that were close to 1 or equaled 1 (see Table 24).

I also compared the averages of surplus ratios, revenue ratios and yields of the last 10 round of the auctions under different information disclosure policies. The general patterns were also present in these auctions for sensitivity test. (see Table 25)

These results showed that the different ways of initiating the auctions in the sensitivity tests were special cases of the methods I used in my dissertation.

APPENDIX VI

PROGRAM TESTING PROCEDURE

I used the procedure to test the program. Therefore, the programs will generate reliable and valid outputs.

The way that the programs work is to take 10 numbers larger than 0.1 as inputs, and then generate numbers as bids following the rules that are described in chapter 3. The output of these programs are bids and the best responses of the bids of ten advertiser agents, search engine revenue, and surplus of the auctions.

To ensure that the programs worked properly and followed expectations, I followed two major steps.

1. I tested every function in the program. There were two types of inputs I used to test the functions. First, I used inputs in correct format such as positive numbers and “NA”. Among the inputs, there were some extreme cases such as the values at the boundaries of ranges. Second, I used wrong inputs, such as negative numbers and special symbols. For the first type of inputs, I compared the results with hand calculation results of two persons. For the second type of input, I mainly checked whether the functions would detect the existence of wrong inputs and report errors. I corrected the mistakes in the functions if there were any.

2. I tested the entire program. I also used the two types of inputs: the correct inputs and the wrong inputs. Similar to what I did for the functions, I also tested some extreme but correct inputs such as ten 0.1s and ten “NA”s.

1) To ensure the program work properly under different settings I repeated the tests under different values of click decay rates and different numbers of rounds. And then I modified the program if I found any mistake.

2) I generated several groups of wrong inputs, such as including letters or negative numbers. As expected, the program stopped and generated an error message. This step was repeated for different values of the click decay rate and numbers of rounds.

APPENDIX VII

THE PROGRAMS

The programs are stored and shared through Github. The link for the repository is:

https://github.com/Wenjuanma/GSP-auction

APPENDIX VIII

THE REGRESSION MODELS TO ESTIMATE THE EFECTS OF BID INFORMATION DISCLOSURE POLICY AND OTHER AUCTION PARAMETERS

Table 26: Regressions to estimate effects of bid information disclosure policies, and the auction parameters

	Surplus Ratio		Revenue Ratio		Yield	
Variables	Est.	s.e.	Est.	s.e.	Est.	s.e.
Intercept	0.9693	0.0002	0.8010	0.0004	0.7124	0.0007
Round	0.0000	0.0000	0.0004	0.0000	0.0002	0.0000
Partial information	-0.0003	0.0001	0.0306	0.0003	0.0196	0.0004
Perfect information	-0.0004	0.0001	0.0049	0.0003	0.0034	0.0004
GAMMA distribution	0.0012	0.0001	-0.0151	0.0002	-0.0725	0.0003
Identical and independent	-0.0001	0.0001	-0.0327	0.0003	-0.0695	0.0005
Different and Correlated	0.0001	0.0001	-0.0637	0.0003	-0.1455	0.0005
Different and independent	0.0010	0.0001	-0.0788	0.0003	-0.1908	0.0005
Varying mix of advertisers	0.0075	0.0001	0.0253	0.0002	0.0159	0.0003
Fictitious (s = 0)	0.0047	0.0001	0.0292	0.0003	0.0142	0.0005
Fictitious (s = 0.25)	0.0100	0.0001	0.0493	0.0003	0.0227	0.0005
Fictitious (s = 0.5)	0.0098	0.0001	0.0441	0.0003	0.0199	0.0005
Fictitious (s = 0.75)	0.0094	0.0001	0.0466	0.0003	0.0216	0.0005
Click decay	0.0194	0.0001	-0.0059	0.0003	-0.0501	0.0005
Common algorithm	-0.0021	0.0001	-0.0038	0.0002	-0.0067	0.0003

These regression models use the auction data that was generated by the simulations for the dissertation. The results show that changing the information disclosure policy has no appreciable effect on surplus ratios. But the revenue ratios and yields are higher under partial information disclosure than those for the other two information disclosures policies after controlling for other auction parameters. These relationships are consistent with those inferred from the comparisons of maximum, minimum and mean values for surplus ratios, revenue ratios and yields for the three information disclosure policies presented in tables in Chapter 4.

The effects of all other parameters I used to generate scenarios on surplus ratios, revenue ratios and yields are aligned with the corresponding relationships reported in the tables in Chapter 4. The other auction parameters include the click value distribution, statistical relationships among the click value distributions, whether the mix of advertisers is fixed or

varied, the bid learning algorithm, the click decay rate, and whether advertiser are using the same learning algorithm or different ones. The round number for an auction was also included as a covariate in the regression models to control for the effects of previous rounds on decisions made in later rounds if there are any.

Compared to the click value distributions generated by the Beta distribution, the click value distributions generated by a Gamma distribution generate slightly higher surplus ratios and lower revenue ratios and yields. The surplus ratio is not sensitive to the statistical relationship among advertisers' click value distributions. But when the click value distributions are identical and correlated, the means for revenue ratios and yields are higher than those when advertisers' click value distributions have other relationships with each other, including identical and independent, different and independent, and different and correlated. When the mix of the advertisers participating in an auction varies across the rounds, the surplus ratio, revenue ratio and yield are all higher than when the mix of advertisers does not vary. When the advertisers use fictitious play as their learning algorithm, auctions generate higher values for surplus ratios, revenue ratios and yields regardless of the size of the squashing factor. Click decay rate is positively associated with Surplus ratios increase with the size of the click decay rate, but the effect on revenue ratios and yields is negative. When all the advertisers use a common learning algorithm, auctions generate lower surplus ratios, revenue ratios and yields than when advertisers employ different learning algorithms.

BIBILIOGRAPHY

BIBILIOGRAPHY

Aggarwal, G., Goel, A., & Motwani, R. (2006). Truthful auctions for pricing search keywords. In EC '06 Proceedings of the 7th ACM conference on Electronic commerce, 1–7.

Andreoni, J., Che, Y.-K., & Kim, J. (2007). Asymmetric information about rivals' types in standard auctions: an experiment. *Games and Economic Behavior, 59(2),* 240–259.

Aschenfelter, O. (1989). How auctions work for wine and art. *Journal of Economic Perspectives, 3,* 23–36.

Athey, S., & Ellison, G. (2011). Position auctions with consumer search. *The Quarterly Journal of Economics, 126(3),* 1213–1270.

Athey, S., & Nekipelov, D. (2011). A structural model of sponsored search advertising auctions. Unpublished paper. Retrieved from: http://eml.berkeley.edu/~nekipelov/pdf_papers/paper16.pdf

Barlow, R., & Marshall, R. (1964). Bounds for distributions with monotone hazard rate. *Annals of Mathematicals Statistics 35(3),* 1234–1257.

Bergemann, D., & Horner, J. (2010). Should auctions be transparent? Cowles Foundation Discussion Paper No. 1764. Retreived from: https://ssrn.com/abstract=1657947.

Börgers, T., Cox, I., Pesenorfer, M., & Petricek, V. (2013). Equilibrium bids in sponsored search auctions: theory and evidence. *American Economic Journal: Microeconomics, 5(4),* 163–187.

Brown, G. W. (1951). Iterative solutions of games by fictitious play. In T. C. Koopmans. Activity Analysis of Production and allocation. New York: Wiley. Retrieved from: http://cowles.econ.yale.edu/P/cm/m13/m13-24.pdf

Bu, T., Deng, X., & Qi, Q. (2008). Forward looking Nash equilibrium for keyword auction. Information *Processing Letters, 105,* 41–46.

Caragiannis, I., Kaklamanis, C., Kanellopoulous, P., & Kyropoulou, M. (2012). Revenue guarantees in sponsored search auctions. In Algorithms – ESA 2012, Lecture Notes in Computer Science, 7501, 253–264.

Caragiannis, I., Kaklamanis, C., Kanellopoulous, P., Kyropoulou, M., Lucier, B., Paes Leme, R., & Tardos. E. (2011). On the efficiency of equilibria in generalized second price auctions. In EC '11 Proceedings of the 12th ACM conference on Electronic commerce, 81–90.

Cary, M., Das, A., Edelman, B., Giotis, I., Heimerl, K., Karlin, A. R., Kominers, S. D., Mathieu, C., & Schwarz, M. (2014). Convergence of position auctions under myopic best-response dynamics. In the proceeding of ACM Transactions on Economics and Computation archive, 2(3), Article No. 9.

Cary, M., Das, A., Edelman, B., Giotis, I., Heimerl, K., Karlin, A. R., Mathieu, C., Schwarz, M. (2007). Greedy bidding strategies for keyword auctions. In *EC '07 Proceedings of the 8th ACM conference on Electronic commerce*, 262–271.

Cason, T. N., Kannan, K. N., & Siebert, R. (2011). An experimental study of information revelation policies in sequential auctions. *Management Science 57(4),* 667–688.

Clarke, E. H. (1971). Multipart pricing of public goods. *Public Choice, 11(1),* 17–33.

Damaceanu, R. C. (2010). Applied computational mathematics in social sciences. Sharjah: Bentham Science Publishers.

Das Varma, G. (2003). Bidding for a process innovation under alternative modes of competition. *International Journal of Industrial Organization. 21(1),* 15–37.

de Silva, D., T. Dunne, A. Kankanamge, G. Kosmopoulou. (2008). The impact of public information on bidding in highway procurement auctions. *European Economic Review, 52(1),* 150–181.

Deltas, D, & Jeitschko, T.D. (2007). Auction hosting site pricing and market equilibrium with endogenous bidder and seller participation. *International Journal of Industrial Organization, 25(2007),* 1190–1212.

Dent, A. (2013). What Google's new layout means for search marketers. Search engine land. Retrieved from: http://searchengineland.com/googles-new-layout-means-ppc-242912

Duong, Q., & Lahaie, S. (2011). *Discrete choice models of bidder behavior in sponsored search.* In the procee Jeitschko of 7th International Workshop.

Edelman, B., & Ostrovsky, M. (2007). Strategic bidder behavior in sponsored search auctions. *Decision Support Systems, 43,* 192–198.

Edelman, B., Ostrovsky, M., & Schwarz, M. (2007). Internet advertising and the generalized second-price auction: selling billions of dollars worth of keywords. *The American Economic Review, 97(1),* 242–259.

Edelman, B., & Schwarz, M. (2006). Optimal auction design in a multi-unit environment: the case of sponsored search auctions. In the proceeding of ACM Conference on Electronic Commerce.

Edelman, B., & Schwarz, M. (2010). Optimal auction design and equilibrium selection in sponsored search auctions. *American Economic Review: Papers & Proceedings, 100,* 597–602.

Elmaghraby, W., & Keskinocak, P. (2000). *Technology for transportation bidding at the Home Depot. Teaching case,* http://www2.isye.gatech.edu/people/faculty/Pinar_Keskinocak/home-depot-case.zip.

Feng, J., Bhargava, H. K., & Pennock, D. M. (2007). Implementing sponsored search in web search engine: computational evaluation of alternative mechanisms. *Informs Journal on Computing, 19(1),* 137–148.

Goeree, J. K. (2003). Bidding for the future: signaling in auctions with an aftermarket. *Journal of Economic Theory, 108(2),* 345–364.

Giotis, I., & Karlin, A. R. (2008). On the equilibria and efficiency of the GSP mechanism in keyword auctions with externalities. *In Internet and Network Economics Lecture Notes in Computer Science, 5385,* 629–638.

Goldman, M., & Rao, J. M. (2014). Experiments as instruments: heterogeneous position effects in sponsored search auctions. Working paper. Retrieved from: http://research.chicagobooth.edu/~/media/70F205DE79214CA396DBA2ECC2304B94.pdf

Golrezaei, N., & Nazerzadeh, H. (2013). Auctions with dynamic costly information acquisition. Retrieved from: http://ssrn.com/abstract=2444429

Gomes, R., & Sweeney, K. (2014). Bayes–Nash equilibria of the generalized second-price auction. *Games and Economic Behavior, 86,* 421–437.

Groves, T. (1973). Incentives in teams. *Econometrica, 41(4),* 617–631.

Gummadi, R., Key, P. B., & Proutiere, A. (2013). Repeated auctions under budget constraints: Optimal bidding strategies and equilibria. Working paper. Retrieved from: http://papers.ssrn.com/sol3/papers.cfm?abstract_id=2066175

Hailu, A., & Schilizzi, S. (2005). Learning in a “basket of crabs”: an agent-based computational model of repeated conservation auctions.

Hailu, A., Rolfe, J., Windle, J., & Greiner, R. (2011). Auction design and performance: an agent-based simulation with endogenous participation. In Agents and Artificial Intelligence Communications in Computer and Information Science, 129, 214–226.

Hart, S., & Colell, A. M. (2000). A simple adaptive procedure leading to correlated equilibrium. *Econometrica, 68(5),* 1127–1150.

Hashimoto, T. (2013). Equilibrium selection, inefficiency, and instability in Internet advertising auctions. Working paper. Retrieved from: http://ssrn.com/abstract=1483531

Interactive Advertising Bureau (2016). IAB internet advertising revenue report, 2015 full year results. Retrieved from: http://www.iab.net/media/file/IAB_Internet_Advertising_Revenue_Report_FY_2015.pdf

Katzman, B., & Rhodes-Krop, M. (2008). The consequences of information revealed in auctions. *Applied Economics Research Bulletin 2(1),* 53–87.

Kempe, D., & Mahdian, M. (2008). A cascade model for externalities in sponsored search. International Workshop on Internet and Network Economics, WINE 2008: Internet and Network Economics, 585–596.

Lahaie, S. (2006). An analysis of alternative slot auction designs for sponsored search. In Proceeding. 7th ACM Conference on Electronic Commerce, 218–227, Ann Arbor, MI, June 2006.

Lahaie, S., & Pennock, D. M. (2007). Revenue analysis of a family of ranking rules for keyword auctions. In Proceeding of the 8th ACM conference on Electronic Commerce, 50–56.

Li, L., Zeng, D., & Zhao, H. (2012). Pure-strategy Nash equilibria of GSP keyword auction. *Journal of the Association for Information Systems, 13(2),* 57–87.

Liang, L., & Qi, Q. (2007). *Cooperative or vindictive: bidding strategies in sponsored search auction.* In *Internet and Network Economics.* Springer, 167–178.

Littlestone, N., & Warmuth, M. K. (1994). The weighted majority algorithm. *Information and Computation, 108(2),* 212–261.

Liu, D., Chen, J., & Whinston, A. B. (2010). Ex ante information and the design of keyword auctions. *Information Systems Research, 21(1),* 133–153.

Lucier, B., & Paes Leme, R. (2011). GSP auctions with correlated types. In *EC '11 Proceedings of the 12th ACM conference on Electronic commerce,* 71–80.

Lucier, B., Paes Leme, R., & Tardos, E. (2012). On revenue in the generalized second price auction. In WWW '12 Proceedings of the 21st international conference on World Wide Web, 361–370.

Maillé, P., Markakis, E., Maurizio, N., Stamoulis, G. D., Tuffin, B. (2012). Sponsored search auctions: an overview of research with emphasis on game theoretic aspects. *Electronic Commerce Research, 12(3),* 265–300.

Myerson, R. (1981). Optimal auction design. *Mathematics of operations Research, 6(1),* 58–73.

Nabout, N. A. (2015). A novel approach for bidding on keywords in newly set-up search

Noti, G., Nisan, N., & Yaniv, I. (2014). An experimental evaluation of bidders' behavior in ad auction. In Proceeding of International World Wide Web Conference Committee (IW3C2).

Ostrovsky, M., & Schwarz, M. (2011). Reserve prices in internet advertising auctions: a field experiment. In EC ’11: Proceedings of the 12th ACM Conference on Electronic Commerce, 59–60.

Paes Leme, R., & Tardos, E. (2010a). Pure and Bayes-Nash price of anarchy for generalized second price auction. In *the proceeding of 51st Annual IEEE Symposium on Foundations of Computer Science (FOCS).*

Paes Leme, R., & Tardos, E. (2010b). *Bayes-Nash price of anarchy for GSP*. In AdAuctions Workshop 2010. Retrieved from: http://www.renatoppl.com/papers/paesleme_tardos_aaw10.pdf

Pin, F., & Key, P. (2011). Stochastic variability in sponsored search auctions: Observations and models. In Proceedings of the 12th ACM Conference on Electronic Commerce, 61–70.

Receptional (2012). 1st position in Google receives 53% of clicks. Retrieved from: http://www.receptional.com/blog/1st-position-in-google-receives-53-of-clicks

Rothkopf, M.H., & Harstad, R.M. (1994) Modeling competitive bidding: a critical essay. *Management Science, 40(3),* 364–84.

Skiera, B., Eckert, J., & Hinz, O. (2009). An analysis of the importance of the long tail in search engine marketing. *Electronic Commerce Research and Applications, 9,* 488–494.

Sodomka, E., Lahaie, S., & Hillard, D. (2013) A predictive model for advertiser value-per-click in sponsored search. International World Wide Web Conference Committee (IW3C2). May 13–17, 2013, Rio de Janeiro, Brazil.

Thompson, D. R. M., & Leyton-Brown, K. (2009). Computational analysis of perfect-information position auctions. In the *Proceedings of the 10th ACM conference on Electronic commerce,* 51–60.

Thompson, D. R. M., & Leyton-Brown, K. (2013). Revenue optimization in the generalized second-price auction. In EC '13 *Proceedings of the fourteenth ACM conference on Electronic commerce,* 837–852.

Yang, S., Zhou, Y., & Deng, X. (2014). Optimal reserve prices in weighted GSP auctions.

Varian, H. R. (2007). Position auctions. *International Journal of Industrial Organization, 25,* 1163–1178.

Varian, H. R. (2009). Online ad auctions. *American Economic Review: Papers & Proceedings, 99(2),* 430–434.

Vickrey, W. (1961). Counterspeculation, auctions, and competitive sealed tenders. *Journal of Finance, 16(1),* 8–37.

Vorobeychik, Y. (2009). Simulation-based game theoretic analysis of keyword auctions with low-dimensional bidding strategies. In UAI '09 Proceedings of the Twenty-Fifth Conference on Uncertainty in Artificial Intelligence, 583–590.

Yao, L., Chen, W., & Liu, T.-Y. (2012). Convergence analysis for weighted joint strategy fictitious play in generalized second price auction. In *Internet and Network Economics.* Springer, 489–495.

Yuan, J. (2011). Estimating the efficiency improvement of the resource allocation in the Yahoo! keyword auction. *International Journal of Humanities and Social Science,1(18),* 272–284.

Yuan, J. (2012). Examining the Yahoo! sponsored search auctions: a regression discontinuity design approach. *International Journal of Economics and Finance, 4(3),* 139–151.

Zhou, Y., & Lukose, R. (2007). Vindictive bidding in keyword auctions. ICEC '07 Proceedings of the ninth international conference on Electronic commerce, 141–146.

---

[i] In the agent-based model, the computer agents are called turtles.