МАГИСТЕРСКАЯ ДИССЕРТАЦИЯ

MASTER THESIS

Тема: __Аукционы контекстной рекламы при ограниченных мощностях покупателей: эмпирические свидетельства____

____

Title: __Sponsored Search Auctions with Capacity Constrained Bidders: Empirical Evidence____

Студент/ Student:

___*Natalya Naumenko*
(Ф.И.О. студента, выполнившего работу)

Научный руководитель/ Advisor:

_______*Sergei Izmalkov*_
(ученая степень, звание, место работы, Ф.И.О.)

Оценка/ Grade:

____

Подпись/ Signature:

____

Москва 2012

Abstract

Sponsored search is a billion dollar business. Search engines place advertisement on a web page, and merchants pay them for received user clicks. Slots for advertisement are sold via auction-like mechanism where the more agent bids, the higher is the probability to obtain slot that receives more clicks. Traditionally agents valuation of a slot is modeled as a value per click multiplied by the number of clicks that slot delivers. Thus, every click has the same value, no matter how many clicks is already received by the agent. In my masters thesis I suggest an alternative approach. I suppose that the value of click decreases with the number of clicks received by the agent, and that there even might be a point at which it is not optimal for agent to receive more clicks. Thus, I introduce capacity constraints in sponsored search auctions. I consider the properties of equilibria arising in that setting show that in the incomplete information setting the equilibrium is inefficient. I suggest using an alternative mechanism of click allocation among agents based on Ausubel multi-unit auction. Additionally, I use real data to analyze the behavior of agents buying slots and categorize them to show that there really are capacity constrained agents.

1 Introduction

Sponsored search auctions attract a lot of attention during last 10 years. Early works on sponsored search introduced formal models and investigated equilibria arising in different settings. The standard model is as follows. There are N agents willing to advertise their products via Internet, and K slots for advertisement allocation. Usually it is assumed that each slot attracts constant number of clicks, and the higher the slot is allocated on the screen, the more clicks it receives. Since position on the screen defines the attractiveness of a slot, sponsored search auctions are often called position auctions. The number of clicks received by each slot is a common knowledge. Agents have constant click valuation, and therefore agent's payoff from a slot can be written as a number of clicks received by the slot multiplied by the value per click minus pay per click. Thus, in the standard setting auctions for clicks become auctions for slots providing clicks.

In this work I relax the assumption of constant value per click and instead introduce decreasing marginal click valuation. I do so by assuming that agents are capacity constrained: there is a limit of number of clicks agents can effectively handle, and therefore if the agent receives too many clicks his marginal valuation may even become negative. I claim that budget constraint might be seen as consequence of agents' capacity constraints. I show that in that setting the current mechanism of slot allocation is inefficient and suggest using Ausubel mechanism for efficient multi-unit allocation. In addition, I analyze an experiment run by large Russian search engine Yandex. In that experiment some advertisers were unexpectedly moved to the positions on the screen receiving more clicks. I show that there is some evidence supporting the presence of capacity constraints.

The rest of the paper is organized as follows. Section 2 reviews current literature on sponsored search auctions. Section 3 introduces model of sponsored search with capacity constraints. Section 4 analyzes efficient allocation of slots and clicks. Section 5 demonstrates general inefficiency of current mechanism in the presence of capacity constraints. Section 6 suggests an alternative mechanism. Section 7 analyzes Yandex experiment, and Section 8 concludes.

2 Literature overview

Edelman, Ostrovsky, and Schwarz (2007) investigate properties of generalized second-price auction (GSP) used by search engines to sell online advertising. They show that GSP is not equivalent to Vickrey mechanism, and that truth-telling is not an equilibrium strategy in that auction. They introduce the notion of locally envy-free equilibrium – the equilibrium in which no agent can improve his payoff by switching positions with the agent above him. They show that there is a locally envy-free equilibrium in which advertisers receive payoffs coinciding with payoffs in Vickrey mechanism, while in any other locally envy-free equilibrium seller's revenue is at least as high as in that equilibrium. Finally, authors introduce generalized English auction corresponding to GSP and show that it has unique equilibrium in which agents' payments and slot allocation are the same as in Vickrey mechanism.

Independently and around the same time Varian (2007) describes the model of sponsored search auctions which has become standard since then, and analyzes equilibria arising in that setting. He introduces the notion of symmetric Nash equilibria which turns out to be the alternative way to define EOS locally envy-free equilibria. He also analyzes auction revenue. The author points out the similarity between slot allocation problem and the long-known Shapley-Shubik (Shapley and Shubik, 1971) assignment problem. He shows that the distribution of slots among agents is a competitive equilibrium in Shapley-Shubik assignment game in which utilities are of special structure. Finally, using the notion of symmetric Nash equilibria author derives the bounds for agents' click valuations and investigates empirical relationship between bids and values.

Lahaie (2006) analyzes equilibria and welfare implications of sponsored search auctions using both first- and second-price payment rules, and ranking bidders by their bids or by the revenue generated. By revenue Lahaie assumes the product of bid and an individual characteristic of the agent - the click through rate (CTR), the probability of an ad to be clicked if viewed. He shows that in private information setting ranking agents by revenue leads to efficient allocation, while ranking by bid does not. He also proves that in full information setting with first-price payment rule no pure-strategy equilibria exist, while in a second-price payment rule multiple equilibria are possible, and one that leads to the efficient slot allocation is among them.

Subsequent literature concentrated on particular aspects and properties of sponsored search auctions. Budget constraints are usually discussed when the question of distributing budget among different keywords arises. In the presence of budget constraints the question of efficiency or revenue maximizing becomes much more complicated.

Ashlagi et al Ashlagi et al. (2010) suggest an elegant extension of EOS Generalized English Auctions to account for budget constraints. They introduce Generalized Position Auction. In Generalized Position Auction there are K rounds to sell K slots (assuming the number of players N is greater or equal than K). Each round the highest available slot is sold but the price rule is more sophisticated than the rule used in EOS Generalized English Auction. Each round the price ascends from zero to the point at which the number of agents willing to buy the lowest free slot equals the number of free slots. At that point all agents start participating again, the price continues to rise from the previous stop point until the moment at which the number of agents willing to buy previous-to-lowest slot equals the number of free slots minus one, and so forth. The price continues to rise until there is only one participant willing to buy the highest free slot. This participant pays the price and which all others refused to buy that slot. Indeed, this mechanism seems very similar to EOS Generalized English Auction, but there is one significant difference. The price starts not from zero but from the price at which there were two players willing to buy second-highest slot (maybe the winner is not among these two), and the price for second-highest slot starts from the point at which there are exactly three agents willing to buy third-highest slot, and so forth. And at each stop-point all agents can participate (even if the agent was not among those who were willing to buy lower slot she still can participate in the auction for higher slot). The authors prove that there is a truth-telling equilibrium in which all agents truthfully report their valuation and budget. They also show that truth-telling equilibrium results in Pareto-efficient and envy-free distribution of slots, and that any other mechanism which leads to Pareto-efficient and envy-free outcomes necessarily lead to the same distribution of slots and payments as Generalized Position Auction.

Since accounting for budget constraints is both theoretically difficult and practically important, a lot of literature is devoted to algorithmic solutions to maximizing revenue or increasing efficiency. For instance, Abrams et al. (2007) account for bidders' budget constraints and suggest algorithm forecasting query frequencies, pricing and ranking schemes to optimize advertisement delivery to the agents. They test proposed algorithm via simulations, and claim to achieve significant improvements in both seller's revenue and efficiency of the mechanism. However, authors do not take into account the possible change in bidding strategies by the agents in response to suggested changes in advertising delivery mechanism. Borgs et al. (2007) study behavior of budget constrained bidders bidding on more than one keyword. They assume that bidders optimize to achieve equal return-on-investment from each keyword they bid on. That approach allows authors to show that that heuristic may lead to a cycling behavior of the system. To overcome it authors introduce bid perturbations. It allows the system to converge to the point which authors call market equilibrium. For the first-price auction they prove it, and for the second-price auction show it by simulations. Mehta et al. (2005) suggest an algorithm which given agents, their daily budgets, keywords agents bid on and bids for each keyword calculates the matching of bidders with keywords maximizing revenue for the seller. Under the assumption that bids are small compared to budgets the suggested algorithm is time efficient. Again, authors do not take into account the possible change in bidders' behavior in response to introduction of a new algorithm.

In contrast with traditional view on budget constraints there is another one. Traditionally budget constraint is something agent cannot break, cannot physically spend more money than he or she has. Such constraint is called hard budget constraint. On the other hand, sometimes bankrupt firms or unprofitable projects are bailed out by governments or investors. When a firm continues to spend money exceeding its initial budget it is said to have soft budget constraint. The overview of works on soft budgets is given in Maskin Maskin (1999).

I suggest an alternative view on position auctions. Assuming that it is clicks agents are interested in, not the position on the screen, and that all clicks are identical for the agent (thus, the source of click is of no interest itself) I consider position auction as multi-unit auctions in which clicks are sold in bundles and every agent can buy only one such bundle. Good overview of multi-unit auctions is given in Khrishna Krishna (2010).

3 Model

The standard model of position auction is as follows. There are N players and K slots. Each slot receives γ_k clicks, and without loss of generality, we assume that $\gamma_1 > \gamma_2 > \cdots > \gamma_K$. Each agent i has a constant valuation of a click v_i. Thus, if the agent obtains slot number k and pays p per click, his payoff is $u_i = \gamma_k(v_i - p)$. The number of slots, agents, and clicks received by each slot are common knowledge. In the full information game all valuations are common knowledge as well, whereas in incomplete information setting v_i is derived according to the commonly known distribution function.

The game proceeds as follows. Each agent submits bid s_i, then bids are ordered and K highest bids are granted corresponding slots. Thus, the highest bid gets the first slot, the second highest gets the second slot and so on. Ties are broken randomly. For simplicity of the notation let us renumber bids so that $s_1 > s_2 > \cdots > s_N$. Every agent who obtained a slot pays next-highest bid for clicks that slot receives. For instance, the agent getting slot number one pays $\gamma_1 s_2$. If $N \leq K$ then payment for slot number N (the lowest slot sold) is zero.

I want to concentrate on a situation when it is not optimal for agent to receive unlimited number of clicks, even for free. I claim that if number of clicks made by users is positively correlated with the number of purchases users make, then our assumption is indeed relevant. Every agent is capacity constrained. Even Amazon cannot handle infinite number of orders at a time. If there are too many clicks on agent's advertisement (and as a result too many orders), then some users get unsatisfied with service provided by the agent. Therefore it might be optimal for agent to restrict number of clicks received to some reasonable amount.

The simplest (although not the only one) way to introduce capacity constraint is to add a parameter reflecting agent's capacity. Therefore, I modify the standard model in the following way. Each agents still has a valuation of click a_i, $i = 1, \dots, N$. The value of click may be interpreted as expected profit from one click. I also assume that there is a parameter b_i corresponding to agent's capacity (the higher b_i is, the more clicks agent wants to process, all else being equal). Thus, I write agent's payoff u_i as quadratic function:

$$u_i(x) = a_i x - \frac{1}{2b_i} x^2 - px$$

where x is the number of clicks received ($x = \gamma_k$ for some k, or $x = 0$ is the agent did not get a slot), and p is the price per click. Note that if $p = 0$ then the optimal number of clicks for agent i is $a_i b_i$, and agent does not want to get more than $2a_i b_i$ clicks even for free.

In the full information setting all (a_i, b_i) are common knowledge. In the incomplete information setting (a_i, b_i) are distributed according to the distribution function $F_i(\cdot)$, $i = 1, \dots, N$. The rules of the auction remain standard: all agents submit bids, bids are ordered from highest to the lowest, K highest bids receive slots, and pay next-highest bid per click.

4 Efficient allocation of slots and clicks

By efficient allocation of either slots or clicks I understand such allocation of slots or clicks among agents that generates maximum aggregated payoff to agents. Without capacity constraints efficient allocation of slots is obvious: we give slot receiving maximum amount of clicks to the agent with highest click valuation, the second-highest slot to the agent with second-highest valuation, and so forth. However, when capacity constraints are present, the question of efficient slot allocation becomes a non-trivial problem. Therefore in this section I consider in more details efficient allocation of slots and clicks among capacity-constrained agents.

4.1 Efficient slot allocation

Let $\sigma : K \to N$ be an injective function mapping slots to agents. Thus, $\sigma(k)$ is the number of agent that gets slot number k, and every agent can get at most one slot. The problem of efficient slot distribution can then be written as

$$\sum_{k=1}^{K}(a_{\sigma(k)}\gamma_k - \frac{1}{2b_{\sigma(k)}}\gamma_k^2) \to \max_{\sigma(\cdot)}$$

Since the number of possible slot distributions among agents is finite, the optimal slot allocation always exists. There is a famous Hungarian algorithm suggested by Kuhn and Yaw (1955) solving allocation problem in polynomial time, but for our purposes it is only important to say that the allocation is non-trivial, and that there is no simple allocating rule. Below are two numerical examples to illustrate that point.

`Example 1.` Let $N = K = 2$, $\gamma_1 = 100$, $\gamma_2 = 80$, and $(a_1, b_1) = (1, 200)$, $(a_2, b_2) = (2, 50)$. We may interpret parameters saying that Agent 1 is a discounter with high sales volumes but low margin, and Agent 2 is a small firm with higher margin but lower sales volume. Then it is straightforward to check that in efficient slot distribution higher slot (slot number 1) should go to an agent with higher capacity (Agent 1), and lower slot should be given to a smaller agent (Agent 2).

`Example 2.` Let us slightly change the numbers. All parameters are the same as in the previous example but $\gamma_2 = 20$. Then the reader may easily check that giving first slot to Agent 1 and second slot to Agent 2 generates total payoff of 111, while giving first slot to Agent 2 and second slot to Agent 1 yields total payoff of 119. Thus, in efficient slot allocation higher slot should go to a smaller agent! The interpretation is as follows. The total amount of clicks received by two slots is relatively small and therefore it turns out to be more efficient to give most part of those clicks to the smaller agent since she can extract higher revenue from it.

4.2 Efficient click allocation

I have already mentioned that slots might be viewed as bundles of clicks. Obtaining particular slot can be considered as buying a predefined bundle or flow of clicks. Then it is natural to suggest breaking up the bundles and distributing clicks among agents, not slots.

Formally, let μ_{ik} be the "share" of clicks from slot k that goes to agent i, $k = 1, \dots, K$, $i = 1, \dots, N$. Then μ_{ik} can be viewed as probability for Agent i's advertisement to be placed on slot k. Two important conditions should be met: first, each slot shares cannot in total make greater than 1, since each slot cannot present more than one advertisement at a time. And second, each agent cannot be allocated more than one slot at a time. Therefore, each agent's sum of slot shares also cannot be greater than 1. Thus, the problem of efficient click allocation among agents can be written as:

$$\sum_{i=1}^{N}(a_i x_i - \frac{1}{2b_i}x_i^2) \to \max_{x_1,\dots,x_N}$$

$$s.t. \quad x_i = \sum_{k=1}^{K} \mu_{ik}\gamma_k$$

$$\forall i = 1, \dots, N \quad \sum_{k=1}^{K} \mu_{ik} \leq 1$$

$$\forall k = 1, \dots, K \quad \sum_{i=1}^{N} \mu_{ik} \leq 1$$

$$0 \leq \mu_{ik} \leq 1$$

The total amount of clicks is $\sum_{k=1}^{K} \gamma_k$, and those clicks are divided among agents. The only important constraint is that no agent can be placed in more than one slot at a time, and therefore no one can get more than 1 shares of slots in total. The solution to the problem $M = (\mu_{ik})$ and $x_i = \sum_{k=1}^{K} \mu_{ik}\gamma_k$ is the optimal click distribution among agents. It is simple linear optimization problem and the solution can be found using standard methods. Since optimal slot allocation calculated in previous section is always among feasible click distributions, in general breaking up slots and distributing clicks among agents leads to an increase in efficiency. To illustrate that point, let us calculate efficient click allocations for two previous examples.

`Example 1 (continued).` Remember, we had two agents and two slots, $\gamma_1 = 100$, $\gamma_2 = 80$. In efficient slot allocation we gave first slot to a bigger agent (Agent 1), and the second slot to a smaller agent (Agent 2). It turns out that it would be the best to give 104 clicks to Agent 1, and 76 clicks to Agent 2. However, giving 104 clicks to the Agent 1 would mean that this agent's advertisement should be allocated on both available slots sometimes. Since it is impossible, an efficient click allocation coincides with efficient slot allocation in that example: we give 100 clicks to Agent 1, and 80 clicks to Agent 2.

`Example 2 (continued).` In that example the second slot was much smaller: $\gamma_1 = 100$, $\gamma_2 = 20$. Solving linear optimization problem of Efficient click allocation among agents we get $x_1 = 56$ and $x_2 = 64$. Thus, Agent 1 receives 56 clicks, and Agent 2 gets 64 clicks. In terms of slot shares it means that Agent 1 receives a share of 0.45 of slot 1 and 0.55 of slot 2; and Agent 2, vice versa, gets share of 0.55 of slot 1 and 0.45 of slot 2. The aggregate payoff is 135.2. In efficient slot allocation first slot was given to Agent 2, and second slot to Agent 1, and the aggregate payoff was 119. Thus, the ability to "break up" slots selling clicks separately may sometimes noticeably increase efficiency.

5 Equilibrium analysis

Let us consider incomplete information setting. Each agent i knows his own private type (a_i, b_i), and submits bid s_i. All bids are then ordered, the highest bid receives first slot and pays second-highest bid per click, the second bid receives second slot and pays third-highest bid per click, and so on. Thus, each agent i has to choose bid s_i to maximize his expected payoff:

$$\mathbb{E}u_i = \mathbb{E}(a_i x - \frac{1}{2b_i}x^2 - px) \to \max_{s_i}$$

where x is the number of clicks ($x = \gamma_k$ for some k, or x is zero if the agent does not get any slot), and p is the price per click. Note that the higher submitted bid s_i is, the higher are both expected number of clicks $\mathbb{E}x$ and expected payment per click $\mathbb{E}p$.

Claim 1. *Equilibrium strategies s_i are continuous in (a_i, b_i).*

Proof. Since $\mathbb{E}u_i$ is a strictly concave function of x it is a simple consequence of the Maximum Theorem. □

Claim 2. *Equilibrium strategies s_i are nondecreasing in a_i and b_i.*

Proof. Note that both $\frac{\partial \mathbb{E}u_i}{\partial a_i}$ and $\frac{\partial \mathbb{E}u_i}{\partial b_i}$ are nondecreasing in x. Therefore expected payoff has increasing differences property with respect to (x, a_i) and (x, b_i). Thus (see Milgrom and Shannon (1994)), optimal bid s_i should be a nondecreasing function of both a_i and b_i. The intuition is straightforward: if it is optimal for agent to bid some amount s when the agent's type is (a, b), then it is still optimal to bid at least s when either a or b increases. □

Theorem 1. *In incomplete information setting equilibrium allocation of slots is inefficient.*

Proof. On the contrary, suppose that for any profile of types $(a_i, b_i)_{i=1}^N$ the equilibrium allocation is efficient. Let us consider a game with one slot ($K = 1$) and two agents ($N = 2$), with types (a, b_1) and (a, b_2), $b_1 > b_2$. Thus, the agents differ only in their capacity, and Agent 1 is bigger. Let as assume also that the amount of clicks delivered by the slot, γ, is not too big, so both agents are better with slot than without it: $\gamma < ab_2$. Agents submit bids $s_1(a, b_1)$ and $s_2(a, b_2)$. Since by assumption equilibrium allocation is efficient, the slot should go to the bigger agent (Agent 1). Therefore the following inequality should hold:

$$s_1(a, b_1) > s_2(a, b_2)$$

Because the inequality is strict and because strategies are continuous in (a, b), the inequality should hold in sufficiently small neighborhood of (a, b_2). Thus, there exists $\epsilon > 0$ such that for any $(a'_2, b'_2) \in U_\epsilon(a, b_2)$

$$s_1(a, b_1) > s_2(a'_2, b'_2)$$

Let us now consider $(a', b_2) \in U_\epsilon(a, b_2)$ such that $a' > a$. Since $(a', b_2) \in U_\epsilon(a, b_2)$, $s_1(a, b_1) > s_2(a', b_2)$, and in equilibrium the slot still goes to the first agent. But if

$$\gamma < \frac{2(a' - a)}{\frac{1}{b_2} - \frac{1}{b_1}}$$

then

$$a'\gamma - \frac{1}{2b_2}\gamma^2 > a\gamma - \frac{1}{2b_1}\gamma^2$$

and it is more efficient to give slot the the second agent. Thus, the equilibrium allocation is inefficient. A contradiction. □

Note that we did not rely on the assumption of symmetric equilibrium strategies or ex-ante symmetric agents.

6 Alternative auction design

Since we have shown that current auction design generally leads to inefficient slot allocation, the natural question to ask is what design would lead to the efficient one? As it was shown earlier, to allocate clicks efficiently we should consider selling not slots but clicks to agents. Thus, the mechanism takes form of a much more familiar multi-unit auction.

We know that in single-unit case sealed-bid Vickrey mechanism implements efficient allocation, and that there is generalization for multi-unit case (see, for example, Krishna (2010)). There is an elegant implementation of Vickrey mechanism suggested by Ausubel (see Ausubel (2004)). In that implementation the auctioneer gradually increases price and the units are allocated to bidders at current price whenever they are "clinched", that is, whenever aggregate demand of all other participants is lower than the number of units to sell. See the original paper for the explicit example.

Ausubel's mechanism allows for external constrains like "no more than three units" to a bidder. Thus, all we need to implement efficient click allocation is to ensure that no agent gets more clicks than one slot can provide, that is, no agent's advertisement is placed on more than one slot at a time (see Section 4.2). Therefore I define "clinch" rule as follows: the bidder is guaranteed clicks at current price any time when aggregate demand of all other participants becomes less than current aggregate supply of clicks without the highest unsold yet slot. That is, any time aggregate demand of all other participants can be allocated on all unsold slots without the largest one, we sell that largest slot. This rule prevents selling more clicks than one slot can provide, and therefore implements efficient allocation described in Section 4.2. Thus, there are all means to allocate clicks efficiently among agents.

Note that in this section I did not actually rely on a particular form of agents' payoff function. To apply Ausubel auction I used only the nonincresing marginal utility of clicks. Thus, the suggested mechanism may be applied to any form of payoff function with nonicreasing marginal utility property. In particular, in case of constant click valuation the resulting click allocation coincides with efficient slot allocation (the highest slot goes to the agent with highest click valuation, and so forth), and resulting prices coincide with Vickrey prices.

7 Testing for capacity constraints

I have shown that presence of capacity constraints may significantly affect agents' behavior and might lead to generally inefficient outcomes. Thus, testing for the presence of capacity constraints is important. I use data from large Russian search engine Yandex to analyze if there are capacity constraints.

7.1 Experiment Design

Yandex generally displays advertisement on three different areas on the screen: *spetsrazmeschenie* (SR), *garantia* (GR), and *dinamica*. *SR* locates below the search string, there are no more than three advertisement slots in it. *Garantia* locates to the right to the results of generic search, there are maximum four slots. See Figure 1 for the example of *SR* and *garantia* advertisement slots. *Dinamica* locates below the *garantia* positions, and again no more than four advertisements is displayed there. Because of their location, *SR* slots, if displayed, usually receive much more clicks than any other slots. Thus, although, all else being equal, slot located higher on the search screen receives more clicks, first three slots receive much more clicks than all others.

For advertisement to be displayed in *SR* an important condition should be met. On the basis of advertisement's history Yandex estimates advertisers's click-through rate (CTR). Estimated CTR multiplied by bid should be higher than a certain threshold. Minimum bid needed for advertisement to be located in *SR* is displayed in advertiser's web interface, and therefore is known to the advertiser. Each time some user executes search an auction similar to the described in Section 3 runs, winning ads is then displayed on the page with search results. If no ads meets threshold condition then no *SR* is displayed on the search screen.

During summer of 2011 Yandex run a series of experiments with *SR* thresholds. For the advertisement displayed in St. Petersburg and all Russian regions except Moscow (Regions) it decreased the threshold by 10% on June 30, then on July 19, and finally on August 11. The threshold for advertisement displayed in Moscow was not changed. As a result, in St. Petersburg and Regions for some advertisement the probability of being displayed in *SR* increased, and therefore some ads started receiving much more clicks.

7.2 Data

For analysis I have random sample from Yandex search logs[1] in Moscow and Regions for August 1 – 30, 2011. Those logs contain all information on search, clicks, etc. Thus, I can analyze only one stage of the experiment: the outcomes of lowering the thresholds on August 11.

I take advertisement campaign as a unit of observation. Generally speaking, advertisement campaign is an advertisement linked to a list of keywords, with specified budget and bid. In addition, the geographical area of displaying ads is usually specified. For example, some ads might be displayed only for Moscow, some only for Novosibirsk, other – for Moscow area except Moscow, etc. When a user executes search, all suitable advertisements participate in an auction similar to described in Section 2. If at least one advertisement meets the threshold, it is displayed in *SR* area, otherwise winning advertisements are displayed in *garantia* area.

From the logs I extract those advertisement campaigns that were displayed only in Moscow and only in Regions during August 1 – 10, before lowering the threshold. I check for the migration between groups during August 11 – 28. Table 1 represents the summary statistics on groups' size and migration rate between groups. The number of advertisement campaigns in group is displayed in Column (I). The number of campaigns that were displayed in other geographical area during August 11 – 28 (after lowering the thresholds) is in columns (II) – (III). It turns out that very few advertisement campaigns migrated.

Moscow is generally very different from Regions. The absolute number of clicks sold in Moscow is several times higher than in Regions, the are more advertisers in Moscow, and the price for clicks is higher (see Figure 2). Nevertheless, trends are generally similar for both groups. Figures 3 and 4 demonstrate average bid made and average number of clicks received by campaign during a day. The graphs a

[1]The size of sample is from %10 to %50 of all logs. I do not report the number more precisely since it might reveal commercial secret information.

rescaled to show the similarities in trends.

It can be seen from the figures that the amount of clicks sold decreases on weekends. It is well known phenomena – during weekends traffic decreases, and many advertisers prefer not to display their ads on weekends and holidays. In addition, on August 19 was a catastrophic event – Yandex failed. In all my estimates below I control for weekends, and exclude the data on August 19.

It is hard to evaluate the effect of lowering the thresholds from the plots of average clicks received and average bid made. Though from Figure 4 we can see that the next week after August 11 the total number of clicks sold increased in Regions compared to Moscow. Still, careful analysis is needed to test if there are capacity constraints on number of clicks received by advertisers.

7.3 Testing for capacity constraints

I can only indirectly test for the presence of capacity constraints among advertisers. The idea is to check if advertisers receiving more clicks than expected after lowering the threshold turned off their campaigns or significantly decreased their bids to stop receiving so many clicks.

First, I calculate the amount of clicks received by each campaign before and after lowering the threshold, and use simple difference-in-difference analysis to check if there are differences in trends in Moscow and Regions after lowering the threshold. Using aggregated amounts of clicks allows to account for the fact that not receiving clicks is an important event which I do not observe directly. Then I do the same exercise with the amount of money paid by advertisers before and after lowering the thresholds. The results are presented in Table 2, Panel A. Both the amount of clicks received and the amount of money paid changed significantly after lowering the threshold. The estimates, however, are very rough, and more detailed analysis is needed.

Next, I disaggregate the group participants to estimate the effect of lowering the threshold on different campaigns. For each advertisement campaign I calculate the number of clicks received while displayed in *SR* during August 1 – 10, before lowering the thresholds. I also calculate the number of clicks received while displayed on position 1 in *garantia*, and on positions 2 and 3 in *garantia*. The reasoning is simple: if adveritisement was placed on top positions before the experiment, it should be influenced more by lowering the thresholds, than the adveritisement from the lower positions. I mark as "SR mostly" advertisement campaigns that received more clicks from *SR* positions than from top 3 positions in *garantia*; I mark as "GR 1 mostly" those that received most of their clicks from first position in *garantia*; finally, I mark as "GR 2 – 3 mostly" those campaigns that received most clicks from positions 2 and 3 in *garantia*.

I analyze the impact of lowering the threshold on the amount of clicks received and the amount of money paid controlling for the position on the screen before lowering the threshold. The results are presented in Table 2, Panel B. As it can be seen from the table, the difference in trends is still significant.

Advertisers might need time to note that they started receiving more clicks and are paying more money. Learning new "rules of the game" and reoptimizing strategies takes some time. Therefore it is reasonable to expect that the reaction to lowering the threshold varies during time. Luckily for me, the threshold was lowered on August 11, so I can divide available information into three roughly equal parts. Using difference-in-difference analysis I compare first decade of the month (before lowering the threshold) separately with second and third decades. The results for clicks are presented in Table 3, and the results for the amount paid are in Table 4. We see that though second decade differs significantly from first decade, both without controls and with position controls, the third decade of the month seem to be unaffected by lowering the threshold. There is no differences in trends for both clicks received and money paid between decade 1 and decade 3. Thus, there is some evidence supporting the idea that advertisers react to increased number of clicks in a way to put things back and to return received clicks to the previous level.

To examine the impact of lowering the threshold more precisely I estimate a panel regression with the number of clicks received per day as a dependent variable. I control for time adding dummies for second and third decades. The results are reported in Table 5, Column (I). It can be seen that lowering the threshold significantly increased the number of received clicks in both decades. However, when I add position controls (Column (II)), I do observe strong negative effect in the third decade for those campaigns which were mostly in *SR* before lowering the threshold. Thus, it appears that those campaigns that were mostly on top positions before lowering the threshold in third decade started receiving fewer clicks.

The estimates might be biased since I do no observe directly no-click events. To account for that fact I treat days with no observations for particular campaign as days when that campaign received zero clicks. I repeat my estimates using this artificially enriched panel, both controlling only for time and for time and position on the screen. The results are presented in Table 5, Columns (III) and (IV). The estimates are qualitatively the same. Effect of lowering the threshold on clicks is positive and significant in both decades, and effect for SR campaigns is significantly negative. Thus, there is even more indirect evidence that top campaigns reacted somehow to decrease the number of clicks after lowering the threshold. However, using information on clicks an alternative explanations cannot be ruled out. Top campaigns might have started receiving less clicks because of increased competition in *SR*.

8 Conclusion

It is still very much work in progress. Nevertheless, in the simplest setting I show that presence of capacity constraint makes current mechanism of slots allocation inefficient. I also suggest an alternative mechanism of click allocation based on

Ausubel multi-unit auction. Finally, I provide empirical evidence indicating that capacity constraints indeed exist.

There are two major ways for the further research. First, empirical estimates should be continued to accept or reject the existence of capacity constraints. Currently there is only indirect evidence, and advertisers' behavior should be studied more carefully. Since not all data on experiments is explored now there are all conditions to do so. Second, the model should be closed to incorporate the reasons for capacity constraints.

Though very large literature is devoted to position auctions, it seems that the current theory is still a way behind full explanation of advertisers' behavior. From the practitioner's point of view a lot of algorithms and optimization procedures are suggested already. Theory still has no clear answers on the sources of click valuations and budget constraints. It limits the ability of evaluating suggested algorithms.

References

Zoe Abrams, Ofer Mendelevitch, and John A. Tomlin. Optimal delivery of sponsored search advertisements subject to budget constraints. In *Proceedings of the 8th ACM conference on Electronic commerce*, pages 272–278, 2007.

Itai Ashlagi, Mark Braverman, Avinatan Hassidim, Ron Lavi, and Moshe Tennenholtz. Position auctions with budgets: Existence and uniqueness. *The B.E. Journal of Theoretical Economics*, Vol. 10, 2010.

Lawrence M. Ausubel. An efficient ascending-bid auction for multiple objects. *The American Economic Review*, Vol. 94(5):1452–1475, 2004.

Jean-Pierre Benoit and Vijay Krishna. Multiple-object auctions with budget constrained bidders. *The Review of Economic Studies*, 68:155–179, 2001.

Christian Borgs, Nicole Immorlica, Jennifer Chayes, and Kamal Jain. Dynamics of bid optimization in online advertisement auctions. In *Proceedings of the 16th International World Wide Web Conference*, pages 713–723, 2007.

Benjamin Edelman, Michael Ostrovsky, and Michael Schwarz. Internet advertising and the generalized second price auction: Selling billions of dollars worth of keywords. *The American Economic Review*, 97:242–259, 2007.

Jon Feldman, S. Muthukrishnan, Martin Pal, and Cliff Stein. Budget optimization in search-based advertising auctions. In *Proceedings of the 8th ACM conference on Electronic commerce*, 2007.

Maria florina Balcan, Avrim Blum, Jason D. Hartline, and Yishay Mansour. Sponsored search auction design via machine learning.

Vijay Krishna. *Auction theory, 2nd edition.* Academic Press, 2010.

H. W. Kuhn and Bryn Yaw. The hungarian method for the assignment problem. *Naval Research Logistics Quarterly*, pages 83–97, 1955.

Sebastien Lahaie. An analysis of alternative slot auction designs for sponsored search. In *Proceedings of the 7th ACM conference on Electronic commerce*, pages 218–227. ACM Press, 2006.

Eric S. Maskin. Recent theoretical work on the soft budget constraint. *The American Economic Review*, pages 421–425, 1999.

Aranyak Mehta, Amin Saberi, Umesh Vazirani, and Vijay Vazirani. Adwords and generalized on-line matching. In *FOCS 05: Proceedings of the 46th Annual IEEE Symposium on Foundations of Computer Science*, pages 264–273. IEEE Computer Society, 2005.

Paul Milgrom. *Putting auction theory to work.* Cambridge University Press, 2004.

Paul Milgrom and Chris Shannon. Monotone comparative statics. *Econometrica*, 62: 157–180, 1994.

L. S. Shapley and M. Shubik. The assignment game I: The core. *International Journal of Game Teory*, 1:111–130, 1971.

Hal R. Varian. Position auctions. *International Journal of Industrial Organization*, 25:1163–1178, 2007.

Figure 1: **Yandex search page.** *Spetsrazmeschenie* and *garantia* areas are marked by large black bars.

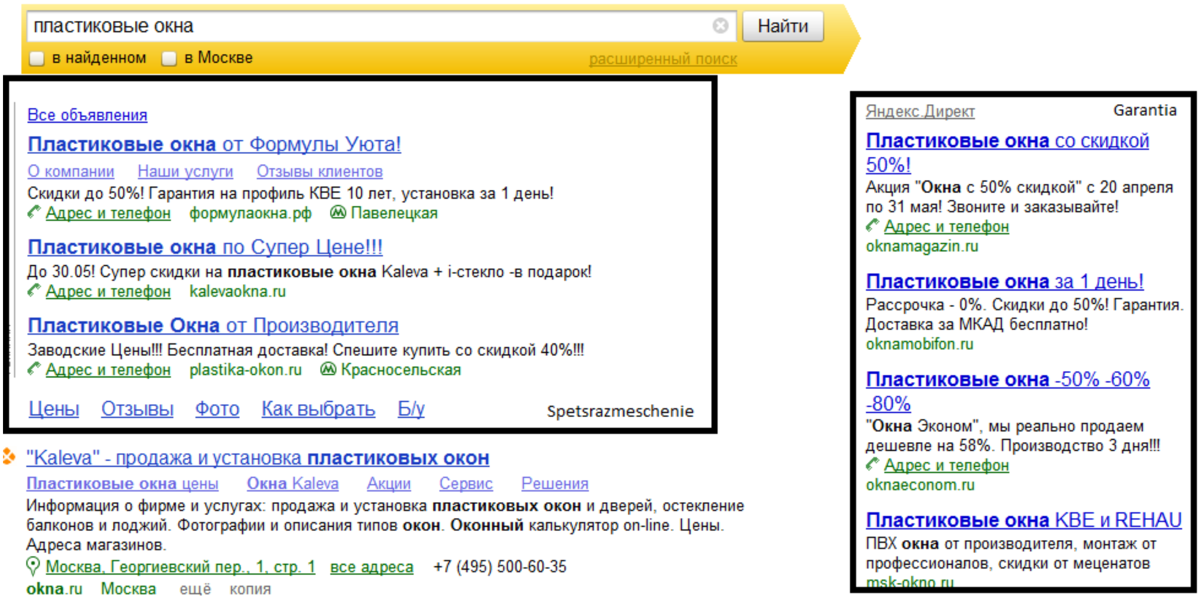

Table 1: **Summary statistics on groups**. Column (I) represents the number of advertisement campaigns is group on August 1 – 10. Columns (II) – (III) show the migration between groups: number of advertisement campaigns that were displayed in other geographical area during August 11 – 30.

	August 1 – 10	Migration during August 11 – 28	
Group	N. of campaigns (I)	Moscow (II)	Regions (III)
Moscow	56 973	–	988 (1.73%)
Regions	39 206	941 (2.40%)	–

Figure 2: **Number of clicks sold.** Aggregate number of clicks sold per day. Solid line: Moscow, dashed line: Regions.

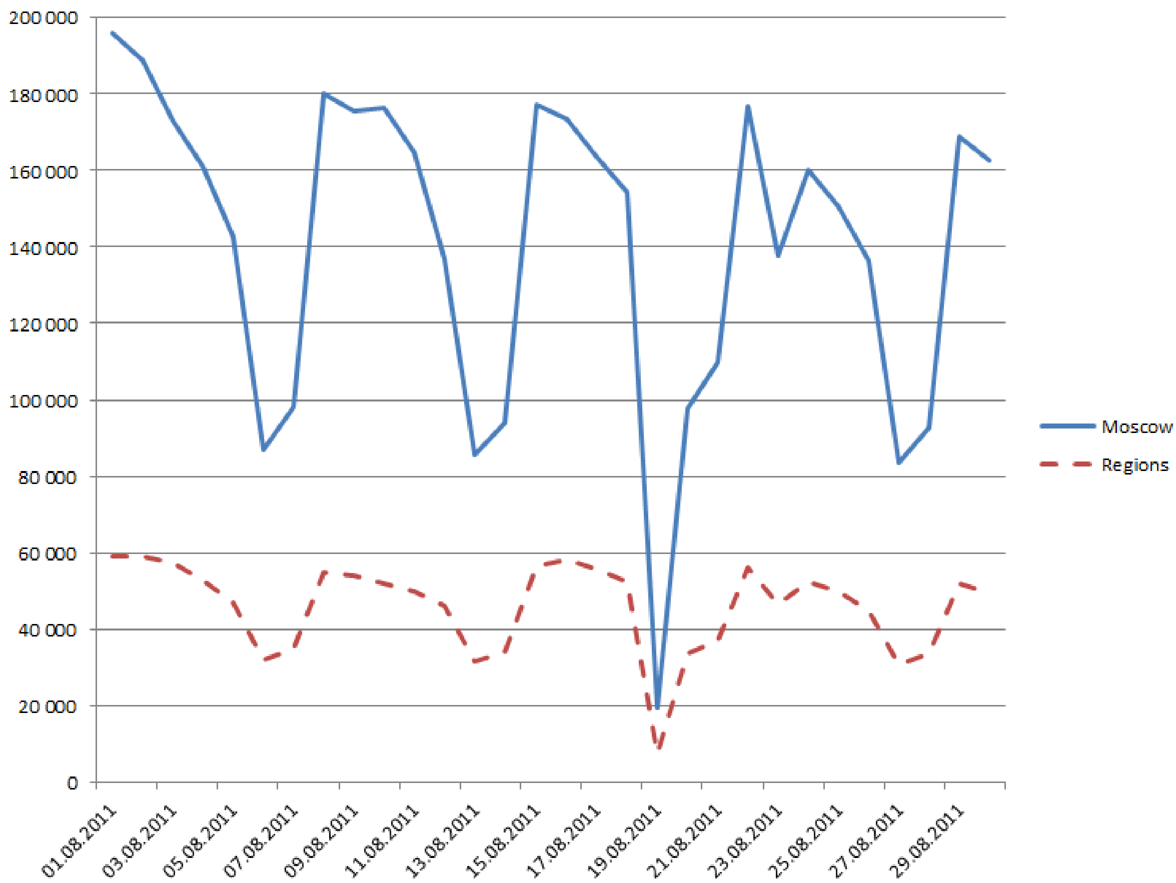

Figure 3: **Average bid.** Average daily bid made by advertisers rescaled to show similarities in trends. Solid line: Moscow, dashed line: Regions.

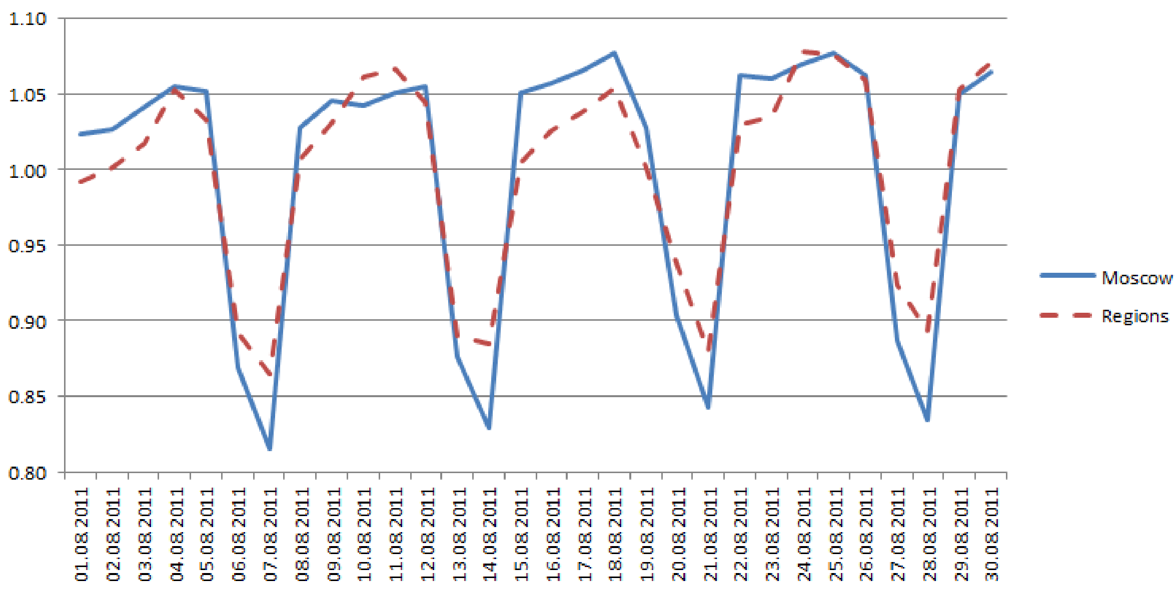

Figure 4: **Clicks per order.** Average daily number of clicks per order received by advertisers, rescaled to show similarities in trends. Solid line: Moscow, dashed line: Regions.

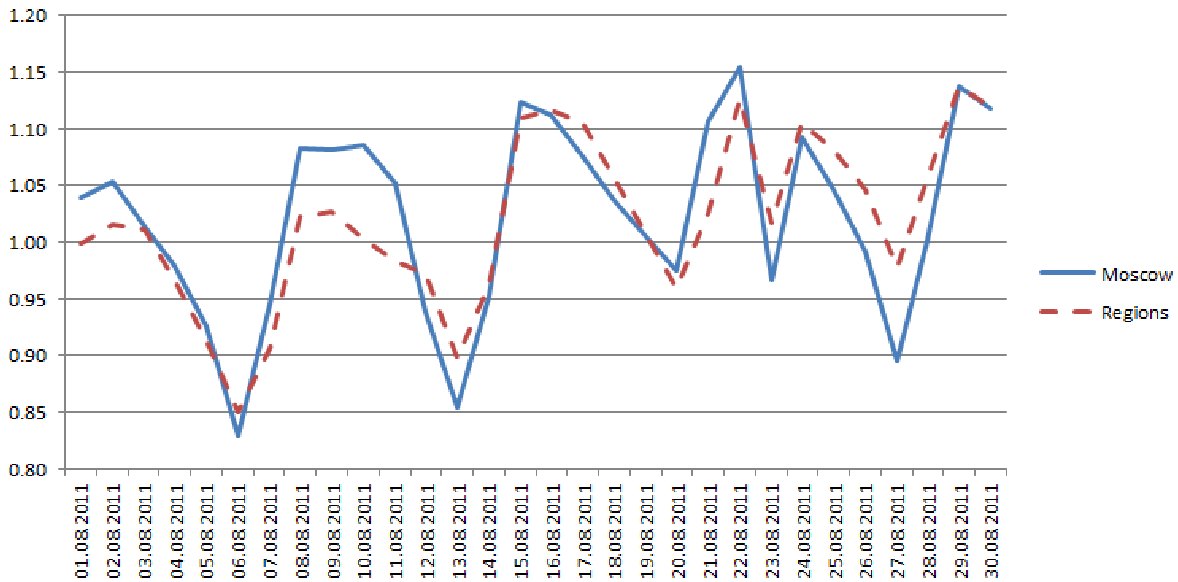

Table 2: **Differences in trends between Moscow and Regions**. This table analyzes outcome of lowering the *spetsrazmeschenie* threshold on the amount of clicks received by advertising campaign and on the total payments made by campaign. *Clicks* is the aggregate amount of clicks received by campaign during August 1 – 10, before lowering the threshold, and during August 11 – 30, after lowering the threshold. *Aggregate cost* is the total amount of money paid by campaign for the received clicks before and after lowering the threshold, in cents. *Panel A* represents plain difference-in-difference analysis, while *Panel B* controls for the position of advertisement on the screen before lowering the threshold: mostly *spetsrazmeschenie*, mostly first place in the *garantia*, or mostly second of third place in the *garantia*. Standard errors are in parentheses.

	Clicks			Aggregate cost		
Panel A: No controls						
	Moscow	Regions	Difference	Moscow	Regions	Difference
Before	40.482***	19.577***	-20.905***	3,957.214***	683.325	-3,273.889***
	(0.818)	(1.008)	(1.298)	(83.530)	(102.934)	(132.562)
After	70.253***	34.085***	-36.168***	6,782.606***	1,105.934***	-5,676.672***
	(0.832)	(1.007)	(1.306)	(84.994)	(102.798)	(133.385)
Difference	29.771***	14.508***	-15.263***	2,825.392***	422.609***	-2,402.784***
	(1.167)	(1.425)	(1.841)	(119.169)	(145.475)	(188.054)
Panel B: Controlling for the position on the screen						
	Moscow	Regions	Difference	Moscow	Regions	Difference
Before	30.980***	13.796***	-17.184***	2,343.067***	-486.9	-2,829.926***
	(1.639)	(1.758)	(1.306)	(166.096)	(178.2)	(132.376)
After	60.959***	28.827***	-32.132***	5,201.00***	28.42	-5,172.58***
	(1.648)	(1.758)	(1.314)	(167.0)	(178.1)	(133.149)
Difference	29.979***	15.031***	-14.948***	2,857.933***	515.32***	-2,342.246***
	(1.161)	(2.486)	(1.833)	(235.535)	(251.942)	(185.760)
Number of campaigns	128,090			128,090		

***, **, and * denote significance at the 1, 5, and 10 percent levels respectively

Table 3: **Clicks: differences in trends between Moscow and Regions by decades**. This table analyzes the outcome of lowering the *spetsrazmeschenie* threshold on the amount of clicks received by advertising campaign comparing first decade separately with second and third decades. *Clicks* is the aggregate amount of clicks received by campaign during August 1 – 10, before lowering the threshold, and during August 11 – 30, after lowering the threshold. *Aggregate cost* is the total amount of money paid by campaign for the received clicks before and after lowering the threshold, in cents. *Panel A* represents plain difference-in-difference analysis, while *Panel B* controls for the position of advertisement on the screen before lowering the threshold: mostly *spetsrazmeschenie*, mostly first place in the *garantia*, or mostly second of third place in the *garantia*. Standard errors are in parentheses.

Dependent variable: Aggregate number of clicks received by campaign						
	Decade 1 vs Decade 2			Decade 1 vs Decade 3		
Panel A: No controls						
	Moscow	Regions	Difference	Moscow	Regions	Difference
Before	40.482***	19.577***	-20.905***	40.482***	19.577***	-20.905***
	(0.533)	(0.657)	(0.846)	(0.570)	(0.703)	(0.905)
After	36.692***	18.844***	-17.848***	41.538***	20.935***	-20.603***
	(0.567)	(0.701)	(0.901)	(0.618)	(0.767)	(0.985)
Difference	-3.790***	-0.733	3.057**	1.056	1.358	0.303
	(0.778)	(0.961)	(1.236)	(0.841)	(1.040)	(1.338)
Panel B: Controlling for the position on the screen						
	Moscow	Regions	Difference	Moscow	Regions	Difference
Before	33.063***	14.623***	-18.441***	33.118***	14.840***	-18.279***
	(1.125)	(1.202)	(0.852)	(1.196)	(1.278)	(0.912)
After	29.116***	13.879***	-15.237***	34.021***	16.125***	-17.90***
	(1.140)	(1.225)	(0.906)	(1.219)	(1.314)	(0.989)
Difference	-3.947***	-0.744	3.204***	0.903	1.285	0.384
	(0.774)	(1.716)	(1.230)	(0.837)	(1.833)	(1.331)
Number of campaigns	121,784			119,477		

***, **, and * denote significance at the 1, 5, and 10 percent levels respectively

Table 4: **Aggregate cost: differences in trends between Moscow and Regions by decades**. This table analyzes the outcome of lowering the *spetsrazmeschenie* threshold on the payments made by advertising campaign comparing first decade separately with second and third decades. *Clicks* is the aggregate amount of clicks received by campaign during August 1 – 10, before lowering the threshold, and during August 11 – 30, after lowering the threshold. *Aggregate cost* is the total amount of money paid by campaign for the received clicks before and after lowering the threshold, in cents. *Panel A* represents plain difference-in-difference analysis, while *Panel B* controls for the position of advertisement on the screen before lowering the threshold: mostly *spetsrazmeschenie*, mostly first place in the *garantia*, or mostly second of third place in the *garantia*. Standard errors are in parentheses.

Dependent variable: Aggregate cost paid by campaign						
	Decade1 vs Decade2			Decade1 vs Decade3		
Panel A: No controls						
	Moscow	Regions	Difference	Moscow	Regions	Difference
Before	3,957.214***	683.325***	-3,273.889***	3,957.214***	683.325***	-3,273.889***
	(56.063)	(69.087)	(88.972)	(58.851)	(72.523)	(93.397)
After	3,545.347***	613.911***	-2,931.436***	4,007.258***	676.666***	-3,330.592***
	(59.588)	(73.664)	(94.748)	(63.772)	(79.096)	(101.602)
Difference	-411.867***	-69.414	342.454***	50.044	-6.659	-56.702
	(81.816)	(100.992)	(129.973)	(86.777)	(107.312)	(138.007)
Panel B: Controlling for the position on the screen						
	Moscow	Regions	Difference	Moscow	Regions	Difference
Before	2,732.901***	-243.0***	-2,975.890***	2,693.082***	-263.8***	-2,956.882***
	(117.290)	(125.3)	(88.822)	(122.315)	(130.8)	(93.258)
After	2,294***	-309.3***	-2,603.3***	2,718***	-268.6***	-2,986.6***
	(118.9)	(127.7)	(94.439)	(124.7)	(134.4)	(101.148)
Difference	-438.901***	-66.3	372.409***	24.918	-4.8	-29.374
	(167.015)	(178.906)	(128.263)	(174.674)	(187.542)	(136.183)
Number of campaigns	121,784			119,477		

***, **, and * denote significance at the 1, 5, and 10 percent levels respectively

Table 5: **Clicks.** This table examines the impact of lowering *spetsrazmeschenie* threshold on the number of clicks received by advertising campaign per day. *D2* and *D3* are controls for the second and third decade of the month. *Threshold lowered* indicated that the threshold was lowered for the campaign. *SR mostly*, *GR 1 mostly*, and *GR 2–3 mostly* are position dummies described in text. *Clicks before* is the amount of clicks received before lowering the threshold. Column (I) reports estimates with only time and treatment controls. Column (II) adds advertisement position covariates. Columns (III) and (IV) repeat estimates treating no observations as zero-click events. Clustered (campaign) standard errors are in parentheses.

	Dependent variable: number of clicks			
	(I)	(II)	(III)	(IV)
D2	-0.022	-0.022	-0.230***	-0.231***
	(0.033)	(0.033)	(0.025)	(0.025)
D3	0.174***	0.174***	0.019	0.019
	(0.041)	(0.041)	(0.028)	(0.028)
D2 ×	0.394***	0.393***	0.325***	0.326***
Threshold lowered	(0.050)	(0.050)	(0.031)	(0.031)
D3 ×	0.303***	0.411***	0.265***	0.323***
Threshold lowered	(0.066)	(0.104)	(0.043)	(0.055)
SR mostly		0.034		-0.037
		(0.127)		(0.080)
GR 1 mostly		0.243**		0.100**
		(0.106)		(0.048)
GR 2–3 mostly		-0.069		-0.034
		(0.054)		(0.031)
SR mostly ×		-0.182**		-0.017
Threshold lowered		(0.085)		(0.050)
GR 1 mostly ×		-0.198*		-0.049
Threshold lowered		(0.114)		(0.050)
GR 2–3 mostly		0.071		0.044
Threshold lowered		(0.075)		(0.037)
D3 × SR mostly ×		-0.347***		-0.207***
Threshold lowered		(0.097)		(0.050)
D3 × GR 1 mostly ×		0.097		0.055
Threshold lowered		(0.121)		(0.062)
D3 × GR 2–3 mostly ×		0.035		0.022
Threshold lowered		(0.105)		(0.054)
Clicks before	0.097***	0.097***	0.095***	0.095***
	(0.003)	(0.003)	(0.003)	(0.003)
Moscow	0.182**	0.117	0.098*	0.116**
	(0.085)	(0.080)	(0.059)	(0.052)
Saturday	-2.265***	-2.264***	-1.834***	-1.834***
	(0.054)	(0.055)	(0.026)	(0.026)
Sunday	-1.439***	-1.440***	-1.629***	-1.629***
	(0.042)	(0.044)	(0.024)	(0.024)
Constant	1.015***	1.063***	0.591***	0.590***
	(0.102)	(0.083)	(0.057)	(0.048)
Number of campaign-days	887,742	887,742	1,485,911	1,485,911

***, **, and * denote significance at the 1, 5, and 10 percent levels respectively