

---

# Optimal Bidding Strategies for Online Ad Auctions with Overlapping Targeting Criteria

---

**Erik Tillberg**  
Department of Computer Science  
University of Toronto

**Peter Marbach**  
Department of Computer Science  
University of Toronto

**Ravi Mazumder**  
Department of Electrical and Computer Engineering  
University of Waterloo

## Abstract

We analyze the problem of how to optimally bid for ad spaces in online ad auctions. For this we consider the general case of multiple ad campaigns with overlapping targeting criteria. In our analysis we first characterize the structure of an optimal bidding strategy. In particular, we show that an optimal bidding strategies decomposes the problem into disjoint sets of campaigns and targeting groups. In addition, we show that pure bidding strategies that use only a single bid value for each campaign are not optimal when the supply curves are not continuous. For this case, we derive a lower-bound on the optimal cost of any bidding strategy, as well as mixed bidding strategies that either achieve the lower-bound, or can get arbitrarily close to it.

## 1 Introduction

Online advertising has grown dramatically in recent years and is by now a massive market. To service the demand, ad exchanges have been set up in several places such as New York and Toronto. Ad exchanges allow advertisers to target their ad campaigns towards specific demographic groups. This leads to a new set of problems for bidding on ad spaces where specific targeting criteria need to be taken into account. These decision problems are challenging due the large number of ad spaces that can be used to potentially place a given ad, as well as the competition between ad campaigns that target the same ad spaces. As a result, how to optimally bid for ad spaces has been an open problem, and commercial platforms bidding for ad spaces have been relying on heuristics to make these decisions.

In this paper we address this problem. More precisely, we consider the situation where rather than directly buying ad spaces, advertisers use a broker, or a co-called demand supply platform (DSP), that acquires ad spaces on their behalf. The interaction between a DSP and advertiser is governed by a contract under which the DSP is obliged to acquire a certain number of ad spaces over a given period of time, where each ad space satisfies the targeting criteria. We refer to this contract as an ad campaign. We assume that a DSP represents one or more campaign of one or more advertiser. Each campaign has its own targeting criteria, as well as a required number of ad spaces that the DSP needs to obtain for the campaign.

A transaction (opportunity to buy an ad space) in an online ad market unfolds as follows. The transactions begins with an end user visiting a web page or mobile app. This visit triggers a bid request that is sent to the ad exchange. The bid request contains information about the ad space, as well as about the end user that has accessed the web page or mobile ad. The ad exchange then forwards the bid request to the demand side platforms (DSPs). A DSP that receives the bid request

can then decide whether or not to bid for the ad space by submitting a bid for the ad space. The ad exchange collects all the bids submitted by DSPs for the ad space, and uses a second-price auction to decide which of the competing DSPs will obtain the ad space. The DSP that wins the auction can display an ad to the end user on that ad space. Once an ad has been displayed, and can be viewed by the end user, this ad space is denoted, and counts, as an impression.

In this paper we analyze how a DSP should bid on a bid requests from the ad-exchange in order to obtain the required number of impressions for the ad campaigns it represents at the lowest possible cost. We formally model this decision problem as an optimization (cost minimization) problem. To do that, we introduce the concept of targeting groups that groups bid requests from ad campaigns based on the bid request types that match the targeting criteria of ad campaigns. In addition, we associate with each targeting group a supply curve that characterizes the number (volume) of impressions the DSP obtains during a given decision as a function of the bid value.

The main contributions of the paper are:

- a) we provide a formal analysis of how a DSP optimal bids for ad spaces in an online ad market for the case where a DSP has several campaigns, and the campaigns can have overlapping targeting criteria. In addition, we do not assume that supply curves to be convex and continuous functions, but only assume them to be right-continuous and non-decreasing functions.
- b) we show that an optimal bidding strategies decomposes the problem into disjoint sets of campaigns and targeting groups, where an optimal bidding strategy can be computed for each component individually.
- c) we show that pure bidding strategies that use only a single bid value for each campaign and targeting group pair are not optimal when the supply curves are not continuous. For this case we derive a lower-bound on the optimal cost of any bidding strategy, and provide mixed bidding strategies that either achieve the lower-bound, or can get arbitrarily close to it.

We note that discontinuous supply curves are the norm in practical applications (see Fig. 1 for an illustration). As a result, the problem considered in this paper can not be solved using standard techniques from optimization theory to compute an optimal bidding strategy, which makes the analysis challenging. In the paper we provide a solution approach and technique that can be used to analyze supply curves that are right-continuous, non-decreasing functions, and thus not necessarily continuous or convex functions.

The rest of the paper is organized as follows. In Section 2 we provide an overview of related work. In Section 3 we describe the second price auction that is used in online ad markets. In Section 4 we present the optimization problem to that we use to model the decision problem of the DSP. In Section 5 we consider the special case of a single campaign with a single targeting group in order to illustrate the main intuition behind our analysis and results. In Section 6, we show that when the supply curves are given by continuous functions, then an optimal bidding strategy decomposes the problem into disjoint sets of campaigns and targeting groups, where an optimal bidding strategy can be computed for each component individually. In Section 7 we present the algorithm to compute an bidding strategy, and show in Section 8 that this bidding strategy is optimal when the supply curves are given by continuous functions. In addition, we derive in Section 8 a lower-bound on the optimal cost of any bidding strategy, and provide mixed bidding strategies that either achieve the lower-bound, or can get arbitrarily close to it. In Section 9 we discuss how the results obtained in the paper can be applied to practical systems.

## 2 Related Work

In this section we provide an overview of the literature related to optimal bidding algorithms for DSPs. We first consider the literature that is directly related to the problem consider in this paper, i.e. existing literature on optimal bidding algorithms for DSPs. Roughly, the existing literature on the DSP bidding problem can be grouped into the following four categories: 1) single campaigns, 2) many campaigns with identical targeting criteria, 3) many campaigns with different but non-overlapping targeting criteria, and 4) many campaigns with overlapping targeting criteria. In this

paper we consider the last case, i.e. the general case where a DSP represents many campaigns that have overlapping targeting criteria.

*The single campaign case:* There exist several results for the basic case of optimal bidding algorithms where a DSP represents a single ad campaign [5, 4, 11], and the goal of the DSP is to maximize the utility (valuation) of the obtained ad spaces given a budget constraint. Compared with our analysis, this literature does not consider the case where the supply curves are not necessarily differentiable, or continuous.

*Many Campaigns with Identical Targeting Criteria:* The first generalization of the single campaign case is to consider the case where a DSP represents many campaigns, but all campaigns have an identical targeting criteria. This case has been studied in [10] where the authors propose a bidding strategy that maximizes revenue subject to a global budget and risk constraint; risk being the variance of revenue. The resulting algorithm decides on the optimal bid value, as well as the probabilities with which bid requests are assigned to a given campaign. This work considers a different problem (with a budget and risk constraint rather than a constraint on the number of impressions that a campaign needs to obtain) than the one considered in this paper, and the bidding strategies of [10] can not be applied, and would not be optimal, for the problem considered in this paper.

*Many Campaigns with Non-Overlapping Targeting Criteria:* In [12], the authors consider the case where a DSP represents many campaigns and the campaigns have different, but disjoint, targeting criteria. For this case, [12] considers the problem of maximizing the total utility (valuation) over all campaigns subject to a global budget constraint. The decision variables of the corresponding optimization problem are the amount of total budget that is allocated to each campaign, and the price that is used for to bid for ad spaces for the individual campaigns. The work in [12] derives an optimal bidding strategy for this case assuming that the supply curves (distribution function of the winning price for a given targeting criteria) are convex. This allows the problem to be formulated as a convex optimization problem, and can be solved using standard techniques from convex optimization. In addition, the analysis in [12] does not provide the structural properties of an optimal solution that we present in Section 8 that shows that an optimal allocation decomposes the problem into disjoint components that can be solved individually, and that pure strategies are not optimal when the supply curves are not continuous (which is the case for the supply curves obtained in practice).

*Many Campaigns with Overlapping Targeting Criteria:* The most general case, and the case that we consider in this paper, is where a DSP represents many campaigns that can have overlapping targeting criteria. Obtaining an optimal bidding strategy for this case has been an open problem, and so far only heuristic strategies have been considered for this case [7, 9]. In [7] the authors propose a heuristic algorithm to compute bidding values, and allocate bid requests to campaign, for the case where the DSP aims at maximizing the number of clicks that are received under a given budget constraint. The proposed algorithm uses a linear (heuristic) bidding function that is proportional to the predicted click through rate of a campaign for a given bid request, and allocates the bid request to the campaign with the highest bid value. In essence, the proposed algorithm allocates bid requests to the campaign that has the highest click-through rate. In [9], the authors consider the special case where two campaigns either have identical targeting criteria (fully overlap), or have disjoint criteria (do not overlap). For this case, the authors propose a heuristic algorithm that uses linear regression to forecast the cost and number of clicks received. This forecast function is then used to formulate the decision problem of a DSP as an optimization problem that can be solved.

A large fraction of the literature on optimal bidding strategies for DSPs considers the case where a DSP aims at acquiring a many ad spaces as possible under a given budget constraint, or at maximizing the utility of the obtained ad spaces under a given budget constraint. In this paper, we consider a slightly different objective where the DSP needs to acquire a predefined number of ad spaces for each campaign, and wants to achieve this at the lowest possible cost. The latter objective reflects that scenario that many DSPs face in practice.

Next we highlight existing literature that is related to the general problem of online ad markets, but considers a fundamentally different problem compared with the one consider in this paper.

In [8], the authors study the interaction between several advertisers/DSPs that are competing for an identical set of bid requests, i.e. each advertiser/DSP has single campaign and the campaigns (from the different advertisers/DSPs) have all the same targeting criteria. Using the game-theoretic

approach, the authors derive a dynamic bidding strategy (for each advertiser/DSP) and show that the bid values under this strategy converges to a Nash equilibrium.

In [1], the authors study online ad markets from the perspective of the ad publishers (owners of the web pages or mobile ads where where the ad spaces are located), rather than the DSP. In particular, [1] considers the situation where a publishers can choose between selling ad spaces either on contract-based market, or on a spot market (ad exchange). A key difference between this problem, and the problem considered in this paper, is that the total number of ad spaces that are available (at the publisher) are given (as part of the problem formulation), and do not depend on a decision variable (bid price) of the optimization problem.

There is a large body of existing literature on the problem where a centralized platform collects available ad spaces, as well as the preferences of advertisers regarding the ad spaces, and the objective of the platform is to match ad spaces with advertisers with respect to a given criteria (see [6] for an overview). This problem corresponds to a bi-partite matching problems that is NP (see discussion in [6]). As a result this body of research focuses on obtaining efficient approximate algorithm for the corresponding bi-partite matching problems. We note that this problem is fundamentally different from the problem considered in this paper.

### 3 Second Price Auction

Before we define in Section 4 the model that we use to study the optimal decisions by a DSP in an online ad market, we describe in this section the second price, sealed bid auction, that is used by ad exchanges use to decide which DSP wins the ad spot for a given bid request.

Suppose that there are  $N$  DSPs submitting a bid for a given bid request, and let  $(b_1, \dots, b_N)$  be the bids that are received by the ad exchange and assume that

$$b_i \neq b_j, \quad i \neq j, i, j = 1, \dots, N.$$

Then the DSP  $i^* \in \{1, \dots, N\}$  with the highest bid  $b_{i^*}$ , i.e. we have that

$$b_{i^*} \geq b_i, \quad i = 1, \dots, N,$$

will win the auction and receive the ad spot of the bid request. The price that DSP  $i^*$  pays for the ad spot is equal to the value of the second highest bid in  $(b_1, \dots, b_N)$ , i.e. the price  $p$  is equal to  $b_s \in \{1, \dots, N\}$  such that

$$b_s \geq b_i, \quad i \in \{1, \dots, N\} \setminus \{i^*\}.$$

We ignore in this paper the floor price of an auction; a floor price can easily be introduced in our analysis without changing the results.

Whenever a DSP wins an ad space for a campaign that it represents, we say that the DSP has received an impression (ad space) for the campaign.

## 4 Problem Definition

Recall that we consider the problem of how a DSP should bid for ad spaces in an online ad market. We assume that the DSP represents one or more ad campaigns. Each campaign is given by a contract that specifies the targeting criteria for the campaign, and the number of impressions the campaign needs to obtain over a given time period (for example, a month). Below we describe the model that we use to analyze this problem.

### 4.1 Decision Period

We focus in our analysis on the case where the DSP is required to spread out the total number of impression over the given time period, and needs to obtain an equal number of impressions in during a number of pre-specified smaller time intervals (for example, a day). In addition, we allow the DSP to break the time intervals over which a required number of impressions needs to be obtained into smaller decision periods. A the length of a decision period could relatively long such as an hour, or shorter such as a few minutes. In the following, we analyze how a DSP optimally decides on the bid values and bid request allocations for a given decision period. To do that, we formally model the decision problem as an optimization problem

## 4.2 Campaigns

Let  $\mathcal{N} = \{1, \dots, N_c\}$  be the set of campaigns that the DSP represents. For each campaign  $i \in \mathcal{N}$ , let  $I_i$  be the total number of impressions that campaign  $i$  is required (scheduled) to obtain during the given decision period.

We assume that each bid request  $r$  that the DSP receives from the ad exchange, is characterized by its properties  $t(r)$ . Examples of a bid request properties could be the location, age, and gender, of the end user who accessed the web page or mobile app and triggered the bid request. We refer to  $t(r)$  as the type of the bid request. The type  $t(r)$  determines whether bid request  $r$  matches the targeting criteria of a given campaign  $i \in \mathcal{N}$ . More precisely, let  $\mathcal{R}$  be the set of all possible bid request types, and let  $\mathcal{R}_i \subseteq \mathcal{R}$  be the set of all bid request types that match the targeting criteria of campaign  $i$ . For example, the targeting criteria of an a campaign could be to show ads to users located in New York City aged between 30-50 years.

If for a given bid request  $r$  matches the targeting criteria of campaign  $i$  and we have that  $t(r) \in \mathcal{R}_i$ , then we say that  $r$  is an admissible bid request for campaign  $i$ .

## 4.3 Targeting Groups

The targeting criteria  $\mathcal{R}_i$  of the campaigns  $i \in \mathcal{N}$  are not necessarily disjoint sets. Moreover, the targeting criteria  $\mathcal{R}_i$  and  $\mathcal{R}_j$  of two ad campaigns  $i$  and  $j$  that overlap, may not fully overlap. As a result the set of targeting criteria  $\{\mathcal{R}_i\}_{i \in \mathcal{N}}$  has very little structure, which makes it difficult to use them directly to formulate an optimization problem for the decision problem of the DSP. To overcome this problem, instead of using the targeting criteria directly, we use the set of targeting criteria  $\mathcal{R}_i$ ,  $i \in \mathcal{N}$ , to construct a set  $\mathcal{T} = \{1, \dots, N_r\}$  of  $N_r$  targeting groups with the following structure.

**Definition 1.** A set of targeting groups  $\mathcal{T}$  is given by  $N_r$  sets  $\mathcal{T}_j \subset \mathcal{R}$ , such that

$$\bigcup_{j=1}^{N_r} \mathcal{T}_j = \bigcup_{i=1}^{N_c} \mathcal{R}_i$$

and

$$\mathcal{T}_m \cap \mathcal{T}_n = \emptyset, \quad m \neq n, m, n = 1, \dots, N_r.$$

Furthermore, for every campaign  $i$  there exists a set  $\mathcal{A}_i \subseteq \{1, \dots, N_r\}$  of targeting groups such that

$$\mathcal{R}_i = \bigcup_{j \in \mathcal{A}_i} \mathcal{T}_j$$

and

$$\mathcal{R}_i \cap \mathcal{T}_j = \mathcal{T}_j, \quad j \in \mathcal{A}_i.$$

For given a targeting group  $j$ , the set  $\mathcal{B}_j$  is given by

$$\mathcal{T}_j \cap \mathcal{R}_i = \mathcal{T}_j.$$

Note that as the number of campaigns is finite, a set of targeting groups as given by Definition 1 can always be constructed. Below we discuss properties of targeting groups as given by Definition 1 that play an important role in our analysis.

The set  $\mathcal{A}_i$ ,  $i \in \mathcal{N}$ , in Definition 1 consists of all the targeting groups that campaign  $i$  can bid on. That is, given a bid request  $r$  such that  $t(r) \in \mathcal{R}_i$ , we have that there exists a  $j \in \mathcal{A}_i$  such that  $t(r) \in \mathcal{T}_j$ . Furthermore as the targeting groups are given by disjoint sets, for given a bid request  $r$  such that  $t(r) \in \mathcal{R}_i$  there exists exactly one  $j \in \mathcal{A}_i$  such that  $t(r) \in \mathcal{T}_j$ .

Similarly, the set  $\mathcal{B}_j$ ,  $j \in \mathcal{T}$ , consists of all of the campaigns that can bid on a request  $r$  such that  $t(r) \in \mathcal{T}_j$  and  $t_r \in \mathcal{R}_i$ . That is, it consists of all the campaigns that can bid on the set of targets in  $\mathcal{T}_j$ .

The key property of the targeting groups  $\mathcal{T}$  is that the sets  $\mathcal{T}_j$ ,  $j \in \mathcal{T}$ , are disjoint, and hence each type  $t(r)$  of a bid request  $r$  maps to exactly one targeting group.

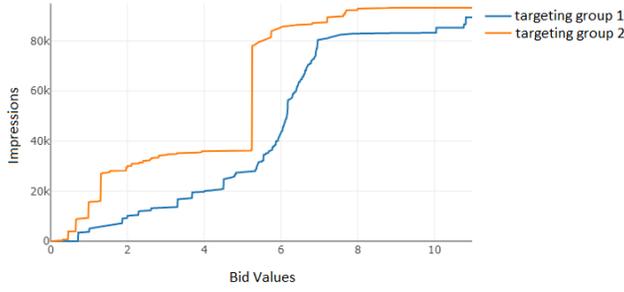


Figure 1: Two examples of supply obtained from real-life data.

#### 4.4 Supply Curves

Given the set  $\mathcal{T}$  of targeting groups, we associate with each targeting group  $j \in \mathcal{T}$  a supply curve  $D_j(b)$ ,  $b \geq 0$ . The supply curve  $D_j(b)$  has the following interpretation. If during the decision period the DSP bids with a fixed bid value  $b_j$  for all bid requests that match targeting group  $\mathcal{T}_j$ ,  $j \in \mathcal{T}$ , then  $D_j(b_j)$  is the total number of impressions that the DSP obtains from targeting group  $j$  over the decision period. Note that by definition, the supply curves  $D_j(b)$ ,  $b \geq 0$ , are non-decreasing function. Fig. 1 shows two supply curves that have been obtained from actual data from an online ad exchange. From Fig. 1 we have that the supply curves are not necessarily given by differentiable, or continuous, functions, and we make the following assumption regarding the supply curves  $D_j(b)$ ,  $j \in \mathcal{T}$ .

**Assumption 1.** For each targeting group  $j \in \mathcal{T}$  the supply curve  $D_j(b) : [0, \infty) \mapsto [0, \infty)$  is a non-decreasing right-continuous function.

Using this definition of a supply curve, we have that if the DSP bids on all bid requests that match a given targeting group  $j \in \mathcal{T}$  with a fixed bid  $b_j$ , then the DSP will obtain  $D_j(b_j)$  impressions for targeting group  $j$  with a cost of (see for example [3])

$$\int_0^{b_j} b dD_j(b) = D_j(b_j)b_j - \int_0^{b_j} D_j(b)db.$$

#### 4.5 Decision Variables

For each targeting group  $j \in \mathcal{T}$  the DSP has to decide on

- the fraction  $\gamma_{i,j}$ ,  $i \in \mathcal{B}_j$ , of bid requests from the targeting group  $j$  that are allocated to campaign  $i$  to bid on, and
- the bid value  $b_{i,j}$  that it uses for bidding on requests in targeting group  $j$ ,  $j \in \mathcal{A}_i$ , that are allocated to campaign  $i$ .

We then use the following notation for the decision variables of the DSP during a given decision period. Let

$$\gamma_j = \{\gamma_{i,j}\}_{i \in \mathcal{B}_j}, \quad i = 1, \dots, N_r,$$

be the fraction of bid request from targeting group  $j$  that are allocated to campaign  $i$ , and let

$$\mathbf{b}_j = \{b_{i,j}\}_{j \in \mathcal{A}_i}, \quad i = 1, \dots, N_c,$$

be the bid values that campaign  $i$  uses for bid request from targeting group  $j$ .

Furthermore, let

$$\mathbf{b} = (\mathbf{b}_1, \dots, \mathbf{b}_{N_c}), \text{ and } \gamma = (\gamma_1, \dots, \gamma_{N_r}),$$

be the set of decision variables for all campaigns, and all targeting groups, and let  $(\mathbf{b}, \gamma)$  denote the corresponding allocation of the DSP for a given decision period.

If the DSP allocates for a given decision a fraction  $\gamma_{i,j}$  of the bid requests of targeting group  $j$  to campaign  $i$ , then the total number of impressions that campaign  $i$  will obtain from targeting group  $j$  is given by  $\gamma_{i,j}D_j(b_{i,j})$ , and the total cost for obtaining these impressions is given by

$$\gamma_{i,j} \left[ D_j(b_{i,j}) - \int_0^{b_{i,j}} D_j(b) db \right].$$

#### 4.6 Optimization Problem

Recall that for a given decision period the goal of the DSP is to obtain the required number of impressions for each campaign  $i \in \mathcal{N}$  at the lowest possible cost. Using the above notations and definitions, the corresponding optimization problem is given by

$$\begin{aligned} & \underset{(\mathbf{b}, \boldsymbol{\gamma})}{\text{minimize}} && \sum_{i=1}^{N_c} \left[ \sum_{j=1}^{N_r} \gamma_{i,j} \left( D_j(b_{i,j}) b_{i,j} - \int_0^{b_{i,j}} D_j(b) db \right) \right] \\ & \text{subject to} && \sum_{j \in \mathcal{A}_i} \gamma_{i,j} D_j(b_{i,j}) \geq I_i, \quad i = 1, \dots, N_c, \\ & && \sum_{i \in \mathcal{B}_j} \gamma_{i,j} \leq 1, \quad j = 1, \dots, N_r, \\ & && b_{i,j} \geq 0, \quad i = 1, \dots, N_c, j = 1, \dots, N_r, \\ & && \gamma_{i,j} \geq 0, \quad i = 1, \dots, N_c, j = 1, \dots, N_r. \end{aligned} \tag{1}$$

The first inequality requires the number of impressions obtained for each campaign to meet the required number of impressions. The second inequality states that the total fraction of bid request that are allocated to campaigns can not exceed 1. The last two constraints state that the bid values and allocations can not be negative.

In the following we refer to an allocation  $(\mathbf{b}, \boldsymbol{\gamma})$  that satisfies the above conditions as a solution to the optimization problem given by Eq. (1)

For our analysis we assume that the DSP can always get the required impressions by using high enough bid values  $b_{i,j}$ ,  $i \in \mathcal{N}$ ,  $j \in \mathcal{T}$ , and thus the optimization problem has a feasible solution.

#### 4.7 Online Bidding Algorithm

In this section discuss how a solution  $(\mathbf{b}, \boldsymbol{\gamma})$  of the optimization problem given by Eq. (1) can be used to implement an online bidding algorithm.

Let  $(\mathbf{b}, \boldsymbol{\gamma})$  be a solution for the optimization problem given by Eq. (1). The solution  $(\mathbf{b}, \boldsymbol{\gamma})$  can then be used to implement an online algorithm as follows. Whenever a new bid request  $r$  is received by the DSP, the DSP first determines the type  $t(r)$  of the bid request. Next, the DSP determines the targeting group the type  $t(r)$  belongs to. Let  $j$  be this targeting group. Using the solution  $(\mathbf{b}, \boldsymbol{\gamma})$ , the DSP then allocates with probability  $\gamma_{i,j}$  the bid request to campaign  $i$ ,  $i \in \mathcal{B}_j$ ; and uses the bid value  $b_{i,j}$  to obtain the ad space for campaign  $i$ .

Note that the solution  $(\mathbf{b}, \boldsymbol{\gamma})$  is (pre-)computed in the background (off-line), and the values of the solution  $(\mathbf{b}, \boldsymbol{\gamma})$  are used to implement the online bidding algorithm described above. The length of a decision period determines how often the solution  $(\mathbf{b}, \boldsymbol{\gamma})$  is re-computed, and the bid values and allocations for the online algorithm are updated. In Section 9 we discuss additional implementation issues of the online algorithm.

## 5 Single Campaign with a Single Targeting Group

Before we start with the analysis of an optimal allocation for the optimization problem given by Eq. (1), we provide in this section the main intuition behind our analysis and results. For this we consider the special case where a single campaign  $i = 1$  gets a total of  $I_1$  impressions from a single targeting group  $j = 1$ . We assume that

$$\lim_{b \rightarrow \infty} D_1(b) > I_1,$$

and there is sufficient supply to support the impressions of the campaign.

Our first result characterizes the optimal allocation  $(\mathbf{b}^*, \gamma^*)$  for the case where there exists a bid value  $b^*$  such that  $D_1(b^*) = I_1$ .

**Proposition 1.** *Consider the special case where a single campaign  $i = 1$  gets impressions from a single targeting group  $j = 1$ , and  $I_1$  is the total number of impressions that the campaign is required to obtain. If there exists a bid value  $b^*$  such that*

$$D_1(b^*) = I_1,$$

then the allocation  $(\mathbf{b}^*, \gamma^*)$  given by

$$\gamma_{1,1}^* = 1 \quad \text{and} \quad b_{1,1}^* = b^*$$

is an optimal allocation. Furthermore, the optimal cost is given by

$$I_1 b^* - \int_0^{b^*} D_1(b) db.$$

We provide a proof of Proposition 1 in Appendix B.1. Proposition 1 implies that if the supply curve  $D_1(b)$  is given by a continuous function, then the allocation  $(\mathbf{b}^*, \gamma^*)$  given in Proposition 1 is an optimal allocation. To see this, note that if the supply curve  $D_1(b)$  is given by a continuous function, then there exists a bid value  $b^*$  such that

$$D_1(b^*) = I_1.$$

If the the supply curve  $D_1(b)$  is not continuous, and there does not exist a bid value  $b^*$  such that  $D_1(b^*) = I_1$ , then we define the bid value  $b^*$  such that

$$D_1(b^*) \geq I_1 \tag{2}$$

and

$$D_1(b) < I_1, \quad 0 \leq b < b^*. \tag{3}$$

In the next subsection we study optimal allocations for this case. To do that we use the following notation. We let

$$D^{(-)}(b^*) = \lim_{b \uparrow b^*} D_1(b) < D_1(b^*).$$

## 5.1 Pure Bidding Strategies

Note that the optimization problem given by Eq. (1) considers allocations under which each campaign  $i \in \mathcal{N}$  uses (as most) a single bid value  $b_{i,j}$  for targeting group  $j \in \mathcal{A}_i$ . We refer such an allocation as a pure bidding strategy. Let  $\mathcal{S}_p$  be the set of all pure bidding strategies. We then have the following result.

**Lemma 1.** *Consider the special case where a single campaign  $i = 1$  gets impressions from a single targeting group  $j = 1$ , and  $I_1$  is the total number of impressions that the campaign is required to obtain. Furthermore, let the bid value  $b^*$  be as given by Eq. (2) and (3). Then the pure bidding strategy given by*

$$b_{1,1}^* = b^* \quad \text{and} \quad \gamma_{1,1}^* = \frac{I_1}{D_1(b^*)}$$

is an optimal pure bidding strategy. Furthermore the cost under the optimal pure strategy is given by

$$I_1 b^* - \gamma_{1,1}^* \int_0^{b^*} D_1(b) db.$$

The result of Lemma 1 follow from the fact that the bid value  $b^*$  be as given by Eq. (2) and (3) is the lowest bid value to obtain the  $I_1$  impressions required by the campaign. Note that if we have that  $D_j(b^*) = I_1$ , then the result of Lemma 1 is identical to the result of Proposition 1.

## 5.2 Mixed Bidding Strategies

We next consider a more general class of bidding strategies, where a campaign can use several bid values for a given targeting group. We refer to such strategies as mixed bidding strategies. For a given mixed bidding strategy  $s$ , let  $\mathcal{P}_{1,1}$  be the bid values that the campaign 1 uses for targeting group 1, and let  $\gamma_{1,1}(b)$ ,  $b \in \mathcal{P}_{1,1}$ , be the fraction of bid requests from targeting group 1 for which the campaign uses bid value  $b$ , where we have the constraint that

$$\sum_{b \in \mathcal{P}_{1,1}} \gamma_{1,1}(b) \leq 1.$$

Let  $\mathcal{S}$  be the set of all such mixed bidding strategies.

The cost  $C(s)$  of a mixed bidding strategy  $s \in \mathcal{S}$  to obtain

$$\sum_{b \in \mathcal{P}_{1,1}} \gamma_{1,1}(b) D_1(b)$$

impressions is given by

$$C(s) = \sum_{b \in \mathcal{P}_{1,1}} \gamma_{1,1}(b) \left[ D_1(b)b - \int_0^b D_1(x) dx \right].$$

We have the following result.

**Proposition 2.** *Consider the special case where a single campaign  $i = 1$  gets impressions from a single targeting group  $j = 1$ , and  $I_1$  is the total number of impressions that the campaign is required to obtain. Let  $s \in \mathcal{S}$  given mixed bidding strategy such that*

$$\sum_{b \in \mathcal{P}_{1,1}} \gamma_{1,1}(b) D_1(b) = I_1.$$

Then we have

$$C(s) \geq I_1 b^* - \int_0^{b^*} D_1(b) db,$$

where the bid value  $b^*$  is as given by Eq. (2) and (3).

We provide a proof for Proposition 2 in Appendix B.2. Proposition 2 provides a lower-bound on the cost obtained by any mixed bidding strategy.

The next result shows that the bound of Proposition 1 is tight.

**Proposition 3.** *Consider the special case where a single campaign  $i = 1$  gets impressions from a single targeting group  $j = 1$ , and  $I_1$  is the total number of impressions that the campaign is required to obtain. Let the bid value  $b^*$  be as given by Eq. (2) and (3), and suppose that  $D^{(-)}(b^*) > 0$ . Furthermore, let the mixed bidding strategy  $s(b_1) \in \mathcal{S}$ ,  $0 \leq b_1 < b^*$ , be given by  $\mathcal{P}_1(s) = \{b_1, b^*\}$ , as well as*

$$\gamma_{1,1}(b_1) = \frac{D_1(b^*) - I_1}{D_1(b^*) - D_1(b_1)} \quad \text{and} \quad \gamma_{1,1}(b^*) = 1 - \gamma_{1,1}(b_1).$$

Then we have that

$$\lim_{b_1 \uparrow b^*} C(s(b_1)) = I_1 b^* - \int_0^{b^*} D_1(b) db.$$

We provide a proof for Proposition 3 in Appendix B.3. Proposition 3 is that it implies that it is sufficient to consider mixed bidding strategies that use two bid values  $b_1$  and  $b^*$ ,  $b_1 < b^*$ , in order to obtain a (near) optimal cost.

## 5.3 Cost Bound for Optimal Pure Strategy

The above results imply that using mixed bidding strategies can lead to a strictly lower overall cost compared with pure strategies. The following result provides a bound on this cost difference.

**Proposition 4.** Consider the special case where a single campaign  $i = 1$  gets impressions from a single targeting group  $j = 1$ , and  $I_1$  is the total number of impressions that the campaign is required to obtain. Furthermore, let the bid value  $b^*$  be as given by Eq. (2) and (3), and let  $s^* \in \mathcal{S}_p$  be the optimal pure bidding strategy as given by Lemma 1. Then we have that

$$\begin{aligned} & C(s^*) - \left[ I_1 b^* - \int_0^{b^*} D_1(b) db \right] \\ & \leq (1 - \gamma_{1,1}^*) \int_0^{b^*} D_1(b^*) db \\ & \leq \frac{D_1(b^*) - D_1^{(-)}(b^*)}{D_1(b^*)} \int_0^{b^*} D_1(b) db, \end{aligned}$$

where  $D^{(-)}(b^*) = \lim_{b \uparrow b^*} D_1(b)$ .

Proposition 4 follows from Lemma 1. Proposition 4 states that the smaller the height (jump)  $(D_1(b^*) - D_1^{(-)}(b^*))$  of the discontinuity at  $b^*$ , the closer the performance of the optimal pure strategy will be to the lower-bound of Proposition 2.

## 5.4 Discussion

The results for case where a single campaign gets impressions from a single targeting group provide important insight into the structure of an optimal bidding strategy. In particular, the results suggest that studying optimal policies under the assumption that the supply curve is continuous provides correct structure of an optimal bidding strategy. More precisely, Proposition 2 suggests that the optimal cost in the case of continuous supply curves provides the correct expression for the (tight) lower-bound of the cost of any mixed strategy for the case where the supply curves are not necessarily continuous. In addition, the optimal value  $b^*$  obtained in Proposition 1 corresponds to the optimal bid value of a pure strategy as given in Lemma 1, as well as one of the bid values of the general bid strategy given in Proposition 3.

In the following we show that the results and insights obtained for the case of a single campaign getting impressions from a single targeting group carry over to the general case where multiple campaigns getting impressions from several targeting groups. To do that, we proceed as follows. We first characterize in Section 6 the structure of an optimal allocation  $(\mathbf{b}^*, \boldsymbol{\gamma}^*)$  under the assumption that the supply curves are given by continuous function. Using this result, we propose in Section 7 an algorithm to compute an allocation  $(\mathbf{b}^*, \boldsymbol{\gamma}^*)$ . In Section 8 we show that this allocation is an optimal bidding strategy when the supply curves are given by continuous functions. In addition, we derive for this case a lower-bound on the optimal cost of any mixed bidding strategy, as well as mixed bidding strategies that either achieve the lower-bound, or can get arbitrarily close to it.

## 6 Structural Properties

In this section we characterize an optimal solution  $(\mathbf{b}^*, \boldsymbol{\gamma}^*)$  for the optimization problem given by Eq. (1) under the simplifying assumption that the supply curves  $D_j(b)$ ,  $j \in \mathcal{T}$ , are given by continuous functions.

To do this, we use the following notation. Let  $(\mathbf{b}^*, \boldsymbol{\gamma}^*)$  be a solution to the optimization given by Eq. (1). For each campaign  $i \in \mathcal{N}$  let the set  $\mathcal{A}_i(\mathbf{b}^*, \boldsymbol{\gamma}^*)$  be given by

$$\mathcal{A}_i(\mathbf{b}^*, \boldsymbol{\gamma}^*) = \{j \in \mathcal{A}_i | \gamma_{i,j}^* D_j(b_{i,j}^*) > 0\}.$$

The set  $\mathcal{A}_i(\mathbf{b}^*, \boldsymbol{\gamma}^*)$  contains all the targeting groups  $j$  from which a campaign receives a positive number  $\gamma_{i,j}^* D_j(b_{i,j}^*)$  of impressions under the solution  $(\mathbf{b}^*, \boldsymbol{\gamma}^*)$ .

Similarly, for each targeting group  $j \in \mathcal{T}$  let the set  $\mathcal{B}_j(\mathbf{b}^*, \boldsymbol{\gamma}^*)$  be given by

$$\mathcal{B}_j(\mathbf{b}^*, \boldsymbol{\gamma}^*) = \{i \in \mathcal{B}_j | \gamma_{i,j}^* D_j(b_{i,j}^*) > 0\}.$$

When we have the result that there always exists an optimal solution  $(\mathbf{b}^*, \gamma^*)$  for the optimization problem given by Eq. (1) that consists of components

$$\mathcal{C}(p^*, \mathbf{b}^*, \gamma^*) = (\mathcal{C}_c(p^*, \mathbf{b}^*, \gamma^*), \mathcal{C}_t(p^*, \mathbf{b}^*, \gamma^*))$$

where  $\mathcal{C}_c(p^*, \mathbf{b}^*, \gamma^*)$  is a set of campaigns, and  $\mathcal{C}_t(p^*, \mathbf{b}^*, \gamma^*)$  is a set of targeting groups, such that for each campaign  $i \in \mathcal{C}_c(p^*, \mathbf{b}^*, \gamma^*)$  we have that

$$b_{i,j}^* = p^*, \quad j \in \mathcal{A}_i(\mathbf{b}^*, \gamma^*).$$

**Proposition 5.** *Suppose that the supply curves  $D_j(b)$ ,  $j \in \mathcal{T}$ , are given by continuous functions. Then there exists an optimal solution  $(\mathbf{b}^*, \gamma^*)$  to the optimization problem given by Eq. (1) such that the following is true. There exists a set of bid values  $\mathcal{P}(\mathbf{b}^*, \gamma^*)$  such that with every bid value  $p^* \in \mathcal{P}(\mathbf{b}^*, \gamma^*)$  we can associate a component*

$$\mathcal{C}(p^*, \mathbf{b}^*, \gamma^*) = (\mathcal{C}_c(p^*, \mathbf{b}^*, \gamma^*), \mathcal{C}_t(p^*, \mathbf{b}^*, \gamma^*))$$

where  $\mathcal{C}_c(p^*, \mathbf{b}^*, \gamma^*)$  is a set of campaigns, and  $\mathcal{C}_t(p^*, \mathbf{b}^*, \gamma^*)$  is a set of targeting groups, such that

$$\mathcal{A}_i(\mathbf{b}^*, \gamma^*) \cap \mathcal{C}_t(p^*, \mathbf{b}^*, \gamma^*) = \mathcal{A}_i(\mathbf{b}^*, \gamma^*), \quad i \in \mathcal{C}_c(p^*, \mathbf{b}^*, \gamma^*),$$

and

$$\mathcal{B}_j(\mathbf{b}^*, \gamma^*) \cap \mathcal{C}_c(p^*, \mathbf{b}^*, \gamma^*) = \mathcal{B}_j(\mathbf{b}^*, \gamma^*), \quad j \in \mathcal{C}_t(p^*, \mathbf{b}^*, \gamma^*),$$

as well as

$$b_{i,j}^* = p^*, \quad i \in \mathcal{C}_c(p^*, \mathbf{b}^*, \gamma^*), j \in \mathcal{A}_i(\mathbf{b}^*, \gamma^*).$$

In addition, we have that

$$\bigcup_{p^* \in \mathcal{P}(\mathbf{b}^*, \gamma^*)} \mathcal{C}_c(p^*, \mathbf{b}^*, \gamma^*) = \mathcal{N}$$

and

$$\bigcup_{p^* \in \mathcal{P}(\mathbf{b}^*, \gamma^*)} \mathcal{C}_t(p^*, \mathbf{b}^*, \gamma^*) = \mathcal{T}.$$

## 7 Algorithm to Compute $(\mathbf{b}^*, \gamma^*)$

Using the results of Section 6, we present in this section an algorithm to compute a solution  $(\mathbf{b}^*, \gamma^*)$  for the optimization problem given by Eq. (1). In the next section, we show that the solution  $(\mathbf{b}^*, \gamma^*)$  is an optimal solution for the case where the supply curves are given by continuous functions. In addition, we derive a lower-bound on the optimal cost of any bidding strategy, as well as mixed bidding strategies that either achieve the lower-bound, or can get arbitrarily close to it.

### 7.1 Algorithm

Using the result of Proposition 5, we derive an algorithm that computes an allocation  $(\mathbf{b}^*, \gamma^*)$  that consists of components as given in Proposition 5. To do that, the algorithm recursively computes components  $(\hat{\mathcal{C}}_c, \hat{\mathcal{C}}_t, p, \mathbf{t})$ , where each component consists of a set of campaigns  $\hat{\mathcal{C}}_c \subseteq \mathcal{N}$  and a set of targeting groups  $\hat{\mathcal{C}}_t \subseteq \mathcal{T}$ . Using these components, we then construct an allocation  $(\mathbf{b}^*, \gamma^*)$  as given in Proposition 5.

More precisely, let

$$\mathcal{C}_c = \mathcal{N} \text{ and } \mathcal{C}_t = \mathcal{T}$$

be the initial set of campaigns and targeting groups, and let

$$\mathcal{P} = \emptyset.$$

We then split the set  $(\mathcal{C}_c, \mathcal{C}_t) = (\mathcal{N}, \mathcal{T})$  into two subsets  $(\mathcal{C}_c^1, \mathcal{C}_t^1)$  and  $(\mathcal{C}_c^2, \mathcal{C}_t^2)$  as follows.

We first compute the unique bid value  $p$  such that

$$\sum_{j \in \mathcal{C}_t} D_j(p) \geq \sum_{i \in \mathcal{C}_c} I_i \quad (4)$$

and

$$\sum_{j \in \mathcal{C}_t} D_j(b) < \sum_{i \in \mathcal{C}_c} I_i, \quad 0 \leq b < p. \quad (5)$$

Using the bid value  $p$ , we next compute an optimal solution to the optimization problem  $ALLOC(\mathcal{C}_c, \mathcal{C}_t, p)$  given by

$$\begin{aligned} & \underset{\mathbf{t}}{\text{minimize}} && \sum_{i \in \mathcal{C}_c} \left( I_i - \sum_{j \in \mathcal{A}_i \cap \mathcal{C}_t} t_{i,j} D_j(p) \right)^2 \\ & \text{subject to} && \sum_{i \in \mathcal{B}_j \cap \mathcal{C}_c} t_{i,j} \leq 1, \quad j \in \mathcal{C}_t, \\ & && 0 \leq t_{i,j}, \quad i \in \mathcal{C}_c, j \in \mathcal{C}_t. \end{aligned} \quad (6)$$

Note that the optimization problem  $ALLOC(\mathcal{C}_c, \mathcal{C}_t, p)$  is a convex (quadratic) optimization problem, and can easily be solved using standard algorithms (solvers) of quadratic programming.

If the minimization problem given by Eq. (6) has an optimal solution  $\mathbf{t}^*$  with cost zero, then we stop the recursion, and return the component  $(\hat{\mathcal{C}}_c, \hat{\mathcal{C}}_t, p, \mathbf{t})$  where

$$\mathbf{t} = \mathbf{t}^*$$

is equal to the optimal solution  $\mathbf{t}^*$  for the optimization problem  $ALLOC(\mathcal{C}_c, \mathcal{C}_t, p)$ , and  $p$  is the bid value that we used to in the optimization problem  $ALLOC(\mathcal{C}_c, \mathcal{C}_t, p)$ . In addition, we add the bid value  $p$  to the set  $\mathcal{P}$  and let

$$\mathcal{P} = \mathcal{P} \cup \{p\}.$$

On the other hand, if the minimization problem given by Eq. (6) has an optimal solution  $\mathbf{t}^*$  with a strictly positive cost, then we split the set  $(\mathcal{C}_c, \mathcal{C}_t)$  into two sets  $(\mathcal{C}_c^1, \mathcal{C}_t^1)$  and  $(\mathcal{C}_c^2, \mathcal{C}_t^2)$  as follows. We set

$$\mathcal{C}_c^{(1)} = \left\{ i \in \mathcal{C}_c \mid \left( I_i - \sum_{j \in \mathcal{A}_i \cap \mathcal{C}_t} t_{i,j}^* D_j(p) \right)^2 > 0 \right\}, \quad (7)$$

and

$$\mathcal{C}_t^{(1)} = \bigcup_{i \in \mathcal{C}_c^{(1)}} \mathcal{A}_i. \quad (8)$$

Furthermore, we set

$$\mathcal{C}_c^{(2)} = \mathcal{C}_c \setminus \mathcal{C}_c^{(1)} \text{ and } \mathcal{C}_t^{(2)} = \mathcal{C}_t \setminus \mathcal{C}_t^{(1)}. \quad (9)$$

For each of the two components we then recursively apply the above step until all recursions stop.

More formally we define the algorithm as follows. Let the function

$$p = PRICE(\mathcal{C}_c, \mathcal{C}_t)$$

return the unique price  $p^*$  as defined by Eq. (4) and Eq. (5). Furthermore let the function

$$(\mathbf{t}^*, done) = ALLOC(\mathcal{C}_c, \mathcal{C}_t, p)$$

return an optimal solution to the optimization problem given by Eq (6), as well as a boolean value  $done = TRUE$  if the optimal solution  $\mathbf{t}^*$  for the optimization problem by Eq (6) has a cost equal to 0, i.e. we have that

$$\sum_{i \in \mathcal{C}_c} \left( I_i - \sum_{j \in \mathcal{A}_i \cap \mathcal{C}_t} t_{i,j}^* D_j(p) \right)^2 = 0,$$

and a boolean value  $done = FALSE$  otherwise. Finally let the function

$$((\mathcal{C}_{c,1}, \mathcal{C}_{t,1}), (\mathcal{C}_{c,2}, \mathcal{C}_{t,2})) = SPLIT(\mathcal{C}_c, \mathcal{C}_t, \mathbf{t}^*, p)$$

return the sets  $(\mathcal{C}_{c,1}, \mathcal{C}_{t,1})$  and  $(\mathcal{C}_{c,2}, \mathcal{C}_{t,2})$ , as given by Eq. (7) - (9).

Using these functions, the pseudo-code for the recursion step

$$(\mathcal{C}_{out}, \mathcal{P}_{out}) = REC(\mathcal{C}_c, \mathcal{C}_t, \mathcal{C}_{in}, \mathcal{P}_{in})$$

---

**Algorithm 1:**  $(\mathcal{C}_{out}, \mathcal{P}_{out}) = REC(\mathcal{C}_c, \mathcal{C}_t, \mathcal{C}_{in}, \mathcal{P}_{in})$ 


---

```

1  $p = PRICE(\mathcal{C}_c, \mathcal{C}_t)$ ;
2  $(\mathbf{t}^*, done) = ALLOC(\mathcal{C}_c, \mathcal{C}_t, p)$ ;
3 if  $done$  then
4    $\mathcal{C}_{out} = \mathcal{C}_{in} \cup \{(\mathcal{C}_c, \mathcal{C}_t, \mathbf{t}^*, p)\}$ ;
5    $\mathcal{P}_{out} = \mathcal{P}_{in} \cup \{p\}$ ;
6   return  $(\mathcal{C}_{out}, \mathcal{P}_{out})$ ;
7 else
8    $((\mathcal{C}_{c,1}, \mathcal{C}_{t,1}), (\mathcal{C}_{c,2}, \mathcal{C}_{t,2})) = SPLIT(\mathcal{C}_c, \mathcal{C}_t, \mathbf{t}^*, p)$ ;
9    $(\mathcal{C}_{out,1}, \mathcal{P}_{out,1}) = REC(\mathcal{C}_{c,1}, \mathcal{C}_{t,1}, \mathcal{C}_{in}, \mathcal{P}_{in})$ ;
10   $(\mathcal{C}_{out,2}, \mathcal{P}_{out,2}) = REC(\mathcal{C}_{c,2}, \mathcal{C}_{t,2}, \mathcal{C}_{in}, \mathcal{P}_{in})$ ;
11   $\mathcal{C}_{out} = \mathcal{C}_{out,1} \cup \mathcal{C}_{out,2}$ ;
12   $\mathcal{P}_{out} = \mathcal{P}_{out,1} \cup \mathcal{P}_{out,2}$ ;
13  return  $(\mathcal{C}_{out}, \mathcal{P}_{out})$ ;
14 end

```

---

is then given by Algorithm 1. To simplify the notation, it is assumed that the set  $\mathcal{A}_i$ ,  $i \in \mathcal{N}$ , and  $\mathcal{B}_j$ ,  $j \in \mathcal{T}$ , are available to the function  $REC(\mathcal{C}_c, \mathcal{C}_t, \mathcal{C}_{in}, \mathcal{P}_{in})$  given in Algorithm 1 as global (globally defined) variables.

Using the recursive function  $REC(\mathcal{C}_c, \mathcal{C}_t, \mathcal{C}_{in}, \mathcal{P}_{in})$ , the algorithm to compute the set  $\{(\hat{\mathcal{C}}_c, \hat{\mathcal{C}}_t, p, \mathbf{t})\}_{p \in \mathcal{P}}$  of components is given by

$$(\mathcal{C}_{out}, \mathcal{P}) = REC(\mathcal{N}, \mathcal{T}, \emptyset, \emptyset) \quad (10)$$

where  $\mathcal{P}$  is the set of bid values of the components  $(\hat{\mathcal{C}}_c, \hat{\mathcal{C}}_t, p, \mathbf{t})$ , and

$$\mathcal{C}_{out} = \{(\hat{\mathcal{C}}_c, \hat{\mathcal{C}}_t, p, \mathbf{t})\}_{p \in \mathcal{P}}$$

is the set of components obtained by the algorithm.

## 7.2 Construction of Allocation $(\mathbf{b}^*, \gamma^*)$

Using the set  $\mathcal{P}$  of bid values, and the components  $(\hat{\mathcal{C}}_c, \hat{\mathcal{C}}_t, p, \mathbf{t})$ ,  $p \in \mathcal{P}$ , obtained by the algorithm given by Eq (10), we construct a solution  $(\mathbf{b}^*, \gamma^*)$  as given by Proposition 5 as follows. Let

$$\mathcal{P}(\mathbf{b}^*, \gamma^*) = \mathcal{P},$$

and for each component  $(\hat{\mathcal{C}}_c, \hat{\mathcal{C}}_t, p, \mathbf{t})$ ,  $p \in \mathcal{P}$ , create a component  $\mathcal{C}(p^*, \mathbf{b}^*, \gamma^*)$  with  $p^* = p$ , and

$$\mathcal{C}_c(p^*, \mathbf{b}^*, \gamma^*) = \hat{\mathcal{C}}_c \text{ and } \mathcal{C}_t(p^*, \mathbf{b}^*, \gamma^*) = \hat{\mathcal{C}}_t.$$

Finally using the component  $(\hat{\mathcal{C}}_c, \hat{\mathcal{C}}_t, p, \mathbf{t})$  with  $p = p^*$ , the allocation for component  $\mathcal{C}(p^*, \mathbf{b}^*, \gamma^*)$ ,  $p^* \in \mathcal{P}(\mathbf{b}^*, \gamma^*)$ , is given by

$$b_{i,j}^* = p^*, \quad i \in \mathcal{C}_c(p^*, \mathbf{b}^*, \gamma^*), j \in \mathcal{C}_t(p^*, \mathbf{b}^*, \gamma^*),$$

as well as

$$\gamma_{i,j}^* = t_{i,j}, \quad i \in \mathcal{C}_c(p^*, \mathbf{b}^*, \gamma^*), j \in \mathcal{C}_t(p^*, \mathbf{b}^*, \gamma^*)$$

and

$$\gamma_{i,j}^* = 0, \quad i \in \mathcal{C}_c(p^*, \mathbf{b}^*, \gamma^*), j \notin \mathcal{C}_t(p^*, \mathbf{b}^*, \gamma^*).$$

## 8 Results

In this section we analyze the allocation  $(\mathbf{b}^*, \gamma^*)$  obtained by the algorithm in Section 7, and present our main results. In particular we show in the next subsection that if the supply curves are given by contiguous function, then the allocation  $(\mathbf{b}^*, \gamma^*)$  obtained by the algorithm is an optimal solution to the optimization problem given by Eq. (1). In Subsection 8.2 we use the allocation  $(\mathbf{b}^*, \gamma^*)$  obtained by the algorithm to derive a lower-bound for the cost of any mixed bidding strategy for the

the optimization problem given by Eq. (1), and derive a sequence of mixed bidding strategies that achieves a cost that is equal, or arbitrarily close, to this lower-bound.

We use the following notation to state our results. Let the allocation  $(\mathbf{b}^*, \gamma^*)$  with components  $\mathcal{C}(p^*, \mathbf{b}^*, \gamma^*)$ ,  $p^* \in \mathcal{P}(\mathbf{b}^*, \gamma^*)$ , be the solution to the allocation obtained by the algorithm in Section 7. Then for all bid values  $p^* \in \mathcal{P}(\mathbf{b}^*, \gamma^*)$  let

$$D_{p^*}(b) = \sum_{j \in \mathcal{C}_t(p^*, \mathbf{b}^*, \gamma^*)} D_j(b), \quad b \geq 0,$$

and

$$I_{p^*} = \sum_{i \in \mathcal{C}_c(p^*, \mathbf{b}^*, \gamma^*)} I_i.$$

## 8.1 Continuous Supply Curves

We first analyze the allocation  $(\mathbf{b}^*, \gamma^*)$  obtained by the algorithm presented in Section 7 for the case where the supply curves are given by continuous functions, and show that in this case allocation  $(\mathbf{b}^*, \gamma^*)$  is an optimal allocation.

**Proposition 6.** *Let  $(\mathbf{b}^*, \gamma^*)$  with components  $\mathcal{C}(p^*, \mathbf{b}^*, \gamma^*)$ ,  $p^* \in \mathcal{P}(\mathbf{b}^*, \gamma^*)$ , be the allocation obtained by the algorithm in Section 7. If the supply curves  $D_j(b)$  are given by continuous functions, then  $(\mathbf{b}^*, \gamma^*)$  is an optimal solution for the optimization problem given by Eq. (1), with cost*

$$\sum_{p^* \in \mathcal{P}(\mathbf{b}^*, \gamma^*)} \left[ I_{p^*} p^* - \int_0^{p^*} D_{p^*}(b) db \right].$$

We provide a proof for Proposition 6 in Appendix E.

## 8.2 General Supply Curves

We next consider the general case where the supply curves are not necessarily given by continuous functions. For this case, we consider mixed strategies in addition to pure strategy. Under a mixed strategy campaign  $i$  can use more than one bid value when bidding for impression of a targeting group  $j \in \mathcal{A}_i$ . More precisely, for a mixed strategy  $s$  we let  $\mathcal{P}_{i,j}$  be the set of bid values that campaign  $i$  uses at targeting group  $j$ , and let  $\gamma_{i,j}(b)$ ,  $b \in \mathcal{P}_{i,j}(s)$ , be the fraction of bid requests at targeting group  $j$  for which campaign  $i$  uses the bid value  $b \in \mathcal{P}_{i,j}$ . Let  $\mathcal{S}$  be the set of all such mixed strategies.

In the following we say that a mixed strategy  $s \in \mathcal{S}$  is a solution to the optimization problem given by Eq. 1, if we have that

$$\sum_{j \in \mathcal{A}_i} \sum_{b \in \mathcal{P}_{i,j}} \gamma_{i,j}(b) D_j(b) \geq I_i, \quad i \in \mathcal{N},$$

and

$$\sum_{i \in \mathcal{B}_j} \sum_{b \in \mathcal{P}_{i,j}} \gamma_{i,j}(b) \leq 1, \quad j \in \mathcal{T}.$$

The cost  $C(s)$  of a mixed strategy that is a solution for the optimization problem Eq. 1 is given by

$$C(s) = \sum_{i=1}^{N_c} \left[ \sum_{j=1}^{N_r} \sum_{b \in \mathcal{P}_{i,j}} \gamma_{i,j}(b) \left( D_j(b)b - \int_0^b D_j(b) db \right) \right].$$

We then have the following result.

**Proposition 7.** *Let  $s \in \mathcal{S}$  be a mixed strategy that is a solution to the optimization problem given by Eq. 1. Then we have that*

$$C(s) \geq \sum_{p^* \in \mathcal{P}(\mathbf{b}^*, \gamma^*)} \left[ I_{p^*} p^* - \int_0^{p^*} D_{p^*}(b) db \right].$$

We provide a proof for Proposition 7 in Appendix F. Similar to the case of a single campaign and a single targeting group of Section 5, Proposition 7 states that the expression for the lower-bound of the cost of any mixed strategy  $s$  is equal to the optimal cost in Proposition 6 for the case where the supply curves are given by continuous functions. In addition, Proposition 7 that the allocation  $(\mathbf{b}^*, \gamma^*)$  obtained by the algorithm of Section 7 can be used to compute the lower-bound on the cost of any mixed strategy. This is an important result as it provides an algorithm to compute a lower-bound of the cost of any mixed bidding strategy. We discuss in more details in Section 9 why this result is important in practice.

We next construct a sequence of mixed strategies that either obtains a cost that is equal to the lower-bound given in Proposition 7, or gets arbitrarily close to the lower-bound. More precisely, the next result constructs for each component  $\mathcal{C}(p^*, \mathbf{b}^*, \gamma^*)$ ,  $p^* \in \mathcal{P}(\mathbf{b}^*, \gamma^*)$ , of the solution  $(\mathbf{b}^*, \gamma^*)$  obtained by the algorithm in Section 7 a sequence of mixed bidding strategies under which the cost converges to the lower-bound for this component which is given by

$$I_{p^*} p^* - \int_0^{p^*} D_{p^*}(b) db.$$

**Proposition 8.** *Let  $(\mathbf{b}^*, \gamma^*)$  with components  $\mathcal{C}(p^*, \mathbf{b}^*, \gamma^*)$ ,  $p^* \in \mathcal{P}(\mathbf{b}^*, \gamma^*)$ , be the allocation obtained by the algorithm in Section 7. Then for every component  $\mathcal{C}(p^*, \mathbf{b}^*, \gamma^*)$ ,  $p^* \in \mathcal{P}(\mathbf{b}^*, \gamma^*)$ , the following is true. Let  $j \in \mathcal{C}_t(p^*, \mathbf{b}^*, \gamma^*)$  be a targeting group that belongs to the component  $\mathcal{C}(p^*, \mathbf{b}^*, \gamma^*)$ , and for all campaigns  $i \in \mathcal{B}_j(\mathbf{b}^*, \gamma^*)$  let*

$$\gamma_{i,j} = \frac{\gamma_{i,j}^*}{\sum_{m \in \mathcal{B}_j(\mathbf{b}^*, \gamma^*)} \gamma_{m,j}^*}.$$

Furthermore, let  $s_{p^*}(b_1)$  be a mixed bidding strategy for the component  $\mathcal{C}(p^*, \mathbf{b}^*, \gamma^*)$  such that for all campaigns  $i \in \mathcal{C}_c(p^*, \mathbf{b}^*, \gamma^*)$  and  $j \in \mathcal{C}_t(p^*, \mathbf{b}^*, \gamma^*)$  we have that

$$\mathcal{P}_{i,j} = \{b_1, p^*\}, \quad j \in \mathcal{A}_i(\mathbf{b}^*, \gamma^*),$$

as well as

$$\gamma_{i,j}(b_1) = \frac{\gamma_{i,j} D_1(p^*) - \gamma_{i,j}^* D_1(p^*)}{D_1(p^*) - D_1(b_1)}, \quad j \in \mathcal{A}_i(\mathbf{b}^*, \gamma^*),$$

and

$$\gamma_{i,j}(p^*) = \gamma_{i,j} - \gamma_{i,1}(b_1), \quad j \in \mathcal{A}_i(\mathbf{b}^*, \gamma^*).$$

Finally, let  $C_{p^*}(s_{p^*}(b_1))$  be the cost of strategy  $s_{p^*}(b_1)$  on component  $\mathcal{C}(p^*, \mathbf{b}^*, \gamma^*)$ .

If  $D_{p^*}^{(-)}(p^*) > 0$ , then we have that

$$\lim_{b_1 \uparrow b^*} C_{p^*}(s_{p^*}(b_1)) = I_{p^*} p^* - \int_0^{p^*} D_{p^*}(b) db.$$

We provide a proof for Proposition 8 in Appendix F. In the proof we show that the allocation  $s_{p^*}(b_1)$  is a feasible solution for component  $\mathcal{C}(p^*, \mathbf{b}^*, \gamma^*)$  in the sense that for  $j \in \mathcal{C}_t(p^*, \mathbf{b}^*, \gamma^*)$  we have

$$\sum_{i \in \mathcal{B}_j(\mathbf{b}^*, \gamma^*)} \left[ \gamma_{i,j}(b_1) + \gamma_{i,j}(p^*) \right] = 1,$$

and for  $i \in \mathcal{C}_c(p^*, \mathbf{b}^*, \gamma^*)$  we have

$$\sum_{j \in \mathcal{A}_i(\mathbf{b}^*, \gamma^*)} \left[ \gamma_{i,j}(b_1) D_j(b_1) + \gamma_{i,j}(p^*) D_j(p^*) \right] = I_i.$$

### 8.3 Cost Bound for Pure Strategy $(\mathbf{b}^*, \gamma^*)$

Finally we provide a bound on much we loose by using the pure strategy  $(\mathbf{b}^*, \gamma^*)$  by the algorithm in Section 7 instead of a mixed strategy.

**Corollary 1.** *Let  $(\mathbf{b}^*, \gamma^*)$  be the allocation obtained by the algorithm in Section 7. Furthermore, let  $C(\mathbf{b}^*, \gamma^*)$  be the cost of the allocation  $(\mathbf{b}^*, \gamma^*)$ . Then we have that*

$$\begin{aligned} & C(\mathbf{b}^*, \gamma^*) - \sum_{p \in \mathcal{P}(\mathbf{b}^*, \gamma^*)} \left[ I_{p^*} p^* - \int_0^{p^*} D_{p^*}(b) db \right] \\ & \leq \sum_{p^* \in \mathcal{P}(\mathbf{b}^*, \gamma^*)} \sum_{j \in \mathcal{C}_t(p^*, \mathbf{b}^*, \gamma^*)} \frac{D_j(p^*) - D_j^{(-)}(p^*)}{D_j(p^*)} \int_0^{p^*} D_j(b) db. \end{aligned}$$

Similar to Section 5, Proposition 1 states that the smaller the height  $(D_j(p^*) - D_j^{(-)}(p^*))$  of a discontinuity at  $p^*$  of the supply curve of targeting groups in  $\mathcal{C}_t(p^*, \mathbf{b}^*, \gamma^*)$ ,  $p^* \in \mathcal{P}(\mathbf{b}^*, \gamma^*)$ , the closer the performance of pure strategy  $(\mathbf{b}^*, \gamma^*)$  will be to the lower-bound of Proposition 7.

## 9 Practical Considerations

The focus of this paper is to analyze the optimization problem given by Eq. (1), and derive an algorithm to compute an optimal allocation  $(\mathbf{b}^*, \gamma^*)$ . In Section 4.7 we discussed how a solution to the optimization problem given by Eq. (1) can be used to implement an online bidding algorithm. In this section we briefly discuss some additional aspects of how the results obtained in the paper are relevant for understanding, and implementing, bidding algorithms on practical systems.

### 9.1 Benchmark

An important result of our analysis is Proposition 7 which provides a tight lower-bound on the cost of any mixed strategy. We note that this lower-bound can be easily computed using the algorithm given in Section 7. This lower-bound can be used to benchmark the performance of other online bidding algorithms. In addition, Proposition 1 can be used to determine the cost difference of the allocation  $(\mathbf{b}^*, \gamma^*)$  obtained by the algorithm in Section 7 and any mixed bidding strategies. Note that the allocation  $(\mathbf{b}^*, \gamma^*)$  is a pure bidding strategy that is (slightly) easier to compute and implement in practice. If the cost difference is not too large, then using allocation  $(\mathbf{b}^*, \gamma^*)$  might be preferred over using a mixed strategy.

### 9.2 Structural Properties

Proposition 5, and Proposition 6 and 8, state that an optimal bidding strategy “decomposes” the problem given by Eq. (1) into disjoint components consisting of a set of campaigns and targeting group; an optimal allocation can be obtained by computing an optimal solution for each component individually. In addition, for each component there exists a single bid value  $p^*$  that is used by all campaigns in the component to bid for impressions. This means that rather than having to decide on a bid value for each admissible campaign and targeting group pair in a given component, it suffices to compute a single bid value  $p^*$  for each component. This result is also important for designing online algorithms that dynamically update bid values and allocations as bid requests arrive as this result allows to significantly reduce the search space for optimal bid values, i.e. the online algorithm has to tune/learn only a single bid value for each component.

### 9.3 Mixed Bidding Strategies

An important result of our analysis is to show that in the case (which is the norm in practice) where supply curves are not necessarily given by continuous functions, then pure bidding strategies might not be optimal. Moreover, the result in Proposition 8 states it is sufficient to consider mixed strategies that use with at most 2 bid values for each targeting group, where one if the bid values is equal to the bid value obtained  $p^*$  by the algorithm of Section 7, and then second bid value is chosen so that it is close to  $p^*$ . This implies that (near) optimal mixed bidding strategies are relatively simple to compute and implement.

## 9.4 Dynamic Bidding Strategies

Rather than pre-computing the bid values and allocations at the beginning of a decision period, one can also consider the approach where bid values and allocations are dynamically updated during a decision period. As noted in the previous subsection, the results on the structural property of an optimal allocation can be used in this case to significantly reduce the search space. Analyzing this approach is future research.

## 9.5 Estimating Supply Curves

The online algorithm of Section 4.7 assumes that (a good estimate of) the supply curves are available at the beginning of a decision period. A discussion of how to estimate the supply curves is outside the scope of this paper, and this question has been addressed in related work such as [2]. The result of the experimental case study in [2], as well as experiments that we carried out, suggest that the supply curves can be estimate quite well using historical data. In particular, using historical data over the past few days, the supply curve for decision periods of a few minutes, or even an hour, can be well estimated (predicted). One reason for this is that each targeting group consists of bid requests that come from several applications. While the supply curve of each application may have random fluctuations (noise) over a decision period of a few minutes or an hour, the aggregated supply curves tend to “smooth out” the noise of individual targeting groups and are easier to estimate/predict. We also note that small estimation errors can be “corrected” over a decision period by using the allocation obtained by the algorithm as a starting point, and dynamically make small adjustments to the bid value and allocations. A detailed analysis and discussion of this approach is outside the scope of this paper, and future research.

We note that the supply curves used in the optimization problem should evolve from one decision period to another. That is, for each decision period one would use the newest estimates/predictions of the supply curves that is available based on historical data.

## 9.6 Curse of Dimensionality

In the case where a bid request type consists of a small number of variables/properties such as for example the location, gender, and age of the end user, then targeting groups as given by Definition 1 can be easily computed/determined. We note that in current practical applications, where the number of properties the bid request type is typically small (around 3-5 properties). However if the bid request types consists of a large number of properties (more than 20 properties) then the number of targeting groups can become too large to be computed and/or stored. This problem corresponds to the well-known “curse-of-dimensionality”. Fortunately, in this case it is possible to compute in this case “approximate” optimal solution by using “approximate targeting groups” that can be efficiently computed. This “approximate” optimal solution maintains the structural properties on an optimal solution for the optimization problem given by Eq. (1). The discussion/derivation of such approximate solutions is outside the scope of this paper. We note that in order to be able to derive such approximate optimal solution, it is crucial to first characterize the exact optimal solution to the the optimization problem given by Eq. (1).

## 10 Conclusions

We studied the optimal bidding strategies for a DSP that represents several ad campaigns with overlapping targeting criteria. Using the concept of targeting groups and supply curves, we formally modelled the problem as an optimization problem, and characterize optimal bidding strategies.

While the focus of the paper is the formal analysis of optimal bidding strategies, we discussed in Section 4.7 how the solution obtained by the algorithm of Section 7 can be used to implement an online bidding algorithm. In addition, we discussed in Section 9 several aspects of the formal analysis, and the online algorithm of Section 4.7, that are relevant for the implementation of bidding algorithms in practice. In collaboration with a DSP, we have implemented and evaluated the online algorithm of Section 4.7 using historical data obtain from actual ad exchanges, and actual campaigns that the DSP was representing. Our experiments showed that the online algorithm (using the alloca-

tion  $(b^*, \gamma^*)$  obtained by the algorithm of Section 7) allows the DSP to significantly reduce the cost of getting the required impressions for the ad campaigns it represents.

There are several interesting direction for future research based on the results in this paper. In particular, the question of how to approximate an optimal bidding strategy in the case where the number of targeting groups becomes too large in order to be used on a practical system (curse of dimensionality). In addition, an interesting question is how to tune/dynamically change the bid values and allocations during a decision period in order to account for small changes, or estimation errors, in the supply curves. Preliminary research suggests that the results obtained in this paper can be used to address these two problems.

## References

- [1] Santiago R. Balseiro, Jon Feldman, Vahab Mirrokni, and S. Muthukrishnan. Yield optimization of display advertising with ad exchange. *Management Science*, 60(12):2886–2907, 2014.
- [2] Ying Cui, Ruofei Zhang, Wei Li, and Jianchang Mao. Bid landscape forecasting in online ad exchange marketplace. In *Proceedings of the 17th ACM SIGKDD International Conference on Knowledge Discovery and Data Mining*, KDD '11, pages 265–273, New York, NY, USA, 2011. ACM.
- [3] Robert G. Gallager. *Discrete Stochastic Processes*. Springer, 1996.
- [4] R. Gummadi, P. B. Key, and A. Proutiere. Optimal bidding strategies in dynamic auctions with budget constraints. In *2011 49th Annual Allerton Conference on Communication, Control, and Computing (Allerton)*, pages 588–588, Sep. 2011.
- [5] Chong Jiang, Carolyn L. Beck, and R Srikant. Bidding with limited statistical knowledge in online auctions. *ACM SIGMETRICS Performance Evaluation Review*, 41:38–41, 04 2014.
- [6] Aranyak Mehta. Online matching and ad allocation. *Found. Trends Theor. Comput. Sci.*, 8(4):265–368, October 2013.
- [7] Claudia Perlich, Brian Dalessandro, Rod Hook, Ori Stitelman, Troy Raeder, and Foster Provost. Bid optimizing and inventory scoring in targeted online advertising. In *Proceedings of the 18th ACM SIGKDD International Conference on Knowledge Discovery and Data Mining*, KDD '12, pages 804–812, New York, NY, USA, 2012. ACM.
- [8] Santiago R. Balseiro and Yonatan Gur. Learning in repeated auctions with budgets: Regret minimization and equilibrium. *SSRN Electronic Journal*, 02 2017.
- [9] Panagiotis Spentzouris, Iordanis Koutsopoulos, Kasper Grud Madsen, and Tommy Vestergaard Hansen. *Advertiser Bidding Prediction and Optimization in Online Advertising*, pages 413–424. 05 2018.
- [10] Weinan Zhang and Jun Wang. Statistical arbitrage mining for display advertising. 06 2015.
- [11] Weinan Zhang, Shuai Yuan, and Jun Wang. Optimal real-time bidding for display advertising. In *Proceedings of the 20th ACM SIGKDD International Conference on Knowledge Discovery and Data Mining*, KDD '14, pages 1077–1086, New York, NY, USA, 2014. ACM.
- [12] Weinan Zhang, Ying Zhang, Bin Gao, Yong Yu, Xiaojie Yuan, and Tie-Yan Liu. Joint optimization of bid and budget allocation in sponsored search. *Proceedings of the ACM SIGKDD International Conference on Knowledge Discovery and Data Mining*, 08 2012.

## A Properties of Mixed Strategies

In this appendix we derive properties of mixed strategies that we use in our analysis.

### A.1 Mixed Strategies for A Single Targeting Group

We first consider the case of Section 5 where a single campaign  $i = 1$  gets impressions from a single targeting group  $j = 1$ , and  $I_1$  is the total number of impressions that the campaign is required to obtain. Furthermore, let  $s \in \mathcal{S}$  be a mixed strategy as defined in Section 5.2.

Our first result states that for this case it suffices to consider mixed bidding strategies that obtain exactly the required number of impressions from the targeting group.

**Lemma 2.** *Consider the special case where a single campaign  $i = 1$  gets impressions from a single targeting group  $j = 1$ , and let  $I_1$  be the total number of impressions that the campaign is required to obtain. If  $s \in \mathcal{S}$  is a mixed bidding strategy such that*

$$\sum_{b \in \mathcal{P}(s)} \gamma_{1,1}(b) D_1(b) > I_1,$$

*then there exists a mixed bidding strategy  $s' \in \mathcal{S}$  with*

$$\sum_{b \in \mathcal{P}'_{1,1}} \gamma'_{1,1}(b) D_1(b) = I_1,$$

*that leads to a strictly lower cost, and we have that*

$$C(s') < C(s).$$

This result can be proved by showing that reducing the allocations  $\gamma_{1,1}(b)$ ,  $b \in \mathcal{P}_{1,1}$ , of strategy  $s$  until the exact number of impressions is obtained will lead to a strictly lower cost compared with  $C(s)$ . We omit here a detailed derivation.

The next result provides a condition under which it suffices to consider mixed strategies such that

$$\sum_{b \in \mathcal{P}} \gamma_{1,1}(b) = 1.$$

**Lemma 3.** *Consider the special case where a single campaign  $i = 1$  gets impressions from a single targeting group  $j = 1$ , and let  $I_1$  be the total number of impressions that the campaign is required to obtain. Suppose that for a mixed bidding strategy  $s \in \mathcal{S}$  we have that*

$$\sum_{b \in \mathcal{P}} \gamma_{1,1}(b) < 1,$$

*and there exists bid values  $b_1, b_2 \in \mathcal{P}_{1,1}$  such that*

$$0 < D_1(b_1) < D_1(b_2).$$

*Then there exists a mixed bidding strategy  $s' \in \mathcal{S}$  with*

$$\sum_{b \in \mathcal{P}'_{1,1}} \gamma'_{1,1}(b) D_1(b) = I_1$$

*and*

$$\sum_{b \in \mathcal{P}'_{1,1}} \gamma'_{1,1}(b) = 1,$$

*that leads to a strictly lower cost, and we have that*

$$C(s') < C(s).$$

*Proof.* We prove the result by constructing a mixed bidding strategy  $s' \in \mathcal{S}$  that leads to a strictly lower cost than the given strategy  $s$

Let

$$\gamma_{1,1}(b_1)D_1(b_1) + \gamma_{1,1}(b_2)D_1(b_2) = I,$$

i.e.  $I$  is the amount of impressions that are obtained by bid values  $b_1$  and  $b_2$  combined.

Let  $\gamma > 0$  be such that

$$\gamma + \sum_{b \in \mathcal{P}_{1,1} \setminus \{b_1, b_2\}} \gamma_{1,1}(b) = 1.$$

As by assumption we have that

$$\sum_{b \in \mathcal{P}} \gamma_{1,1}(b) < 1,$$

it follows that

$$\gamma > \gamma_{1,1}(b_1) + \gamma_{1,1}(b_2).$$

Let  $\Delta_\gamma$  be such that

$$\Delta_\gamma D_1(b_1) + (\gamma - \Delta_\gamma) D_1(b_2) = I.$$

Note that such a  $\Delta_\gamma$  exists as we have that

$$\gamma_{1,1}(b_1)D_1(b_1) + \gamma_{1,1}(b_2)D_1(b_2) = I,$$

as well as

$$0 < D_1(b_1) < D_1(b_2)$$

and

$$\gamma_{1,1}(b_1) + \gamma_{1,1}(b_2) < \gamma.$$

Furthermore, these properties imply that

$$\Delta_\gamma > \gamma_{1,1}(b_1)$$

and

$$(\gamma - \Delta_\gamma) < \gamma_{1,1}(b_2).$$

Using these definitions, we construct the mixed bidding strategy  $s' \in \mathcal{S}$  as follows. We set

$$\mathcal{P}'_{1,1} = \mathcal{P}_{1,1},$$

and let

$$\gamma'_{1,1}(b_1) = \Delta_\gamma,$$

and

$$\gamma'_{1,1}(b^*) = \gamma - \Delta_\gamma.$$

For all  $b \in \mathcal{P}'_{1,1}$ ,  $b \neq b_1, b_2$ , we set

$$\gamma'_{1,1}(b) = \gamma_{1,1}(b),$$

and the new strategy  $s'$  differs from strategy  $s$  only the the allocations for the bid values  $b_1$  and  $b_2$ .

The difference in the cost for the two strategies is then given by

$$\begin{aligned}
& C(s) - C(s') \\
&= \left( \gamma_{1,1}(b_1) - \Delta_\gamma \right) \left[ D_1(b_1)b_1 - \int_0^{b_1} D_1(b)db \right] \\
&\quad + \left( \gamma_{1,1}(b_1) - \gamma + \Delta_\gamma \right) \left[ D_1(b_2)b_2 - \int_0^{b_2} D_1(b)db \right] \\
&= \gamma_{1,1}(b_1) \left[ D_1(b_1)b_1 - \int_0^{b_1} D_1(b)db \right] \\
&\quad + \left( \gamma_{1,1}(b_2) - \gamma \right) \left[ D_1(b_2)b_2 - \int_0^{b_2} D_1(b)db \right] \\
&\quad - \Delta_\gamma \left[ D_1(b_1)b_1 - D_1(b_2)b_2 - \int_{b_1}^{b_2} D_1(b)db \right] \\
&> \gamma_{1,1}(b_1) \left[ D_1(b_1)b_1 - \int_0^{b_1} D_1(b)db \right] \\
&\quad + \left( \gamma_{1,1}(b_2) - \gamma \right) \left[ D_1(b_2)b_2 - \int_0^{b_2} D_1(b)db \right] \\
&\quad - \Delta_\gamma \left[ D_1(b_1)b_1 - \int_0^{b_1} D_1(b)db \right].
\end{aligned}$$

Using the fact that

$$\Delta_\gamma D_1(b_1) + (\gamma - \Delta_\gamma) D_1(b_2) = I,$$

we obtain that

$$\begin{aligned}
& C(s) - C(s') \\
&> \gamma_{1,1}(b_1) \left[ D_1(b_1)b_1 - \int_0^{b_1} D_1(b)db \right] \\
&\quad + \gamma_{1,1}(b_2) \left[ D_1(b_2)b_2 - \int_0^{b_2} D_1(b)db \right] \\
&\quad + \gamma \int_0^{b_2} D_1(b)db \\
&\quad - I + \Delta_\gamma \int_0^{b_1} D_1(b)db.
\end{aligned}$$

Using the fact that

$$\gamma_{1,1}(b_1) D_1(b_1) + \gamma_{1,1}(b_2) D_1(b_2) = I_1,$$

we obtain that

$$\begin{aligned}
& C(s) - C(s') \\
&> -\gamma_{1,1}(b_1) \int_0^{b_1} D_1(b)db \\
&\quad - \gamma_{1,1}(b_2) \int_0^{b_2} D_1(b)db \\
&\quad + \gamma \int_0^{b_2} D_1(b)db \\
&\quad + \Delta_\gamma \int_0^{b_1} D_1(b)db.
\end{aligned}$$

Finally, using the fact that

$$\gamma_{1,1}(b_1) + \gamma_{1,1}(b_2) < \gamma,$$

we obtain that

$$\begin{aligned} & C(s) - C(s') \\ & > \Delta_\gamma \int_0^{b_1} D_1(b) db \\ & > 0. \end{aligned}$$

This establishes the result of the proposition.  $\square$

The next results is the one of Proposition 2.

**Proposition 9.** *Consider the special case where a single campaign  $i = 1$  gets impressions from a single targeting group  $j = 1$ , and let  $I_1$  be the total number of impressions that the campaign is required to obtain. Let  $C(s)$  be the cost of a given mixed bidding strategy  $s \in \mathcal{S}$  such that*

$$\sum_{b \in \mathcal{P}} \gamma_{1,1}(b) D_1(b) = I_1.$$

Then we have that

$$C(s) \geq I_1 b^* - \int_0^{b^*} D_1(b) db,$$

where the bid value  $b^*$  is as given by Eq. (2) and (3).

*Proof.* Let  $s \in \mathcal{S}$  be the mixed bidding strategy as given in the statement of the proposition, and let the set  $B^*(I_1)$  be the set of all bid values  $b \in \mathcal{P}_{1,1}$  such that

$$D_1(b) = D_1(b^*) = I_1.$$

Note that if we have that

$$B^*(I_1) = \mathcal{P}_{1,1},$$

then we have by the definition of the cost function  $C(s)$  that

$$C(s) = I_1 b^* - \int_0^{b^*} D_1(b) db;$$

and the result of the proposition follows. Therefore we assume in the following that

$$B^*(I_1) \neq \mathcal{P}_{1,1}.$$

For this case, let the sets  $B^{(-)}$  and  $B^{(+)}$  be given by

$$B^{(-)} = \{b \in \mathcal{P}_{1,1} | D_1(b) < I_1\}$$

and

$$B^{(+)} = \{b \in \mathcal{P}_{1,1} | D_1(b) > I_1\}.$$

We first consider the case where there exists a bid value  $b_1 \in B^{(-)}$  such that

$$D_1(b_1) > 0.$$

Using Lemma 2 and 3, without loss of generality we can assume in this case that

$$\sum_{b \in \mathcal{P}} \gamma_{1,1}(b) D_1(b) = I_1$$

and

$$\sum_{b \in \mathcal{P}} \gamma_{1,1}(b) = 1,$$

as otherwise the bidding strategy  $s$  is not an optimal bidding strategy. It then follows that

$$\begin{aligned} & \sum_{b \in B^{(-)}} \gamma(b) D_1(b) + \sum_{b \in B^{(+)}} \gamma(b) D_1(b) \\ = & I_1 \left[ \sum_{b \in B^{(-)}} \gamma(b) + \sum_{b \in B^{(+)}} \gamma(b) \right]. \end{aligned}$$

Using these results, we obtain that

$$\begin{aligned} & C(s) - b^* I_1 + \int_0^{b^*} D_1(b) db \\ = & \sum_{b \in B^{(-)}} \gamma(b) \left[ (b D_1(b) - b^* I_1) + \int_b^{b^*} D_1(b) db \right] \\ & + \sum_{b \in B^{(+)}} \gamma(b) \left[ (b D_1(b) - b^* I_1) - \int_{b^*}^b D_1(b) db \right] \\ \geq & \sum_{b \in B^{(-)}} \gamma(b) \left[ (b D_1(b) - b^* I_1) + D_1(b)(b^* - b) \right] \\ & + \sum_{b \in B^{(+)}} \gamma(b) \left[ (b D_1(b) - b^* I_1) - D_1(b)(b - b^*) \right] \\ = & \sum_{b \in B^{(-)}} \gamma(b) \left[ b^* (D_1(b) - I_1) \right] \\ & + \sum_{b \in B^{(+)}} \gamma(b) \left[ b^* (D_1(b) - I_1) \right] \\ = & 0. \end{aligned}$$

This implies that

$$C(s) - b^* I_1 + \int_0^{b^*} D_1(b) db \geq 0,$$

and the result of the proposition follows for the case where there exists a bid value  $b_1 \in B^{(-)}$  such that

$$D_1(b_1) > 0.$$

It remains to consider the case where we have that

$$D_1(b) = 0, \quad b \in B^{(-)}.$$

Without loss of generality we can assume in this case that

$$B^{(-)} = \emptyset,$$

and we have that

$$D_1(b) \geq I_1, \quad b \in \mathcal{P}.$$

Using Lemma 2, without loss of generality we can furthermore assume in this case that

$$\sum_{b \in \mathcal{P}} \gamma_{1,1}(b) D_1(b) = I_1.$$

Let

$$0 \leq \Delta_\gamma = 1 - \sum_{b \in \mathcal{P}} \gamma_{1,1}(b).$$

It then follows that

$$\begin{aligned}
& C(s) - b^* I_1 + \int_0^{b^*} D_1(b) db \\
&= \sum_{b \in B^{(+)}} \gamma(b) \left[ (b D_1(b) - b^* I_1) - \int_{b^*}^b D_1(b) db \right] \\
&\quad - \Delta_\gamma \left[ b^* I_1 - \int_0^{b^*} D_1(b) db \right] \\
&\geq \sum_{b \in B^{(+)}} \gamma(b) \left[ (b D_1(b) - b^* I_1) - D_1(b)(b - b^*) \right] \\
&\quad - \Delta_\gamma \left[ b^* I_1 - \int_0^{b^*} D_1(b) db \right] \\
&= \sum_{b \in B^{(+)}} \gamma(b) (D_1(b) - I_1) b^* \\
&\quad - \Delta_\gamma \left[ b^* I_1 - \int_0^{b^*} D_1(b) db \right] \\
&\geq \sum_{b \in B^{(+)}} \gamma(b) (D_1(b) - I_1) b^* \\
&\quad - \Delta_\gamma \left[ b^* I_1 - \int_0^{b^*} D_1(b) db \right] \\
&\geq \sum_{b \in B^{(+)}} \gamma(b) (D_1(b) - I_1) b^* \\
&\quad - \Delta_\gamma b^* I_1 \\
&= b^* \sum_{b \in B^{(-)}} \gamma(b) D_1(b) \\
&\quad - I_1 b^*.
\end{aligned}$$

Using the fact that

$$\sum_{b \in \mathcal{P}} \gamma_{1,1}(b) D_1(b) = I_1,$$

we obtain that

$$C(s) - b^* I_1 + \int_0^{b^*} D_1(b) db \geq 0.$$

The result of the proposition then follows. □

## A.2 Mixed Strategies on a Set of Targeting Groups

Next we consider a set  $\mathcal{T} = \{1, \dots, k\}$  of targeting groups, and provide a lower bound on the cost  $C(s)$  of any mixed strategy  $s \in \mathcal{S}$  to get a total of  $I_t$  impressions from these targeting groups. More precisely, we have the following result.

**Proposition 10.** *Let  $\mathcal{T} = \{1, \dots, k\}$  be a set of targeting groups, and let*

$$D_t(b) = \sum_{j=1}^k D_j(b), \quad b \geq 0.$$

Furthermore, let  $b_t$  be the bid value such that

$$\sum_{j=1}^k D_j(b_t) \geq I_t$$

and

$$\sum_{j=1}^k D_j(b) < I_t, \quad 0 \leq b < b_t.$$

Then for any mixed strategy  $s \in \mathcal{S}$  that obtains a total of  $I_t$  impressions from the set of targeting groups  $\mathcal{T}$  we have that

$$C(s) \geq I_t b_t - \int_0^{b_t} D_b(b) db.$$

*Proof.* Let  $s \in \mathcal{S}$  be a mixed strategy that obtains  $I_t$  impressions from the set  $\mathcal{T} = \{1, \dots, k\}$  of targeting groups, and let

$$I_j(s), \quad j = 1, \dots, k,$$

be the number of impressions that strategy  $s$  obtains from targeting group  $j$ .

Note that in order to prove the result of the proposition, it suffices to show that for  $n = 1, 2, \dots, k$ , the cost for the strategy  $s$  to get a total of

$$I(n, s) = \sum_{j=1}^n I_j(s)$$

impressions from the first  $n$  targeting groups is lower-bounded by

$$I(n, s) b(n, s) - \int_0^{b(n, s)} \sum_{j=1}^n D_j(b) db,$$

where the bid value  $b(n, s)$  is such that

$$\sum_{j=1}^n D_j(b(n, s)) \geq I(n, s)$$

and

$$\sum_{j=1}^n D_j(b) < I(n, s), \quad 0 \leq b < b(n, s).$$

To prove this result, we proceed as follows. First we note that by Proposition 9 we have for each targeting group  $j = 1, \dots, k$ , that the cost for getting  $I_j(s)$  impressions from targeting group  $j$  is lower-bounded by

$$I_j(s) b_j(s) - \int_0^{b_j(s)} D_j(b) db,$$

where the bid value  $b_j(s)$  is such that

$$D_j(b_j(s)) \geq I_j(s)$$

and

$$D_j(b) < I_j(s), \quad 0 \leq b < b_j(s).$$

We first show that the result is true for  $n = 2$ . Let  $C(2, s)$  be the cost for the strategy to get  $I(2, s)$  impressions from the first two targeting groups. Without loss of generality, we can assume that

$$b_1(s) \leq b_2(s).$$

Note that this implies that

$$b_1(s) \leq b(2, s) \leq b_2(s).$$

Using Proposition 9 we have that

$$\begin{aligned} & C(2, s) - I(2, s) b(2, s) + \int_0^{b(2, s)} [D_1(b) + D_2(b)] db \\ & \geq I_1(s) b_1(s) - \int_0^{b_1(s)} D_1(b) db \\ & \quad + I_2(s) b_2(s) - \int_0^{b_2(s)} D_2(b) db \\ & \quad - I(2, s) b(2, s) + \int_0^{b(2, s)} [D_1(b) + D_2(b)] db \end{aligned}$$

It then follows that

$$\begin{aligned}
& C(2, s) - I(2, s)b(2, s) + \int_0^{b(2, s)} [D_1(b) + D_2(b)] db \\
\geq & I_1(s)b_1(s) + I_2(s)b_2(s) - (I_1(s) + I_2(s))b(2, s) \\
& + \int_{b_1(s)}^{b(2, s)} D_1(b) db \\
& - \int_{b(2, s)}^{b_2(s)} D_2(b) db \\
\geq & -I_1(s)(b(2, s) - b_1(s)) \\
& + I_2(s)(b_2(s) - b(2, s)) \\
& + D_1(b_1(s))(b(2, s) - b_1(s)) \\
& - D_2^{(-)}(b_2(s))(b_2(s) - b(2, s)) \\
= & + [D_1(b_1(s)) - I_1(s)](b(2, s) - b_1(s)) \\
& + [I_2(s) - D_2^{(-)}(b_2(s))](b_2(s) - b(2, s))
\end{aligned}$$

As by assumption we have that

$$b_1(s) \leq b(2, s) \leq b_2(s)$$

as well as

$$D_1(b_1(s)) > I_1(s)$$

and

$$D_2(b) < I_2(s), \quad 0 \leq b < b_2(s).$$

it follows that

$$I_2(s) \geq D_2^{(-)}(b_2(s))$$

and

$$C(2, s) \geq I(2, s)b(2, s) - \int_0^{b(2, s)} [D_1(b) + D_2(b)].$$

This establishes the result for the case of  $n = 2$ .

We can then prove the result by induction. Assume that the result is true for  $n, n = 1, \dots, k - 1$ . For this, let the function  $D(n, b)$  be given by

$$D(n, b) = \sum_{j=1}^n D_j(b).$$

The result is then also true for  $n + 1$  as we have that

$$\begin{aligned}
& C(n + 1, s) - I(n + 1, s)b(n + 1, s) + \int_0^{b(n+1,s)} D(n + 1, b)db \\
\geq & I(n, s)b(n, s) + \int_0^{b(n,s)} D(n, b)db \\
& + I_{n+1}(s)b_{n+1}(s) - \int_0^{b_{n+1}(s)} D_{n+1}(b)db \\
& - I(n + 1, s)b(n + 1, s) + \int_0^{b(n+1,s)} D(n + 1, b)db \\
= & I(n, s)b_1(s) + I_{n+1}(s)b_{n+1}(s) - I(n + 1, s)b(n + 1, s) \\
& + \int_{b(n,s)}^{b(n+1,s)} D(n, b)db \\
& - \int_{b(n+1,s)}^{b_{n+1}(s)} D_{n+1}(b)db \\
\geq & -I(n, s)(b(n + 1, s) - b_1(n, s)) \\
& + I_{n+1}(s)(b_{n+1}(s) - b(n + 1, s)) \\
& + D(n, b(n, s))(b(n + 1, s) - b(n, s)) \\
& - D_{n+1}^{(-)}(b_{n+1}(s))(b_{n+1}(s) - b(n, s)) \\
= & \left[ D(n, b(n, s)) - I(n, s) \right] (b(n + 1, s) - b(n, s)) \\
& + \left[ I_{n+1}(s) - D_{n+1}^{(-)}(b_{n+1}(s)) \right] (b_{n+1}(s) - b(n + 1, s)).
\end{aligned}$$

As by assumption we have that

$$b(n, s) \leq b(n + 1, s) \leq b_{n+1}(s),$$

as well

$$D(n, b(n, s)) > I(n, s)$$

and

$$D_{n+1}(b) < I_{n+1}(s), \quad 0 \leq b < b_{n+1}(s),$$

it follows that

$$I_{n+1}(s) \geq D_{n+1}^{(-)}(b_{n+1}(s))$$

and

$$C(n + 1, s) \geq I(n + 1, s)b(n + 1, s) - \int_0^{b(n+1,s)} D(n + 1, b)db.$$

This establishes the induction step. The result of the proposition then follows.  $\square$

## B Proofs for Section 5

In this appendix we prove Proposition 1-3 of Section 5.

### B.1 Proof of Proposition 1

We start out by proving Proposition 1 in a more general setting. We have the following result.

**Proposition 11.** *Consider the special case where a single campaign  $i = 1$  gets impressions from a single targeting group  $j = 1$ , and let  $I_1$  be the total number of impressions that the campaign is required to obtain. If there exists a bid value  $b^*$  such that*

$$D_1(b^*) = I_1,$$

then an optimal bidding strategy in the set of mixed bidding strategies  $\mathcal{S}$  is a pure bidding strategy given by

$$\gamma_{1,1}^*(b^*) = 1.$$

Furthermore, the optimal cost is given by

$$I_1 b^* - \int_0^{b^*} D_1(b) db.$$

Proposition 1 states that if there exists a bid value  $b^*$  such that

$$D_1(b^*) = I_1,$$

then the optimal bidding strategy is the pure strategy given in the statement of the proposition.

*Proof.* By Proposition 9, the cost of getting  $I_1$  impressions from the targeting group 1 is lower-bounded by

$$I^* b^* - \int_0^{b^*} D_1(b) db.$$

Note that the pure strategy given in the statement of the proposition achieves this lower-bound, and hence is an optimal allocation.  $\square$

## B.2 Proof of Proposition 2

Recall the statement of Proposition 2.

**Proposition 12.** *Consider the special case where a single campaign  $i = 1$  gets impressions from a single targeting group  $j = 1$ , and let  $I_1$  be the total number of impressions that that campaign is required to obtain. Let  $C(s)$  be the cost of a given mixed bidding strategy  $s \in \mathcal{S}$ . Then we have that*

$$C(s) \geq I_1 b^* - \int_0^{b^*} D_1(b) db,$$

where the bid value  $b^*$  is as given by Eq. (2) and (3).

This result follows immediately from Proposition 9.

## B.3 Proof of Proposition 3

Recall the statement of Proposition 3.

**Proposition 13.** *Consider the special case where a single campaign  $i = 1$  gets impressions from a single targeting group  $j = 1$ , and let  $I_1$  be the total number of impressions that the campaign is required to obtain. Let the bid value  $b^*$  be as given by Eq. (2) and (3) and suppose that*

$$D^{(-)}(b^*) > 0.$$

Furthermore, let the mixed bidding strategy  $s(b_1) \in \mathcal{S}$ ,  $0 \leq b_1 < b^*$ , be given by

$$\mathcal{P}_{1,1} = \{b_1, b^*\},$$

and

$$\gamma_{1,1}(b_1) = \frac{D_1(b^*) - I_1}{D_1(b^*) - D_1(b_1)} \quad \text{and} \quad \gamma_{1,1}(b^*) = 1 - \gamma_{1,1}(b_1).$$

Then we have that

$$\lim_{b_1 \uparrow b^*} C(s(b_1)) = I_1 b^* - \int_0^{b^*} D_1(b) db.$$

*Proof.* To prove Proposition 3 we construct a sequence of bidding strategies  $s^{(n)}$  given as follows. Consider a sequence of bid values  $b_n$ ,  $n \geq 1$ , such that

$$b_n < b_{n+1}$$

and

$$\lim_{n \rightarrow \infty} b_n = b^*.$$

The strategy  $s^{(n)}$  is then given by

$$\mathcal{P}_{1,1}^{(n)} = \{b_n, b^*\},$$

as well as

$$\gamma^{(n)}(b_n) = \frac{D_1(b^*) - I_1}{D_1(b^*) - D_1(b_n)}$$

and

$$\gamma^{(n)}(b^*) = 1 - \gamma^{(n)}(b_n).$$

By the assumptions made in the statement of the proposition we have that

$$\gamma^{(n)}(b_n) = \frac{D_1(b^*) - I_1}{D_1(b^*) - D_1(b_n)} < 1,$$

and therefore

$$\gamma^{(n)}(b_n)D_1(b_n) + \gamma^{(n)}(b^*)D_1(b^*) = I_1.$$

Note that the bidding strategy  $s^{(n)}$  with  $b_n$  is equal to the bidding strategy  $s(b_1 = b_n)$  given in the statement of the proposition.

The cost  $C(s^{(n)})$  of strategy  $s^{(n)}$  is given by

$$\begin{aligned} C(s^{(n)}) &= \gamma^{(n)}(b_n) \left[ D_1(b_n)b_n - \int_0^{b_n} D_1(b)db \right] \\ &\quad + (1 - \gamma^{(n)}(b_n)) \left[ D_1(b^*)b^* - \int_0^{b^*} D_1(b)db \right] \\ &= \gamma^{(n)}(b_n)D_1(b_n)b_n + (1 - \gamma^{(n)}(b_n))D_1(b^*)b^* \\ &\quad - \int_0^{b^*} D_1(b)db \\ &\quad + \gamma^{(n)}(b_n) \int_{b_n}^{b^*} D_1(b)db \\ &= \gamma^{(n)}(b_n)D_1(b_n)b^* + (1 - \gamma^{(n)}(b_n))D_1(b^*)b^* \\ &\quad + \gamma^{(n)}(b_n)D_1(b_n)(b_n - b^*) \\ &\quad - \int_0^{b^*} D_1(b)db \\ &\quad + \gamma^{(n)}(b_n) \int_{b_n}^{b^*} D_1(b)db \\ &= I_1b^* - \int_0^{b^*} D_1(b)db \\ &\quad + \gamma^{(n)}(b_n)D_1(b_n)(b_n - b^*) \\ &\quad + \gamma^{(n)}(b_n) \int_{b_n}^{b^*} D_1(b)db. \end{aligned}$$

We then obtain that

$$\begin{aligned}
& \lim_{n \rightarrow \infty} C(s^{(n)}) \\
&= I_1 b^* - \int_0^{b^*} D_1(b) db \\
&\quad + \lim_{b_n \rightarrow b^*} \gamma^{(n)}(b_n) D_1(b_n) (b_n - b^*) \\
&\quad + \lim_{b_n \rightarrow b^*} \gamma^{(n)}(b_n) \int_{b_n}^{b^*} D_1(b) db \\
&= I_1 b^* - \int_0^{b^*} D_1(b) db.
\end{aligned}$$

The result of the proposition then follows. □

## C Properties of an Optimal Allocation

In this appendix we derive properties of an optimal solution  $(\mathbf{b}^*, \gamma^*)$  for the optimization problem given by Eq. (1) for the case where the supply curves are given by continuous functions, and provide the proofs for results in Section 6. To help the reader we repeat here the results of Section 6 in order to keep the flow of the argument.

We proceed as follows. We first provide in Subsection C.1 and Subsection C.2 the for results in Section 6. Using these results, we derive in Subsection C.3 sufficient conditions for an allocation  $(\mathbf{b}^*, \gamma^*)$  to be an optimal solution.

### C.1 Properties of the Bid Values of an Optimal Solution

In this subsection we derive structural properties of the bid values  $b^*$  of an optimal solution for the optimization problem given by Eq. (1) for the case where the supply curves are given by continuous functions. For this we use the following notation. Let  $(\mathbf{b}^*, \gamma^*)$  be a solution to the the optimization given by Eq. (1). Then for each campaign

$$i \in \mathcal{N} = \{1, \dots, N_c\}$$

the set  $\mathcal{A}_i(\mathbf{b}^*, \gamma^*)$  is given by

$$\mathcal{A}_i(\mathbf{b}^*, \gamma^*) = \{j \in \mathcal{A}_i | \gamma_{i,j}^* D_j(b_{i,j}^*) > 0\}.$$

Or in other words, the set  $\mathcal{A}_i(\mathbf{b}^*, \gamma^*)$  contains all the targeting groups  $j$  from which campaign receives a positive number  $\gamma_{i,j}^* D_j(b_{i,j}^*)$  of impressions under the solution  $(\mathbf{b}^*, \gamma^*)$ .

Similarly, for each targeting group  $j = 1, \dots, N_r$  the set  $\mathcal{B}_j(\mathbf{b}^*, \gamma^*)$  is given by

$$\mathcal{B}_j(\mathbf{b}^*, \gamma^*) = \{i \in \mathcal{B}_j | \gamma_{i,j}^* D_j(b_{i,j}^*) > 0\}.$$

The next lemma provides a necessary condition for a solution  $(\mathbf{b}^*, \gamma^*)$  to be an optimal solution for the optimization problem given by Eq. (1).

**Lemma 4.** *If the supply curves  $D_j(b)$ ,  $j \in \mathcal{T}$ , are given by continuous functions, then the following is true. If  $(\mathbf{b}^*, \gamma^*)$  is an optimal solution to the optimization problem given by Eq. (1), then we have for each campaign  $i = 1, \dots, N_c$  that*

$$\sum_{j \in \mathcal{A}_i(\mathbf{b}^*, \gamma^*)} \gamma_{i,j}^* D_j(b_{i,j}^*) = I_i.$$

Lemma 4 states that if  $(\mathbf{b}^*, \gamma^*)$  is an optimal solution to the optimization problem given by Eq. (1), then we have for each campaign  $i = 1, \dots, N_c$  that the acquired number of impressions for campaign  $i$  meets the required number of impressions  $I_i$  of campaign  $i$ . Note that if this is not the case, then we can lower the cost for campaign  $i$  by decreasing the total number of impressions that are acquired (by lowering the allocation variables  $\gamma_{i,j}^*$ ,  $j \in \mathcal{A}_i$ ).

**Proposition 14.** *If the supply curves  $D_j(b)$ ,  $j \in \mathcal{T}$ , are given by continuous functions, then the following is true. Let  $(\mathbf{b}^*, \gamma^*)$  be an optimal solution to the optimization problem given by Eq. (1). Then for each targeting group  $j = 1, \dots, N_r$  there exists a bid value  $q_j^*$  such that*

$$D_j(b_{i,j}^*) = D_j(q_j^*), \quad i \in \mathcal{B}_j(\mathbf{b}^*, \gamma^*).$$

*Proof.* We prove the lemma by contradiction. That is we assume that  $(\mathbf{b}, \gamma)$  is an optimal solution to the optimization problem given by Eq. (1), and there exists a targeting group  $j$ ,  $j = 1, \dots, N_r$ , and two campaigns  $m, n \in \mathcal{B}_j(\mathbf{b}, \gamma)$  such that

$$D_j(b_{m,j}) < D_j(b_{n,j}).$$

Note that this implies that

$$b_{m,j} < b_{n,j}.$$

We then construct a new allocation  $(\mathbf{b}^*, \gamma^*)$  that is a feasible solution to the optimization problem given by Eq. (1), and leads to a strictly lower overall cost compared with the allocation  $(\mathbf{b}, \gamma)$ . Note that proving this result suffices to prove the proposition as this leads to a contradiction with the assumption that  $(\mathbf{b}, \gamma)$  is an optimal allocation.

To construct the new allocation  $(\mathbf{b}^*, \gamma^*)$  we use the following definitions. Let

$$I_{m,j} = \gamma_{m,j} D_j(b_{m,j})$$

and

$$I_{n,j} = \gamma_{n,j} D_j(b_{n,j}).$$

Furthermore, let

$$I_t = I_{m,j} + I_{n,j},$$

and let the bid value  $q_j^*$  be such that

$$(\gamma_{m,j} + \gamma_{n,j}) D_j(q_j^*) = I_t.$$

As we have that

$$D_j(b_{m,j}) < D_j(b_{n,j}),$$

a bid value  $q_j^*$  with the above property exists and we have that

$$D_j(b_{m,j}) < D_j(q_j^*) < D_j(b_{n,j})$$

and

$$b_{m,j} < q_j^* < b_{n,j}.$$

Using these definitions, let  $(\mathbf{b}^*, \gamma^*)$  be the allocation that is identical to  $(\mathbf{b}, \gamma)$ , except that

$$b_{m,j}^* = q_j^*$$

and

$$b_{n,j}^* = q_j^*,$$

as well as

$$\gamma_{m,j}^* = \frac{I_{m,j}}{D_j(q_j^*)}$$

and

$$\gamma_{n,j}^* = \frac{I_{n,j}}{D_j(q_j^*)}.$$

Compared with the allocation  $(\mathbf{b}, \gamma)$ , the new allocation  $(\mathbf{b}^*, \gamma^*)$  only affects (changes) the allocations and bid values of campaign  $m$  and  $n$  at targeting group  $j$ .

As by construction we have that

$$(\gamma_{m,j} + \gamma_{n,j}) D_j(q_j^*) = I_{m,j} + I_{n,j},$$

we have that

$$\gamma_{m,j}^* + \gamma_{n,j}^* = \gamma_{m,j} + \gamma_{n,j}.$$

Furthermore, by construction we have that

$$\gamma_{m,j}^* D_j(q_j^*) = I_{m,j}.$$

Similarly, we have that

$$\gamma_{n,j}^* D_j(q_j^*) = I_{n,j}.$$

These two results show that under the allocation  $(\mathbf{b}^*, \boldsymbol{\gamma}^*)$  both campaign  $m$  and  $n$  receive the required number of impressions  $I_{m,j}$ , and  $I_{n,j}$ , respectively, from the targeting group  $j$ . It then follows that the allocation  $(\mathbf{b}^*, \boldsymbol{\gamma}^*)$  is again a feasible solution to the optimization problem given by Eq. (1).

Therefore in order to show that the allocation  $(\mathbf{b}^*, \boldsymbol{\gamma}^*)$  leads to a strictly lower cost compared with the allocation  $(\mathbf{b}, \boldsymbol{\gamma})$ , it suffices to show that

$$\begin{aligned} & \left[ I_{m,j} b_{m,j} - \gamma_{m,j} \int_0^{b_{m,j}} D_j(b) db \right] \\ & + \left[ I_{n,j} b_{n,j} - \gamma_{n,j} \int_0^{b_{n,j}} D_j(b) db \right] \\ & > I_t q_j^* - (\gamma_{m,j}^* + \gamma_{n,j}^*) \int_0^{q_j^*} D_j(b) db \\ & = (I_{m,j} + I_{n,j}) q_j^* - (\gamma_{m,j} + \gamma_{n,j}) \int_0^{q_j^*} D_j(b) db, \end{aligned}$$

or

$$\begin{aligned} & I_{m,j}(b_{m,j} - q_j^*) - \gamma_{m,j} \int_{b_{m,j}}^{q_j^*} D_j(b) db \\ & + I_{n,j}(b_{n,j} - q_j^*) - \gamma_{n,j} \int_{q_j^*}^{b_{n,j}} D_j(b) db \\ & > 0. \end{aligned}$$

We have that

$$\begin{aligned} & I_{m,j}(b_{m,j} - q_j^*) - \gamma_{m,j} \int_{b_{m,j}}^{q_j^*} D_j(b) db \\ & + I_{n,j}(b_{n,j} - q_j^*) - \gamma_{n,j} \int_{q_j^*}^{b_{n,j}} D_j(b) db \\ & > I_{m,j}(b_{m,j} - q_j^*) + \gamma_{m,j} D_j(b_{m,j})(q_j^* - b_{m,j}) \\ & + I_{n,j}(b_{n,j} - q_j^*) - \gamma_{n,j} D_j(b_{n,j})(b_{n,j} - q_j^*) \\ & = I_{m,j}(b_{m,j} - q_j^*) + I_{m,j}(q_j^* - b_{m,j}) \\ & + I_{n,j}(b_{n,j} - q_j^*) - I_{n,j}(b_{n,j} - q_j^*) \\ & = 0. \end{aligned}$$

This implies that the allocation  $(\mathbf{b}^*, \boldsymbol{\gamma}^*)$  achieves a strictly lower cost compared to the allocation  $(\mathbf{b}, \boldsymbol{\gamma})$ . This leads to a contradiction as we assumed that the allocation  $(\mathbf{b}, \boldsymbol{\gamma})$  is an optimal allocation. The result of the lemma then follows.  $\square$

The following corollary follows immediately from Proposition 14.

**Corollary 2.** *Suppose that the supply curves  $D_j(b)$ ,  $j \in \mathcal{T}$ , are given by continuous functions. Then there exists an optimal solution  $\mathbf{b}^*, \boldsymbol{\gamma}^*$  to the optimization problem given by Eq. (1) such that the following is true. For each targeting group  $j = 1, \dots, N_r$  we have that*

$$b_{i,j}^* = q_j^*, \quad i \in \mathcal{B}_j(\mathbf{b}^*, \boldsymbol{\gamma}^*).$$

**Proposition 15.** *If the supply curves  $D_j(b)$ ,  $j \in \mathcal{T}$ , are given by continuous functions, then the following is true. Let  $(\mathbf{b}^*, \boldsymbol{\gamma}^*)$  be an optimal solution to the optimization problem given by Eq. (1). Then for each campaigns  $i = 1, \dots, N_c$  we have that*

$$D_m(b_{i,m}^*) = D_m(b_{i,n}^*) = D_m(p_i^*), \quad m, n \in \mathcal{A}_i(\mathbf{b}^*, \boldsymbol{\gamma}^*).$$

*Proof.* We prove the lemma by contradiction. That is we assume that  $(\mathbf{b}, \gamma)$  is an optimal solution to the optimization problem given by Eq. (1), and there exists a campaign  $i \in \mathcal{N}$  and targeting groups  $n, m \in \mathcal{A}_i(\mathbf{b}^*, \gamma^*)$  such that

$$D_m(b_{i,m}) < D_m(b_{i,n})$$

or

$$D_m(b_{i,m}) > D_m(b_{i,n}).$$

We then show that there exists an allocation  $(\mathbf{b}^*, \gamma^*)$  such that

$$b_{i,m}^* = b_{i,n}^* = p_i^*$$

which leads to a strictly lower cost compared with the allocation  $(\mathbf{b}, \gamma)$ . This leads to a contradiction as we assumed that the allocation  $(\mathbf{b}, \gamma)$  is an optimal allocation.

In the following we consider the case where

$$D_m(b_{i,m}) < D_m(b_{i,n}).$$

The case where

$$D_m(b_{i,m}) > D_m(b_{i,n}).$$

can be proved using the same argument.

To show the result we use the following definitions. Let

$$I_{i,m} = \gamma_{i,m} D_m(b_{i,m})$$

and

$$I_{i,n} = \gamma_{i,n} D_n(b_{i,n}).$$

Furthermore let  $p_i^*$  be such that

$$\gamma_{i,m} D_m(p_i^*) + \gamma_{i,n} D_n(p_i^*) = I_{i,m} + I_{i,n} = I_t.$$

Note that such a bid value  $p_i^*$  exists. Furthermore, as we have that

$$D_m(b_{i,m}) < D_m(b_{i,n}^*),$$

it follows that

$$D_m(b_{i,m}) < D_m(p_i^*) < D_m(b_{i,n}^*)$$

and

$$b_{i,m} < p_i^* < b_{i,n}.$$

Using these definitions, let  $(\mathbf{b}^*, \gamma^*)$  be the allocation that is identical to  $(\mathbf{b}, \gamma)$ , except that

$$b_{i,m}^* = p_i^*$$

and

$$b_{i,n}^* = p_i^*.$$

Compared with the allocation  $(\mathbf{b}, \gamma)$ , the new allocation  $(\mathbf{b}^*, \gamma^*)$  only affects (changes) the bid values that campaign  $i$  uses to obtain impressions from targeting group  $m$  and  $n$ . Furthermore, as we have that

$$\gamma_{i,m} D_m(p_i^*) + \gamma_{i,n} D_n(p_i^*) = I_{i,m} + I_{i,n} = I_t,$$

it follows the allocation  $(\mathbf{b}^*, \gamma^*)$  is again a feasible solution to the optimization problem given by Eq. (1).

Therefore in order to show that the allocation  $(\mathbf{b}^*, \gamma^*)$  leads to a strictly lower cost compared with the allocation  $(\mathbf{b}, \gamma)$ , it suffices to show that

$$\begin{aligned} & \left[ I_{m,j} b_{m,j} - \gamma_{m,j} \int_0^{b_{m,j}} D_j(b) db \right] \\ & + \left[ I_{n,j} b_{n,j} - \gamma_{n,j} \int_0^{b_{n,j}} D_j(b) db \right] \\ & > I_t p_i^* - \int_0^{q_j^*} [\gamma_{i,m}^* D_m(b) + \gamma_{i,n}^* D_n(b)] db \\ & = (I_{m,j} + I_{n,j}) p_i^* - \int_0^{q_j^*} [\gamma_{i,m} D_m(b) + \gamma_{i,n} D_n(b)] db \end{aligned}$$

or

$$\begin{aligned}
& I_{m,j}(b_{m,j} - p_i^*) + \gamma_{m,j} \int_{b_{m,j}}^{p_i^*} D_j(b) db \\
& + I_{n,j}(b_{n,j} - p_i^*) - \gamma_{n,j} \int_{p_i^*}^{b_{n,j}} D_j(b) db \\
& > 0.
\end{aligned}$$

We have that

$$\begin{aligned}
& I_{m,j}(b_{m,j} - p_i^*) + \gamma_{m,j} \int_{b_{m,j}}^{p_i^*} D_j(b) db \\
& + I_{n,j}(b_{n,j} - p_i^*) - \gamma_{n,j} \int_{p_i^*}^{b_{n,j}} D_j(b) db \\
& > I_{m,j}(b_{m,j} - p_i^*) + \gamma_{m,j} D_m(b_{m,j})(p_i^* - b_{m,j}) \\
& + I_{n,j}(b_{n,j} - p_i^*) - \gamma_{n,j} D_n(b_{n,j})(b_{n,j} - p_i^*) \\
& = 0.
\end{aligned}$$

This implies that the allocation  $(\mathbf{b}^*, \gamma^*)$  achieves a strictly lower cost compared to the allocation  $(\mathbf{b}, \gamma)$ . This leads to a contradiction as we assumed that the allocation  $(\mathbf{b}, \gamma)$  is an optimal allocation. The result of the lemma then follows.  $\square$

The following corollary follows immediately from Proposition 15.

**Corollary 3.** *Suppose that the supply curves  $D_j(b)$ ,  $j \in \mathcal{T}$ , are given by continuous functions. Then there exists an optimal solution  $\mathbf{b}^*, \gamma^*$  to the optimization problem given by Eq. (1) such that the following is true. For each campaign  $i = 1, \dots, N_c$  we have that*

$$b_{i,j}^* = p_i^*, \quad j \in \mathcal{A}_i(\mathbf{b}^*, \gamma^*).$$

## C.2 Decomposition of an Optimal Solution

In this appendix, we prove Proposition 5. Recall the statement of the proposition.

**Proposition 16.** *Suppose that the supply curves  $D_j(b)$ ,  $j \in \mathcal{T}$ , are given by continuous functions. Then there exists an optimal solution  $(\mathbf{b}^*, \gamma^*)$  to the optimization problem given by Eq. (1) such that the following is true. There exists a set of bid values  $\mathcal{P}(\mathbf{b}^*, \gamma^*)$  such that with every bid value  $p^* \in \mathcal{P}(\mathbf{b}^*, \gamma^*)$  we can associate component*

$$\mathcal{C}(p^*, \mathbf{b}^*, \gamma^*) = (\mathcal{C}_c(p^*, \mathbf{b}^*, \gamma^*), \mathcal{C}_t(p^*, \mathbf{b}^*, \gamma^*))$$

where  $\mathcal{C}_c(p^*, \mathbf{b}^*, \gamma^*)$  is a set of campaigns, and  $\mathcal{C}_t(p^*, \mathbf{b}^*, \gamma^*)$  is a set of targeting groups, such that

$$\mathcal{A}_i(\mathbf{b}^*, \gamma^*) \cap \mathcal{C}_t(p^*, \mathbf{b}^*, \gamma^*) = \mathcal{A}_i(\mathbf{b}^*, \gamma^*), \quad i \in \mathcal{C}_c(p^*, \mathbf{b}^*, \gamma^*),$$

and

$$\mathcal{B}_j(\mathbf{b}^*, \gamma^*) \cap \mathcal{C}_c(p^*, \mathbf{b}^*, \gamma^*) = \mathcal{B}_j(\mathbf{b}^*, \gamma^*), \quad j \in \mathcal{C}_t(p^*, \mathbf{b}^*, \gamma^*),$$

as well as

$$b_{i,j}^* = p^*, \quad i \in \mathcal{C}_c(p^*, \mathbf{b}^*, \gamma^*), j \in \mathcal{A}_i(\mathbf{b}^*, \gamma^*)$$

and

$$b_{i,j}^* = p^*, \quad j \in \mathcal{C}_t(p^*, \mathbf{b}^*, \gamma^*), i \in \mathcal{B}_j(\mathbf{b}^*, \gamma^*),$$

as well as

$$\bigcup_{p^* \in \mathcal{P}(\mathbf{b}^*, \gamma^*)} \mathcal{C}_c(p^*, \mathbf{b}^*, \gamma^*) = \mathcal{N}$$

and

$$\bigcup_{p^* \in \mathcal{P}(\mathbf{b}^*, \gamma^*)} \mathcal{C}_t(p^*, \mathbf{b}^*, \gamma^*) = \mathcal{T}.$$

*Proof.* Let  $(\mathbf{b}^*, \gamma^*)$  be an optimal solution to the optimization problem given by Eq. (1) such that for each campaign  $i = 1, \dots, N_c$  we have that

$$b_{i,j}^* = p_i^*, \quad j \in \mathcal{A}_i(\mathbf{b}^*, \gamma^*).$$

Note that such an optimal solution exists by Proposition 3.

We then iteratively construct a set  $\mathcal{P}(\mathbf{b}^*, \gamma^*)$ , and a set of components

$$\mathcal{C}(p^*, \mathbf{b}^*, \gamma^*) = (\mathcal{C}_c(p^*, \mathbf{b}^*, \gamma^*), \mathcal{C}_t(p^*, \mathbf{b}^*, \gamma^*)), \quad p^* \in \mathcal{P}(\mathbf{b}^*, \gamma^*),$$

as given in the statement of the proposition as follows.

Let

$$\mathcal{N}^{(0)} = \mathcal{N} = \{1, \dots, N_c\}$$

and

$$\mathcal{T}^{(0)} = \mathcal{T} = \{1, \dots, N_r\}$$

Furthermore, set

$$\mathcal{P}(\mathbf{b}^*, \gamma^*) = \emptyset.$$

If the iteration hasn't stopped at iteration  $k - 1$ ,  $k \geq 1$ , then we proceed with iteration step  $k$  as follows. If

$$\mathcal{N}^{(k-1)} = \emptyset,$$

then we stop the iteration. Otherwise, we set

$$\mathcal{N}^{(k)} = \mathcal{N}^{(k-1)}$$

and

$$\mathcal{T}^{(k)} = \mathcal{T}^{(k-1)}.$$

As we have in this case that

$$\mathcal{N}^{(k)} \neq \emptyset,$$

we can pick campaign  $i$  from the set  $\mathcal{N}^{(k)}$ . Let the bid value  $p^*$  be such that for this campaign  $i$  we have that

$$b_{i,j}^* = p^*, \quad j \in \mathcal{A}_i(\mathbf{b}^*, \gamma^*).$$

Note that such a  $p^*$  exists by assumption. We then consider the following two cases. If we have that

$$p^* \notin \mathcal{P}(\mathbf{b}^*, \gamma^*),$$

then we set

$$\mathcal{P}(\mathbf{b}^*, \gamma^*) \leftarrow \mathcal{P}(\mathbf{b}^*, \gamma^*) \cup \{p^*\},$$

and create a new set

$$\mathcal{C}(p^*, \mathbf{b}^*, \gamma^*) = (\mathcal{C}_c(p^*, \mathbf{b}^*, \gamma^*), \mathcal{C}_t(p^*, \mathbf{b}^*, \gamma^*))$$

where

$$\mathcal{C}_c(p^*, \mathbf{b}^*, \gamma^*) = \{i\}$$

and

$$\mathcal{C}_t(p^*, \mathbf{b}^*, \gamma^*) = \mathcal{A}_i(\mathbf{b}^*, \gamma^*).$$

Finally, we set

$$\mathcal{A} = \{i\}$$

and repeat the following steps until we have that

$$\mathcal{A} = \emptyset.$$

1. Pick a campaign  $i \in \mathcal{A}$  and set

$$\mathcal{A} \leftarrow \mathcal{A} \setminus \{i\}$$

$$\mathcal{N}^{(k)} \leftarrow \mathcal{N}^{(k)} \setminus \{i\}.$$

2. Update component

$$\mathcal{C}(p^*, \mathbf{b}^*, \gamma^*) = (\mathcal{C}_c(p^*, \mathbf{b}^*, \gamma^*), \mathcal{C}_t(p^*, \mathbf{b}^*, \gamma^*))$$

of the bid value  $p^*$  to

$$\mathcal{C}_c(p^*, \mathbf{b}^*, \gamma^*) \leftarrow \mathcal{C}_c(p^*, \mathbf{b}^*, \gamma^*) \cup \{i\}$$

and

$$\mathcal{C}_t(p^*, \mathbf{b}^*, \gamma^*) \leftarrow \mathcal{C}_t(p^*, \mathbf{b}^*, \gamma^*) \cup \mathcal{A}_i(\mathbf{b}^*, \gamma^*).$$

3. Set

$$\mathcal{N}^{(k)} \leftarrow \mathcal{N}^{(k)} \setminus \{i\}$$

and

$$\mathcal{T}^{(k)} \leftarrow \mathcal{T}^{(k)} \setminus \mathcal{A}_i(\mathbf{b}^*, \gamma^*).$$

4. For all targeting groups  $j \in \mathcal{A}_i(\mathbf{b}^*, \gamma^*)$  set

$$b_{m,j}^* = p^*, \quad m \in \mathcal{B}_j(\mathbf{b}^*, \gamma^*).$$

5. For each targeting group

$$j \in \mathcal{A}_i(\mathbf{b}^*, \gamma^*)$$

set

$$\mathcal{A} \leftarrow \mathcal{A} \cup \left( \mathcal{B}_j(\mathbf{b}^*, \gamma^*) \cap \mathcal{N}^{(k)} \right).$$

6. For all campaigns  $m \in \mathcal{A}$  set

$$b_{m,j}^* = p^*, \quad j \in \mathcal{A}_i(\mathbf{b}^*, \gamma^*).$$

This completes the iteration step  $k$ .

We make the following observation regarding the above update rules. By construction (initialization and Step (6)), we have for all targeting group  $j \in \mathcal{A}_i(\mathbf{b}^*, \gamma^*)$  in Step (4) that

$$b_{i,j}^* = p^*.$$

Using Lemma 14, we then have for each targeting group  $j \in \mathcal{A}_i(\mathbf{b}^*, \gamma^*)$  in Step (4) that

$$D_j(b_{m,j}^*) = D_j(b_{i,j}^*) = D_j(p^*), \quad m \in \mathcal{B}_j(\mathbf{b}^*, \gamma^*).$$

This implies that changing the bid value  $b_{m,j}^*$  to  $p^*$  does not affect that number of impressions that campaign  $m$  obtains from targeting group  $j$ , nor the cost for obtaining the impressions. As such, the allocation (*opt*) is still an optimal allocation if we makes this change in the bid value  $b_{m,j}^*$ .

Similarly, by construction (Step (4)) we have for all campaigns  $m \in \mathcal{A}$  in Step (6) that there exist a targeting group  $n \in \mathcal{A}_i(\mathbf{b}^*, \gamma^*)$  such that

$$b_{m,n}^* = p^*.$$

Using Lemma 15, we then have for each campaign  $m \in \mathcal{A}$  in Step (6) that

$$D_j(b_{m,j}^*) = D_j(b_{m,n}^*) = D_j(p^*), \quad j \in \mathcal{A}_i(\mathbf{b}^*, \gamma^*).$$

This implies that changing the bid value  $b_{m,j}^*$  to  $p^*$  does not affect that number of impressions that campaign  $m$  obtains from targeting group  $j$ , nor the cost for obtaining the impressions. As such, the allocation (*opt*) is still an optimal allocation if we makes this change in the bid value  $b_{m,j}^*$ .

By construction of the set  $\mathcal{P}(\mathbf{b}^*, \gamma^*)$  and the components

$$\mathcal{C}(p^*, \mathbf{b}^*, \gamma^*), \quad p \in \mathcal{P}(\mathbf{b}^*, \gamma^*),$$

we have that for each campaign  $i = 1, \dots, N_c$ , that there exists a unique  $p^* \in \mathcal{P}(\mathbf{b}^*, \gamma^*)$  such that

$$i \in \mathcal{C}_c(p^*, \mathbf{b}^*, \gamma^*).$$

Similarly, we have for each targeting group  $j = 1, \dots, N_r$ , that there exists a unique  $p^* \in \mathcal{P}(\mathbf{b}^*, \gamma^*)$  such that

$$j \in \mathcal{C}_t(p^*, \mathbf{b}^*, \gamma^*).$$

Furthermore, by construction we have that for all components  $\mathcal{C}(p^*, \mathbf{b}^*, \gamma^*)$ ,  $p^* \in \mathcal{P}(\mathbf{b}^*, \gamma^*)$ , that

$$\mathcal{A}_i(\mathbf{b}^*, \gamma^*) \cap \mathcal{C}_t(p^*, \mathbf{b}^*, \gamma^*) = \mathcal{A}_i(\mathbf{b}^*, \gamma^*), \quad i \in \mathcal{C}_c(p^*, \mathbf{b}^*, \gamma^*),$$

and

$$\mathcal{B}_j(\mathbf{b}^*, \gamma^*) \cap \mathcal{C}_c(p^*, \mathbf{b}^*, \gamma^*) = \mathcal{B}_j(\mathbf{b}^*, \gamma^*), \quad j \in \mathcal{C}_t(p^*, \mathbf{b}^*, \gamma^*),$$

as well as

$$b_{i,j}^* = p^*, \quad i \in \mathcal{C}_c(p^*, \mathbf{b}^*, \gamma^*), j \in \mathcal{A}_i(\mathbf{b}^*, \gamma^*)$$

and

$$b_{i,j}^* = p^*, \quad j \in \mathcal{C}_t(p^*, \mathbf{b}^*, \gamma^*), i \in \mathcal{B}_j(\mathbf{b}^*, \gamma^*).$$

This establishes the result of the proposition.  $\square$

### C.3 Sufficient Conditions

We next provide sufficient conditions for an allocation  $(\mathbf{b}, \gamma)$  that has the structure of an optimal solution as given in Proposition 5 to indeed be an optimal solution.

We proceed as follows. We first consider the special case where we are given an allocation  $(\mathbf{b}^*, \gamma^*)$  as in Proposition 5 for which we have that

$$\mathcal{P}(\mathbf{b}^*, \gamma^*) = \{p^*\},$$

the allocation  $(\mathbf{b}^*, \gamma^*)$  consists of a single component  $\mathcal{C}(p^*, \mathbf{b}^*, \gamma^*)$ . Note that we do not assume that the allocation  $(\mathbf{b}^*, \gamma^*)$  is an optimal solution, but only that the allocation has a structure as the allocation given in Proposition 5.

**Proposition 17.** *If the supply curves  $D_j(b)$ ,  $j \in \mathcal{T}$ , are given by continuous functions, then the following is true. Let  $(\mathbf{b}^*, \gamma^*)$  be an allocation such that for all campaigns  $i = 1, \dots, N_c$  we have that*

$$b_{i,j}^* = p^*, \quad j \in \mathcal{A}_i(\mathbf{b}^*, \gamma^*),$$

as given in Proposition 5, i.e. we have that

$$\mathcal{P}(\mathbf{b}^*, \gamma^*) = \{p^*\}$$

and

$$\mathcal{C}_c(p^*, \mathbf{b}^*, \gamma^*) = \mathcal{N}.$$

If we have that

$$\sum_{j \in \mathcal{A}_i(\mathbf{b}^*, \gamma^*)} \gamma_{i,j}^* D_j(p^*) = I_i, \quad i \in \mathcal{C}_c(p^*, \mathbf{b}^*, \gamma^*),$$

and

$$\sum_{i \in \mathcal{C}_c(p^*, \mathbf{b}^*, \gamma^*)} \sum_{j \in \mathcal{A}_i(\mathbf{b}^*, \gamma^*)} D_j(p^*) = \sum_{i \in \mathcal{C}_c(p^*, \mathbf{b}^*, \gamma^*)} I_i,$$

then the allocation  $(\mathbf{b}^*, \gamma^*)$  is an optimal solution to the optimization problem given in Eq. (1) with cost

$$\sum_{i \in \mathcal{C}_c(p^*, \mathbf{b}^*, \gamma^*)} I_i p^* - \int_0^{p^*} \sum_{j \in \mathcal{A}_i(\mathbf{b}^*, \gamma^*)} D_j(b) db.$$

*Proof.* This result can be proved using the same argument as the one given in the proof for Proposition 10.  $\square$

Next we consider a more general, but still a special case. In particular we consider an allocation  $(\mathbf{b}^*, \gamma^*)$  as given by Proposition 5 such that for  $p^*, q^* \in \mathcal{P}(\mathbf{b}^*, \gamma^*)$ ,

$$q^* < p^*$$

we have that

$$\mathcal{A}_i \cap \mathcal{C}_t(q^*, \mathbf{b}^*, \gamma^*) = \emptyset, \quad i \in \mathcal{C}_c(p^*, \mathbf{b}^*, \gamma^*).$$

As it turns out, this case is the case that is relevant for our analysis of the proposed algorithm to compute an optimal allocation  $(\mathbf{b}^*, \gamma^*)$  in Section 7.1.

Again we do not assume that the allocation  $(\mathbf{b}^*, \gamma^*)$  is an optimal solution, but only that the allocation has a structure as the allocation given in Proposition 5.

**Proposition 18.** *If the supply curves  $D_j(b)$ ,  $j \in \mathcal{T}$ , are given by continuous functions, then the following is true. Let  $(\mathbf{b}^*, \gamma^*)$  be an allocation as given in Proposition 5. If for all  $p^* \in \mathcal{P}(\mathbf{b}^*, \gamma^*)$  we have*

a)

$$\sum_{j \in \mathcal{A}_i(\mathbf{b}^*, \gamma^*)} \gamma_{i,j}^* D_j(p^*) = I_i, \quad i \in \mathcal{C}_c(p^*, \mathbf{b}^*, \gamma^*),$$

and

$$\sum_{i \in \mathcal{C}_c(p^*, \mathbf{b}^*, \gamma^*)} \sum_{j \in \mathcal{A}_i(\mathbf{b}^*, \gamma^*)} D_j(p^*) = \sum_{i \in \mathcal{C}_c(p^*, \mathbf{b}^*, \gamma^*)} I_i,$$

b)

$$\sum_{i \in \mathcal{C}_c(p^*, \mathbf{b}^*, \gamma^*)} \sum_{j \in \mathcal{A}_i \cap \mathcal{T}_0} D_j(p) = 0, \quad 0 \leq p < p^*,$$

where

$$\mathcal{T}_0 = \left\{ j \in \mathcal{T} \mid \sum_{i \in \mathcal{B}_j} \gamma_{i,j}^* D_j(b_{i,j}^*) = 0 \right\},$$

c) for  $q^* < p^*$ ,  $q^*, p^* \in \mathcal{P}(\mathbf{b}^*, \gamma^*)$ , that

$$\mathcal{A}_i \cap \mathcal{C}_t(q^*, \mathbf{b}^*, \gamma^*) = \emptyset, \quad i \in \mathcal{C}_c(p^*, \mathbf{b}^*, \gamma^*),$$

then the allocation  $(\mathbf{b}^*, \gamma^*)$  is an optimal solution to the optimization problem given by Eq (1) with cost

$$\sum_{p^* \in \mathcal{P}(\mathbf{b}^*, \gamma^*)} \left[ \sum_{i \in \mathcal{C}_c(p^*, \mathbf{b}^*, \gamma^*)} I_i p^* - \int_0^{p^*} \sum_{j \in \mathcal{A}_i(\mathbf{b}^*, \gamma^*)} D_j(b) db \right].$$

*Proof.* We use the result of Proposition 10 to prove the proposition. More precisely, using Proposition 10 we have that the cost of any solution to the optimization problem of Eq. (1) that obtains the an amount of

$$\sum_{i \in \mathcal{C}_c(p^*, \mathbf{b}^*, \gamma^*)} I_i$$

impressions from the component

$$\mathcal{C}(p^*, \mathbf{b}^*, \gamma^*), \quad p \in \mathcal{P}(\mathbf{b}^*, \gamma^*),$$

is lower-bounded by

$$\sum_{i \in \mathcal{C}_c(p^*, \mathbf{b}^*, \gamma^*)} I_i p^* - \int_0^{p^*} \sum_{j \in \mathcal{A}_i(\mathbf{b}^*, \gamma^*)} D_j(b) db.$$

As a result, the allocation  $(\mathbf{b}^*, \gamma^*)$  is an optimal allocation for each individual component  $\mathcal{C}(p^*, \mathbf{b}^*, \gamma^*)$ ,  $p \in \mathcal{P}(\mathbf{b}^*, \gamma^*)$ .

Therefore, in order to proof the result, we have to show for each component

$$\mathcal{C}(p, \mathbf{b}, \gamma), \quad p \in \mathcal{P}(\mathbf{b}^*, \gamma^*),$$

that is not possible reduce the cost for getting the required impressions for the campaigns in this component by shifting some of these impressions to another component. We show this result by induction as follows.

To show that this is indeed the case, we use the following notation. For each component  $\mathcal{C}(p^*, \mathbf{b}^*, \gamma^*)$ ,  $p \in \mathcal{P}(\mathbf{b}^*, \gamma^*)$ , let

$$I_{p^*} = \sum_{i \in \mathcal{C}_c(p^*, \mathbf{b}^*, \gamma^*)} I_i$$

and

$$D_{-p^*}(b) = \sum_{j \in \mathcal{C}_i(p^*, \mathbf{b}^*, \gamma^*)} D_j(b), \quad 0 \leq b.$$

We first show that for the component corresponding to the highest bid value  $p^* \in \mathcal{P}(\mathbf{b}^*, \gamma^*)$ , that we can not reduce the for getting the require  $I_{p^*}$  impressions by shifting some of the impressions to another component.

We then do an induction step where we assume that it is not possible to reduce the the cost for obtaining the impressions for the components corresponding to the  $n$  highest bid values  $p^* \in \mathcal{P}(\mathbf{b}^*, \gamma^*)$  by shifting impressions to another component, and then show that in this case it is also not possible to reduce the the cost for obtaining the impressions for the components corresponding to the  $(n + 1)$  highest bid values  $p^* \in \mathcal{P}(\mathbf{b}^*, \gamma^*)$  by shifting some of the impressions to another component.

Let  $p^*$  be the the highest bid value in the set  $\mathcal{P}(\mathbf{b}^*, \gamma^*)$ , and suppose that we move an amount of  $\Delta_I$  of the impressions  $I_{p^*}$  to another component. As we have for  $q^* \in \mathcal{P}(\mathbf{b}^*, \gamma^*)$ ,  $q^* < p^*$ , that

$$\mathcal{A}_i \cap \mathcal{C}_t(q^*, \mathbf{b}^*, \gamma^*) = \emptyset, \quad i \in \mathcal{C}_c(p^*, \mathbf{b}^*, \gamma^*),$$

the only option is to move the  $\Delta_I$  impressions to the set  $\mathcal{T}_0$ , i.e. the set

$$\mathcal{T}_0 = \left\{ j \in \mathcal{T} \mid \sum_{i \in \mathcal{B}_j} \gamma_{i,j}^* D_j(b_{i,j}^*) = 0 \right\}.$$

Let

$$\mathcal{T}_0(p^*) = \mathcal{T}_0 \cap \cup i \in \mathcal{C}_c(p^*, \mathbf{b}^*, \gamma^*) A_i,$$

i.e. the set  $\mathcal{T}_0(p)$  contains all the targeting groups  $j \in \mathcal{T}_0$  from which at least one campaign  $i \in \mathcal{C}_c(p^*, \mathbf{b}^*, \gamma^*)$  can obtain impressions. Furthermore, let  $D_0(b)$  be given by

$$D_0(b) = \sum_{j \in \mathcal{T}_0(p^*)} D_j(b).$$

Note by definition that we have that

$$D_0(p^*) = 0.$$

Finally, suppose that there exists a bid value  $q$  such that

$$D_q(q) = \Delta_I$$

and

$$D_0(b) < \Delta_I, \quad 0 \leq b < q.$$

Using these definitions and the fact that

$$D_0(p) = 0$$

and

$$D_0^{(-)}(q) \leq \Delta_I,$$

we have that that the cost for getting  $\Delta_I$  impressions from the targeting groups in the set  $\mathcal{T}_0(p)$  is at least

$$\begin{aligned} & \Delta_I q - \int_{p^*}^q D_0(b) db \\ & \geq \Delta_I q - D_0^{(-)}(q)(q - p^*) \\ & \geq \Delta_I q - \Delta_I(q - p^*) \\ & = \Delta_I p^*. \end{aligned}$$

Similarly using the bid value  $p'$  such that

$$D_{p^*}(p') = I_{p^*} - \Delta_I,$$

the reduction in the cost for reducing the number of impressions obtained from component  $\mathcal{C}(p^*, \mathbf{b}^*, \gamma^*)$  by  $\Delta_I$  is at most

$$\begin{aligned}
& I_{p^*} p^* - \int_0^{p^*} D_p(b) db \\
& - (I_{p^*} - \Delta_I) p' + \int_0^{p'} D_p(b) db \\
= & I(\mathcal{C}(p, \mathbf{b}, \gamma))(p^* - p') + \Delta_I p' \\
& - \int_{p'}^{p^*} D_p(b) db \\
\leq & I(\mathcal{C}(p, \mathbf{b}, \gamma))(p^* - p') + \Delta_I p' \\
& - D_p(p')(p^* - p') \\
= & (I_{p^*} - D_p(p'))(p^* - p') + \Delta_I p' \\
\leq & \Delta_I (p^* - p') + \Delta_I p' \\
= & \Delta_I p^*.
\end{aligned}$$

Therefore, we obtain that it is not possible to reduce the cost for getting  $I - p^*$  impressions for the campaigns in component  $\mathcal{C}(p^*, \mathbf{b}^*, \gamma^*)$  by shifting some of the impressions to targeting groups in the set  $\mathcal{T}_0(p)$ . This establishes the result that the allocation  $(\mathbf{b}^*, \gamma^*)$  is an optimal allocation for getting  $I_{p^*}$  impressions for the campaigns in component  $\mathcal{C}(p^*, \mathbf{b}^*, \gamma^*)$ .

Having established the result for the component with the highest bid value, we do an induction step. That is, we assume that it is not possible to reduce the the for obtaining the impressions for the campaigns in the components corresponding to the  $n$  highest bid values in  $\mathcal{P}(\mathbf{b}^*, \gamma^*)$ , and then show that it is not possible to reduce the cost for obtaining the impressions for the campaigns in the components corresponding to the  $(n + 1)$  highest bid values in  $\mathcal{P}(\mathbf{b}^*, \gamma^*)$ .

Let  $p^*$  be the  $(n + 1)$ th highest bid value and let  $\mathcal{C}(p^*, \mathbf{b}^*, \gamma^*)$  be the corresponding component. Note that in this case, the campaigns  $i \in \mathcal{C}_c(p^*, \mathbf{b}^*, \gamma^*)$  that belong to this component can only get impressions from targeting groups that belong to the set  $\mathcal{C}_t(p^*, \mathbf{b}^*, \gamma^*)$ , from targeting groups that belong to a component with a higher bid value  $q^* \in \mathcal{P}(\mathbf{b}^*, \gamma^*)$ , or from the set of targeting groups  $\mathcal{T}_0(p^*)$  given by

$$\mathcal{T}_0(p^*) = \mathcal{T}_0 \cap \cup i \in \mathcal{C}_c(p, \mathbf{b}, \gamma) A_i.$$

As we have by assumption for  $D_0(b)$  given by

$$D_0(b) = \sum_{j \in \mathcal{T}_0(p^*)} D_j(b),$$

that

$$D_0(p) = 0,$$

it follows from the same argument as given for the case of the component with the highest bid value, that it is not possible to reduce the cost for obtaining the impressions for the campaigns in the components corresponding to the  $(n + 1)$  highest bid values  $p \in \mathcal{P}(\mathbf{b}^*, \gamma^*)$  by shifting some impressions from the component  $\mathcal{C}(p^*, \mathbf{b}^*, \gamma^*)$  to the set  $\mathcal{T}_0(p^*)$ .

It remains to consider the case where we shift some impressions from the component  $\mathcal{C}(p^*, \mathbf{b}^*, \gamma^*)$  to a component  $\mathcal{C}(q^*, \mathbf{b}^*, \gamma^*)$  with a higher bid value  $q^*$ .

For this case, we use the following definitons. Let

$$D_{p^*}(b) = \sum_{j \in \mathcal{C}_t(p^*, \mathbf{b}^*, \gamma^*)} D_j(b)$$

and

$$D_{q^*}(b) = \sum_{j \in \mathcal{C}_t(q^*, \mathbf{b}^*, \gamma^*)} D_j(b).$$

Furthermore, let  $p'$  be such that

$$D_p(p') = I_{p^*} - \Delta_I,$$

and let  $q'$  be such that

$$D_{q^*}(q') = I_{q^*} + \Delta_I$$

Under these definitions, by shifting an amount of  $\Delta_I$  impressions from component  $\mathcal{C}(p^*, \mathbf{b}^*, \gamma^*)$  to the component  $\mathcal{C}(q^*, \mathbf{b}^*, \gamma^*)$ , we can reduce the cost by at most an amount given by

$$\begin{aligned} & I_{p^*} p - \int_0^{p^*} D_p(b) db \\ & - (I_{p^*} - \Delta_I) p' + \int_0^{p'} D_p(b) db \\ = & I_{p^*} (p^* - p') + \Delta_I p' \\ & - \int_{p'}^{p^*} D_p(b) db \\ \leq & I_{p^*} (p^* - p') + \Delta_I p' \\ & - D_p(p')(p^* - p') \\ = & (I_{p^*} - D_p(p'))(p^* - p') + \Delta_I p' \\ \leq & \Delta_I (p^* - p') + \Delta_I p' \\ = & \Delta_I p^*. \end{aligned}$$

Similarly, by shifting an amount of  $\Delta_I$  impressions from component  $\mathcal{C}(p^*, \mathbf{b}^*, \gamma^*)$  to  $\mathcal{C}(q^*, \mathbf{b}^*, \gamma^*)$ , we increase the cost by at least an amount given by

$$\begin{aligned} & -I_{q^*} q^* + \int_0^{q^*} D_q(b) db \\ & (I_{q^*} + \Delta_I) q' - \int_0^{q'} D_q(b) db \\ = & I_{q^*} (q' - q^*) + \Delta_I q' \\ & - \int_{q^*}^{q'} D_q(b) db \\ \geq & I_{q^*} (q' - q^*) + \Delta_I q' \\ & - D_q^{(-)}(q')(q' - q^*) \\ = & (I_{q^*} + \Delta_I - D_q^{(-)}(q'))(q' - q^*) + \Delta_I q^*, \\ \geq & \Delta_I q^*. \end{aligned}$$

As we have that

$$p^* < q^*,$$

it follows that we can not lower the by shifting some impressions from component  $\mathcal{C}(p^*, \mathbf{b}^*, \gamma^*)$  to component  $\mathcal{C}(q^*, \mathbf{b}^*, \gamma^*)$ .

This establishes the that the allocation  $(\mathbf{b}^*, \gamma^*)$  is optimal for getting the  $I_{p^*}$  impressions for the campaigns in the components  $\mathcal{C}(p^*, \mathbf{b}^*, \gamma^*)$  with the  $(n + 1)$  highest bid values  $p^* \in \mathcal{P}(\mathbf{b}^*, \gamma^*)$ . This completes the induction step, and the proof of the proposition.  $\square$

## D Properties of Solution $(\mathbf{b}^*, \gamma^*)$ obtained by Algorithm

In this appendix we derive properties of the set of components  $(\hat{\mathcal{C}}_c, \hat{\mathcal{C}}_t, p, \mathbf{t})$ ,  $p \in \mathcal{P}$  obtained by the algorithm of Section 7.

The following two results that follow directly from the definition of the the algorithm.

**Lemma 5.** *Let  $(\hat{\mathcal{C}}_c, \hat{\mathcal{C}}_t, p, \mathbf{t})$ ,  $p \in \mathcal{P}$  be the set of components obtained through the algorithm given by Eq. (10). Then we have that*

$$\bigcup_{p \in \mathcal{P}} \hat{\mathcal{C}}_c = \mathcal{N}$$

and

$$\bigcup_{p \in \mathcal{P}} \hat{\mathcal{C}}_t = \mathcal{T}.$$

Lemma 5 states that each campaign  $i \in \mathcal{N}$ , and each targeting group  $j \in \mathcal{T}$ , belongs to a component  $(\hat{\mathcal{C}}_c, \hat{\mathcal{C}}_t, p, \mathbf{t})$ ,  $p \in \mathcal{P}$ .

**Lemma 6.** *Let  $(\hat{\mathcal{C}}_c, \hat{\mathcal{C}}_t, q, \mathbf{t})$ ,  $q \in \mathcal{P}$  be the set of components obtained through the algorithm given by Eq. (10), let  $p \in \mathcal{P}$  be the lowest bid value in the set  $\mathcal{P}$ , and let  $(\hat{\mathcal{C}}_c, \hat{\mathcal{C}}_t, p, \mathbf{t})$  be the corresponding component. Then we have that*

$$\mathcal{T}_0(\mathbf{b}^*, \gamma^*) \subseteq \hat{\mathcal{C}}_t,$$

where

$$\mathcal{T}_0(\mathbf{b}^*, \gamma^*) = \left\{ j \in \mathcal{T} \mid \sum_{q \in \mathcal{P}} \sum_{i \in \mathcal{B}_j \cap \hat{\mathcal{C}}_t} t_{i,j} D_j(q) = 0 \right\}.$$

Next we consider one recursion step of the algorithm. That is, we are given a set  $\mathcal{C}_c$  of campaigns, and a set  $\mathcal{C}_t$  of targeting groups, and use the function  $SPLIT(\mathcal{C}_c, \mathcal{C}_t, t^*, p)$  described in Subsection 7.1 to produce two sets  $(\mathcal{C}_{c,1}, \mathcal{C}_{t,1})$  and  $(\mathcal{C}_{c,2}, \mathcal{C}_{t,2})$ .

**Lemma 7.** *Let  $\mathcal{C}_c \subseteq \mathcal{N}$  be a set of campaigns, and  $\mathcal{C}_t \subseteq \mathcal{T}$  be a set of targeting groups. Furthermore, suppose that the function  $SPLIT(\mathcal{C}_c, \mathcal{C}_t, t^*, p)$  in Algorithm 1 returns the sets  $(\mathcal{C}_{c,1}, \mathcal{C}_{t,1})$  and  $(\mathcal{C}_{c,2}, \mathcal{C}_{t,2})$ , let  $p$  be the bid value used in the function  $ALLOC(\mathcal{C}_c, \mathcal{C}_t, p)$ , and let  $\mathbf{t}^*$  be an optimal solution returned by the function  $ALLOC(\mathcal{C}_c, \mathcal{C}_t, p)$ . Then we have that*

$$\mathcal{A}_i \cap \mathcal{C}_{t,2} = \emptyset, \quad i \in \mathcal{C}_{c,1},$$

as well as

$$t_{i,j}^* = 0, \quad i \in \mathcal{C}_c^2, j \in \mathcal{C}_c^1,$$

and

$$\sum_{i \in \mathcal{B}_j \cap \mathcal{C}_{c,1}} t_{i,j}^* = 1, \quad j \in \mathcal{C}_{t,1}.$$

As the minimization problem given by Eq. (6) is a convex optimization problem. The result of Lemma 7 follows directly from the first order conditions of an optimal solution  $\mathbf{t}^*$  for the optimization problem given by Eq. (6).

**Lemma 8.** *Let  $\mathcal{C}_c \subseteq \mathcal{N}$  be a set of campaigns, and  $\mathcal{C}_t \subseteq \mathcal{T}$  be a set of targeting groups. Furthermore, suppose that the function  $SPLIT(\mathcal{C}_c, \mathcal{C}_t, t^*, p)$  in Algorithm 1 returns the sets  $(\mathcal{C}_{c,1}, \mathcal{C}_{t,1})$  and  $(\mathcal{C}_{c,2}, \mathcal{C}_{t,2})$ , let  $p$  be the bid value used in the function  $ALLOC(\mathcal{C}_c, \mathcal{C}_t, p)$ , and let  $\mathbf{t}^*$  be an optimal solution returned by the function  $ALLOC(\mathcal{C}_c, \mathcal{C}_t, p)$ .*

*Then for all components  $(\hat{\mathcal{C}}_c, \hat{\mathcal{C}}_t, q, \mathbf{t})$  that are obtained from the recursion  $REC(\mathcal{C}_{c,2}, \mathcal{C}_{t,2}, \emptyset, \emptyset)$ , we have that*

$$q \leq p.$$

*Proof.* Suppose that we obtain  $K$  components from the recursion  $REC(\mathcal{C}_{c,2}, \mathcal{C}_{t,2}, \emptyset, \emptyset)$ . We then number the components and create the set  $\{(\hat{\mathcal{C}}_c^{(k)}, \hat{\mathcal{C}}_t^{(k)}, q^{(k)}, \mathbf{t}^{(k)})\}_{k=1}^K$  such that for a given  $k$  we have that

$$\mathcal{A}_i \cap \hat{\mathcal{C}}_t^{(l)} = \emptyset, \quad i \in \hat{\mathcal{C}}_c^{(k)}, l = k + 1, \dots, K.$$

Note that we can create this ordering by Lemma 7.

Recall that we say that the optimization problem

$$ALLOC(\mathcal{C}_c, \mathcal{C}_t, p)$$

has a feasible solution if there exists a solution  $\mathbf{t}^*$  such that

$$\sum_{j \in \mathcal{A}_i \cap \mathcal{C}_t} t_{i,j}^* D_j(p) = I_i, \quad i \in \mathcal{C}_c.$$

In order to show the result of the lemma, we have to show that for all  $k = 1, \dots, K$  we have that the optimization problem  $ALLOC(\hat{C}_c^{(k)}, \hat{C}_t^{(k)}, p)$ , where  $p$  is the bid value given in the statement of the lemma, has a feasible solution. We show this result by induction.

First, we show the result for  $k = 1$ , i.e. we show that optimization problem  $ALLOC(\hat{C}_c^{(1)}, \hat{C}_t^{(1)}, p)$ , where  $p$  is the bid value given in the statement of the lemma, has a feasible solution. By definition of the function  $SPLIT(C_c, C_t, \mathbf{t}^*, p)$ , we have that there exists an optimal solution  $\mathbf{t}^*$  to optimization problem  $ALLOC(C_c, C_t, p)$  such that

$$\sum_{j \in C_t \cap A_i} t_{i,j}^* D_j(p) \geq I_i, \quad i \in \hat{C}_c^{(1)}.$$

Furthermore, by Lemma 7 we have for every the optimal solution  $\mathbf{t}^*$  to the optimization problem  $ALLOC(C_c, C_t, p)$  that

$$t_{i,j}^* = 0, \quad i \in \hat{C}_c^{(1)}, j \notin \hat{C}_t^{(1)}.$$

Combining the above results, we obtain for the optimal solution  $\mathbf{t}^*$  to the the optimization problem  $ALLOC(C_c, C_t, p)$  that

$$\sum_{j \in A_i \cap \hat{C}_t^{(1)}} t_{i,j}^* D_j(p) \geq I_i, \quad i \in \hat{C}_c^{(1)}.$$

This implies that the optimization problem  $ALLOC(\hat{C}_c^{(1)}, \hat{C}_t^{(1)}, p)$  has a feasible solution.

Next we do an induction step. That is, we assume that for the components  $(\hat{C}_c^{(l)}, \hat{C}_t^{(l)}, q^{(l)}, \mathbf{t}^{(l)})$ ,  $l = 1, \dots, k-1$ , we have that the optimization problem  $ALLOC(\hat{C}_c^{(l)}, \hat{C}_t^{(l)}, p)$  has a feasible solution. We then show that this is also true for the component  $(\hat{C}_c^{(k)}, \hat{C}_t^{(k)}, q^{(k)}, \mathbf{t}^{(k)})$ , i.e. we have that the optimization problem  $ALLOC(\hat{C}_c^{(k)}, \hat{C}_t^{(k)}, p)$  has a feasible solution.

By the definition of the algorithm there exists a recursion step  $REC(\tilde{C}_c, \tilde{C}_t)$  where

$$\tilde{C}_c \subseteq C_{c,2}$$

and

$$\tilde{C}_t \subseteq C_{t,2},$$

that returns the sets  $(\tilde{C}_{c,1}, \tilde{C}_{t,1})$  and  $(\tilde{C}_{c,2}, \tilde{C}_{t,2})$  such that

$$\hat{C}_c^{(k)} \subseteq \tilde{C}_{c,2}$$

and

$$\tilde{C}_{c,1} \subseteq \bigcup_{l=1}^{k-1} \hat{C}_c^{(l)}.$$

By the induction assumption, we have for  $l = 1, \dots, k-1$  that

$$\sum_{j \in \hat{C}_t^{(l)}} D_j(p) \geq \sum_{i \in \hat{C}_c^{(l)}} I_i.$$

This implies that for the bid value used for the optimization problem  $ALLOC(\tilde{C}_c, \tilde{C}_t, q)$  that was used in the recursion step  $REC(\tilde{C}_c, \tilde{C}_t, \emptyset, \emptyset)$  we have that

$$q < p.$$

Furthermore, by the definition of the algorithm we have that there exists an optimal solution  $\mathbf{t}^*$  to optimization problem  $ALLOC(\tilde{C}_c, \tilde{C}_t, q)$  such that

$$\sum_{j \in \tilde{C}_t \cap A_i} t_{i,j}^* D_j(q) \geq I_i, \quad i \in \hat{C}_c^{(k)}.$$

By Lemma 7 we have for every optimal solution  $\mathbf{t}^*$  to the optimization problem  $ALLOC(\tilde{C}_c, \tilde{C}_t, q)$  that

$$t_{i,j}^* = 0, \quad i \in \hat{C}_c^{(k)}, j \notin \hat{C}_t^{(k)}.$$

Combining the above results, we obtain for the optimal solution  $\mathbf{t}^*$  to the the optimization problem  $ALLOC(\hat{\mathcal{C}}_c, \hat{\mathcal{C}}_t, q)$  that

$$\sum_{j \in \mathcal{A}_i \cap \hat{\mathcal{C}}_t^{(k)}} t_{i,j}^* D_j(q) \geq I_i, \quad i \in \hat{\mathcal{C}}_c^{(k)}.$$

This implies that the optimization problem

$$ALLOC(\hat{\mathcal{C}}_c^{(k)}, \hat{\mathcal{C}}_t^{(k)}, p)$$

has a feasible solution. This completes the induction step, and establishes the result of the lemma.  $\square$

**Lemma 9.** *Let  $\mathcal{C}_c \subseteq \mathcal{N}$  be a set of campaigns, and  $\mathcal{C}_t \subseteq \mathcal{T}$  be a set of targeting groups. Furthermore, suppose that the the function  $SPLIT(\mathcal{C}_c, \mathcal{C}_t, \mathbf{t}^*, p)$  in Algorithm 1 returns the sets  $(\mathcal{C}_{c,1}, \mathcal{C}_{t,1})$  and  $(\mathcal{C}_{c,2}, \mathcal{C}_{t,2})$ , let  $p$  be the bid value used in the function  $ALLOC(\mathcal{C}_c, \mathcal{C}_t, p)$ , and let  $\mathbf{t}^*$  be an optimal solution returned by the function  $ALLOC(\mathcal{C}_c, \mathcal{C}_t, p)$ .*

*Then for all the components  $(\hat{\mathcal{C}}_c, \hat{\mathcal{C}}_t, q, \mathbf{t})$  that are obtained from the recursion  $REC(\mathcal{C}_{c,1}, \mathcal{C}_{t,1}, \emptyset, \emptyset)$ , we have that*

$$q > p.$$

*Proof.* Suppose that we obtain  $K$  components from the recursion  $REC(\mathcal{C}_{c,1}, \mathcal{C}_{t,1})$ . We then number the components and create the set  $\{(\hat{\mathcal{C}}_c^{(k)}, \hat{\mathcal{C}}_t^{(k)}, q^{(k)}, \mathbf{t}^{(k)})\}_{k=1}^K$  such that for a given  $k$  we have that

$$\mathcal{A}_i \cap \hat{\mathcal{C}}_t^{(l)} = \emptyset, \quad i \in \hat{\mathcal{C}}_c^{(k)}, l = 1, \dots, k-1.$$

Note that we can create this ordering by Lemma 7.

Recall that we say that the optimization problem

$$ALLOC(\mathcal{C}_c, \mathcal{C}_t, p)$$

has a feasible solution if there exists a solution  $\mathbf{t}^*$  such that

$$\sum_{j \in \mathcal{A}_i \cap \mathcal{C}_t} t_{i,j}^* D_j(p) = I_i, \quad i \in \mathcal{C}_c.$$

In order to show the result of the lemma, we have to show that the optimization problem  $ALLOC(\hat{\mathcal{C}}_c^{(k)}, \hat{\mathcal{C}}_t^{(k)}, q)$ ,  $k = 1, \dots, K$ , does not have a feasible solution for  $q \leq p$ , where  $p$  is the bid value given in the statement of the lemma. We show this result by induction.

First, we show the result for  $k = 1$ , i.e. we show that optimization problem  $ALLOC(\hat{\mathcal{C}}_c^{(1)}, \hat{\mathcal{C}}_t^{(1)}, q)$  does not have a feasible solution for  $q \leq p$ , where  $p$  is the bid value given in the statement of the lemma. We prove this case by contradiction. That is we assume that there exists a  $q$ ,  $0 \leq q < p$ , such that the optimization problem  $ALLOC(\hat{\mathcal{C}}_c^{(1)}, \hat{\mathcal{C}}_t^{(1)}, q)$  has a feasible solution, and show that this leads to a contradiction.

Suppose that there exists an optimal solution  $\mathbf{t}$  to the optimization problem  $ALLOC(\hat{\mathcal{C}}_c^{(1)}, \hat{\mathcal{C}}_t^{(1)}, q)$  such that

$$\sum_{j \in \hat{\mathcal{C}}_t^{(1)} \cap \mathcal{A}_i} t_{i,j} D_j(q) \geq I_i, \quad i \in \hat{\mathcal{C}}_c^{(1)}.$$

By Lemma 7 we have for every optimal solution  $\mathbf{t}^*$  to the optimization problem  $ALLOC(\mathcal{C}_c, \mathcal{C}_t, p)$  that

$$\sum_{i \in \hat{\mathcal{C}}_c^{(1)}} t_{i,j}^* = 1, \quad j \in \hat{\mathcal{C}}_t^{(1)}.$$

Combining the above results, we obtain for every optimal solution  $\mathbf{t}^*$  to the the optimization problem  $ALLOC(\mathcal{C}_c, \mathcal{C}_t, p)$  that

$$\sum_{j \in \mathcal{A}_i \cap \hat{\mathcal{C}}_t^{(1)}} t_{i,j}^* D_j(p) \geq I_i, \quad i \in \hat{\mathcal{C}}_c^{(1)},$$

as by assumption we have that

$$q \leq p.$$

However this leads to a contradiction to the fact that

$$\hat{\mathcal{C}}_c^{(1)} \subseteq \mathcal{C}_{c,1}$$

and

$$\hat{\mathcal{C}}_t^{(1)} \subseteq \mathcal{C}_{t,1}.$$

This establishes that the optimization problem

$$ALLOC(\hat{\mathcal{C}}_c^{(1)}, \hat{\mathcal{C}}_t^{(1)}, q)$$

does not have a feasible solution for  $q \leq p$

Next we do an induction step. That is, we that for the components  $(\hat{\mathcal{C}}_c^{(l)}, \hat{\mathcal{C}}_t^{(l)}, q^{(l)}, \mathbf{t}^{(l)})$ ,  $l = 1, \dots, k-1$ , we have that

$$q^{(l)} > p.$$

We then show that this also has to be true for component  $(\hat{\mathcal{C}}_c^{(k)}, \hat{\mathcal{C}}_t^{(k)}, q^{(k)}, \mathbf{t}^{(k)})$ , i.e. we have that

$$q^{(k)} > p.$$

We prove this case again by contradiction. That is we assume that for  $q$ ,  $0 \leq q < p$ , the optimization problem  $ALLOC(\hat{\mathcal{C}}_c^{(k)}, \hat{\mathcal{C}}_t^{(k)}, q)$  has a feasible solution, and show that this leads to a contradiction.

By the definition of the algorithm we have in this case for the optimal solution  $\mathbf{t}$  to the optimization problem

$$ALLOC(\hat{\mathcal{C}}_c^{(k)}, \hat{\mathcal{C}}_t^{(k)}, q)$$

that

$$\sum_{j \in \hat{\mathcal{C}}_t^{(k)} \cap \mathcal{A}_i} t_{i,j} D_j(q) \geq I_i, \quad i \in \hat{\mathcal{C}}_c^{(k)}.$$

Furthermore, by the definition of the algorithm there exists a recursion step  $REC(\tilde{\mathcal{C}}_c, \tilde{\mathcal{C}}_t)$  where

$$\tilde{\mathcal{C}}_c \subseteq \tilde{\mathcal{C}}_{c,1}$$

and

$$\tilde{\mathcal{C}}_t \subseteq \tilde{\mathcal{C}}_{t,1},$$

that returns the sets  $(\tilde{\mathcal{C}}_{c,1}, \tilde{\mathcal{C}}_{t,1})$  and  $(\tilde{\mathcal{C}}_{c,2}, \tilde{\mathcal{C}}_{t,2})$  such that

$$\hat{\mathcal{C}}_c^{(k)} \subseteq \tilde{\mathcal{C}}_{c,1}$$

and

$$\tilde{\mathcal{C}}_{c,2} \subseteq \bigcup_{l=1}^{k-1} \hat{\mathcal{C}}_c^{(l)}.$$

By Lemma 7 we have for every optimal solution  $\mathbf{t}^*$  to the optimization problem  $ALLOC(\mathcal{C}_c, \mathcal{C}_t, p)$  that

$$\sum_{i \in (\hat{\mathcal{C}}_c^{(k)}, \hat{\mathcal{C}}_t^{(k)}, q^{(k)}, \mathbf{t}^{(k)})} t_{i,j}^* = 1, \quad j \in \hat{\mathcal{C}}_t^{(k)}$$

and

$$t_{i,j}^* = 0, \quad j \in \hat{\mathcal{C}}_t^{(k)}, i \notin \hat{\mathcal{C}}_c^{(k)}.$$

Combining the above results, we obtain for every optimal solution  $\mathbf{t}^*$  to the the optimization problem  $ALLOC(\mathcal{C}_c, \mathcal{C}_t, p)$  that

$$\sum_{j \in \mathcal{A}_i \cap (\hat{\mathcal{C}}_c^{(k)}, \hat{\mathcal{C}}_t^{(k)}, q^{(k)}, \mathbf{t}^{(k)})} t_{i,j}^* D_j(p) \geq I_i, \quad i \in \hat{\mathcal{C}}_c^{(k)}.$$

as by assumption we have that

$$q \leq p.$$

This leads to a contradiction that

$$\hat{\mathcal{C}}_c^{(k)} \subseteq \tilde{\mathcal{C}}_{c,1}$$

and

$$\hat{\mathcal{C}}_t^{(k)} \subseteq \tilde{\mathcal{C}}_{t,1}.$$

This implies that the optimization problem

$$ALLOC((\hat{\mathcal{C}}_c^{(k)}, \hat{\mathcal{C}}_t^{(k)}, q^{(k)}, \mathbf{t}^{(k)}), \hat{\mathcal{C}}_t^{(k)}, q)$$

does not have a feasible solution for  $q \leq p$ . This also completes the induction step, and establishes the result of the lemma.  $\square$

The following result follows directly from the definition of the algorithm given by Eq. (10).

**Lemma 10.** *Let  $(\hat{\mathcal{C}}_c, \hat{\mathcal{C}}_t, p, \mathbf{t})$  be a component returned by the algorithm given by Eq. (10), then we have that there exists a feasible solution to the optimization problem  $ALLOC(\hat{\mathcal{C}}_c, \hat{\mathcal{C}}_t, p)$ , and we have that*

$$\sum_{j \in \hat{\mathcal{C}}_t} D_j(b) < \sum_{i \in \hat{\mathcal{C}}_c} I_i, \quad 0 \leq b < p.$$

## E Proof of Proposition 6

In this section we prove Proposition 6. For the reader's convenience we repeat the statement of the proposition.

**Proposition 19.** *Let  $\{(\hat{\mathcal{C}}_c, \hat{\mathcal{C}}_t, p, \mathbf{t})\}_{p \in \mathcal{P}}$  be the components that the recursions given by Eq. (10) returned, and let the allocation  $(\mathbf{b}^*, \gamma^*)$  with components*

$$\mathcal{C}(p^*, \mathbf{b}^*, \gamma^*), \quad p \in \mathcal{P}(\mathbf{b}^*, \gamma^*),$$

be given as follows. We set

$$\mathcal{P}(\mathbf{b}^*, \gamma^*) = \mathcal{P}.$$

Furthermore, for each bid value  $p^* \in \mathcal{P}(\mathbf{b}^*, \gamma^*)$ , we set

$$\mathcal{C}_c(p^*, \mathbf{b}^*, \gamma^*) = \hat{\mathcal{C}}_c$$

and

$$\mathcal{C}_t(p^*, \mathbf{b}^*, \gamma^*) = \hat{\mathcal{C}}_t.$$

Finally, the allocation  $(\mathbf{b}^*, \gamma^*)$  is such that

$$b_{i,j}^* = p^*, \quad i \in \mathcal{C}_c(p^*, \mathbf{b}^*, \gamma^*), j \in \mathcal{C}_t(p^*, \mathbf{b}^*, \gamma^*)$$

as well as

$$\gamma_{i,j}^* = t_{i,j}, \quad i \in \mathcal{C}_c(p^*, \mathbf{b}^*, \gamma^*), j \in \mathcal{C}_t(p^*, \mathbf{b}^*, \gamma^*)$$

and

$$\gamma_{i,j}^* = 0, \quad i \in \mathcal{C}_c(p^*, \mathbf{b}^*, \gamma^*), j \notin \mathcal{C}_t(p^*, \mathbf{b}^*, \gamma^*).$$

Then  $(\mathbf{b}^*, \gamma^*)$  is an optimal solution.

*Proof.* In order to prove Proposition 6 we need to show that the allocation  $(\mathbf{b}^*, \gamma^*)$  given by the components

$$\mathcal{C}(p^*, \mathbf{b}^*, \gamma^*), \quad p \in \mathcal{P}(\mathbf{b}^*, \gamma^*),$$

defined in the statement of the proposition indeed satisfy the sufficient conditions of Proposition 18. We prove this result by induction as follows.

We first show for the lowest bid value  $p^* \in \mathcal{P}(\mathbf{b}^*, \gamma^*)$ , that the component  $\mathcal{C}(p^*, \mathbf{b}^*, \gamma^*)$  satisfies the sufficient conditions given in Proposition 18.

We then do an induction step where we assume that the the sufficient conditions given in Proposition 18 are satisfied for the  $n$  components with the  $n$  lowest bid values in the set  $\mathcal{P}(\mathbf{b}^*, \gamma^*)$ , and then show that the sufficient conditions are satisfied by the component with the  $(n + 1)$ th lowest bid value in  $\mathcal{P}(\mathbf{b}^*, \gamma^*)$ .

Let  $p^*$  be the lowest bid value in the set  $\mathcal{P}(\mathbf{b}^*, \gamma^*)$ . Then by construction, i.e. by the definition of the Algorithm 1, we have that

$$\sum_{j \in \mathcal{A}_i(\mathbf{b}^*, \gamma^*)} \gamma_{i,j}^* D_j(p^*) = I_i, \quad i \in \mathcal{C}_c(p^*, \mathbf{b}^*, \gamma^*).$$

In addition, we have by the definition of Algorithm 1 that

$$\sum_{j \in \mathcal{C}_t(p^*, \mathbf{b}^*, \gamma^*)} D_j(b) < \sum_{i \in \mathcal{C}_c(p^*, \mathbf{b}^*, \gamma^*)} I_i, \quad 0 \leq b < p^*,$$

and

$$\mathcal{T}_0 \subset \mathcal{C}_t(p^*, \mathbf{b}^*, \gamma^*).$$

This implies that

$$\sum_{j \in \mathcal{T}_0} D_j(b) = 0, \quad 0 \leq b < p^*.$$

It then follows that the sufficient conditions given in Proposition 18 are satisfied for the component  $\mathcal{C}(p^*, \mathbf{b}^*, \gamma^*)$  with the lowest bid value  $p^*$ .

Next we assume that the sufficient conditions given in Proposition 18 are satisfied for the  $n$  components with the  $n$  lowest bid values in the set  $\mathcal{P}(\mathbf{b}^*, \gamma^*)$ , and then show that the sufficient conditions are satisfied by the component with the  $(n+1)$ th lowest bid value in  $\mathcal{P}(\mathbf{b}^*, \gamma^*)$ . More precisely, let  $p^* \in \mathcal{P}(\mathbf{b}^*, \gamma^*)$  be the  $(n+1)$ th lowest bid value, and let  $\mathcal{C}(p^*, \mathbf{b}^*, \gamma^*)$  be the corresponding component.

By Lemma 7, for all bid values  $q^* \in \mathcal{P}(\mathbf{b}^*, \gamma^*)$ , such that

$$q^* < p^*,$$

we have that

$$\mathcal{A}_i \cap \mathcal{C}_t(q^*, \mathbf{b}^*, \gamma^*) = \emptyset, \quad i \in \mathcal{C}_c(p^*, \mathbf{b}^*, \gamma^*).$$

In this case, the sufficient conditions given in Proposition 18 are that we have that

$$\sum_{j \in \mathcal{A}_i(\mathbf{b}^*, \gamma^*)} \gamma_{i,j}^* D_j(p^*) = I_i, \quad i \in \mathcal{C}_c(p^*, \mathbf{b}^*, \gamma^*),$$

and

$$\sum_{j \in \mathcal{C}_t(p^*, \mathbf{b}^*, \gamma^*)} D_j(b) < \sum_{i \in \mathcal{C}_c(p^*, \mathbf{b}^*, \gamma^*)} I_i, \quad 0 \leq b < p^*.$$

Note that this is the case by the definition of Algorithm 1. This completes the induction step, and the proof of the proposition.  $\square$

## F Proof of Proposition 7 and 8

In this appendix we proof Proposition 7 and Proposition 8 of Section 8. Before we do that, we derive in the next subsections a bound on the cost of mixed strategies.

Recall the statement of Proposition 7.

**Proposition 20.** *Let  $s \in \mathcal{S}$  be a mixed strategy that is a solution to the optimization problem given by Eq. 1. Then we have that*

$$C(s) \geq \sum_{p^* \in \mathcal{P}(\mathbf{b}^*, \gamma^*)} \left[ I_{p^*} p^* - \int_0^{p^*} D_{p^*}(b) db \right].$$

*Proof.* We use the result of Proposition 10 to prove Proposition 7. To do that, we use the following notation. For each component  $\mathcal{C}(p^*, \mathbf{b}^*, \gamma^*)$ ,  $p \in \mathcal{P}(\mathbf{b}^*, \gamma^*)$ , of the allocation  $(\mathbf{b}^*, \gamma^*)$  obtained by the algorithm in Section 7, let

$$I_{p^*} = \sum_{i \in \mathcal{C}_c(p^*, \mathbf{b}^*, \gamma^*)} I_i$$

and

$$D_{-p^*}(b) = \sum_{j \in \mathcal{C}_t(p^*, \mathbf{b}^*, \gamma^*)} D_j(b), \quad 0 \leq b.$$

Using Proposition 10 we have that the cost of the mixed allocation  $(\mathbf{b}^*, \gamma^*)$  that obtains the an amount of

$$I(\mathcal{C}(p^*, \mathbf{b}^*, \gamma^*))$$

impressions from component

$$\mathcal{C}(p^*, \mathbf{b}^*, \gamma^*), \quad p \in \mathcal{P}(\mathbf{b}^*, \gamma^*),$$

is lower-bounded by

$$I_{p^*} p^* - \int_0^{p^*} D_{p^*}(b) db. \quad (11)$$

Therefore, in order to proof the result, we have to show for each component

$$\mathcal{C}(p^*, \mathbf{b}^*, \gamma^*), \quad p^* \in \mathcal{P}(\mathbf{b}^*, \gamma^*),$$

that is not possible reduce the bound given by Eq. (11) by shifting some of the impressions  $I_{p^*}$  to another component. We show this result by induction as follows.

We first show that for the component corresponding to the highest bid value  $p \in \mathcal{P}(\mathbf{b}^*, \gamma^*)$ , that we can not reduce the bound given by Eq. (11) by shifting some of the impressions  $I_{p^*}$  to another component.

We then do an induction step where we assume that it is not possible to reduce the the bound given by Eq. (11) for obtaining the impressions for the components corresponding to the  $n$  highest bid values  $p^* \in \mathcal{P}(\mathbf{b}^*, \gamma^*)$ , and then show that it is not possible to reduce the the bound given by Eq. (11) for obtaining the impressions for the components corresponding to the  $(n + 1)$  highest bid values  $p^* \in \mathcal{P}(\mathbf{b}^*, \gamma^*)$ .

Let  $p^*$  be the the highest bid value in the set  $\mathcal{P}(\mathbf{b}^*, \gamma^*)$ , and suppose that we move an amount of  $\Delta_I$  of the impressions  $I_{p^*}$  to another component. Note that as the solution satisfies the sufficient conditions given in Proposition 18, the only option is to move the  $\Delta_I$  impressions to the set  $\mathcal{T}_0$ , i.e. the set

$$\mathcal{T}_0 = \left\{ j \in \mathcal{T} \mid \sum_{i \in \mathcal{B}_j} \gamma_{i,j}^* D_j(b_{i,j}^*) = 0 \right\}.$$

Let

$$\mathcal{T}_0(p^*) = \mathcal{T}_0 \cap \cup_i \in \mathcal{C}_c(p^*, \mathbf{b}^*, \gamma^*) A_i,$$

i.e. the set  $\mathcal{T}_0(p^*)$  contains all the targeting groups  $j \in \mathcal{T}_0$  from which at least one campaign  $i \in \mathcal{C}_c(p^*, \mathbf{b}^*, \gamma^*)$  can obtain impressions. Furthermore, let  $D_0(b)$  be given by

$$D_0(b) = \sum_{j \in \mathcal{T}_0(p^*)} D_j(b).$$

Note by definition that we have that

$$D_0(p^*) = 0.$$

Finally, suppose that there exists a bid value  $q$  such that

$$D_0(q) \geq \Delta_I$$

and

$$D_0(b) < \Delta_I, \quad 0 \leq b < q.$$

Using these definitions and the fact that

$$D_0(p^*) = 0$$

and

$$D_0^{(-)}(q) \leq \Delta_I,$$

we have that that the cost for getting  $\Delta_I$  impressions from the targeting groups in the set  $\mathcal{T}_0(p^*)$  is at least

$$\begin{aligned} & \Delta_I q - \int_{p^*}^q D_0(b) db \\ & \geq \Delta_I q - D_0^{(-)}(q)(q - p^*) \\ & \geq \Delta_I q - \Delta_I(q - p^*) \\ & = \Delta_I p^*. \end{aligned}$$

Similarly using the bid value  $p'$  such that

$$D_{p^*}(p') \geq I_{p^*} - \Delta_I$$

and

$$D_{p^*}(p') < I_{p^*} - \Delta_I, \quad b < p',$$

the reduction in the lower bound for reducing the number of impressions obtained from component  $\mathcal{C}(p^*, \mathbf{b}^*, \gamma^*)$  by  $\Delta_I$  is at most

$$\begin{aligned} & I_{p^*} p^* - \int_0^{p^*} D_p(b) db \\ & - (I_{p^*} - \Delta_I) p' + \int_0^{p'} D_p(b) db \\ & = I(\mathcal{C}(p, \mathbf{b}, \gamma))(p^* - p') + \Delta_I p' \\ & - \int_{p'}^{p^*} D_p(b) db \\ & \leq I(\mathcal{C}(p, \mathbf{b}, \gamma))(p^* - p') + \Delta_I p' \\ & - D_p(p')(p^* - p') \\ & = (I_{p^*} - D_p(p'))(p^* - p') + \Delta_I p' \\ & \leq \Delta_I(p^* - p') + \Delta_I p' \\ & = \Delta_I p^*. \end{aligned}$$

Therefore, we obtain that it is not possible to reduce the the bound given by Eq. (11) by shifting some of the impressions  $I_{p^*}$  to targeting groups in the set  $\mathcal{T}_0(p^*)$ . This establishes the result that the cost for getting  $I_{p^*}$  impressions for the campaigns in component  $\mathcal{C}(p^*, \mathbf{b}^*, \gamma^*)$  is bounded from below by

$$I_{p^*} p^* - \int_0^{p^*} D_{p^*}(b) db.$$

This establishes the result for the component  $\mathcal{C}(p^*, \mathbf{b}^*, \gamma^*)$  with the highest bid value  $p^*$ .

Having established the result for the component with the highest bid value, we do an induction step. That is, we assume that it is not possible to reduce the the bound given by Eq. (11) for obtaining the impressions for the components corresponding to the  $n$  highest bid values in  $\mathcal{P}(\mathbf{b}^*, \gamma^*)$ , and then show that in this case it is not possible to reduce the the bound given by Eq. (11) for obtaining the impressions for the components corresponding to the  $(n + 1)$  highest bid values in  $\mathcal{P}(\mathbf{b}^*, \gamma^*)$ .

Let  $p_{n+1}^*$  be the  $(n + 1)$ th highest bid value and let  $\mathcal{C}(p^*, \mathbf{b}^*, \gamma^*)$  be the corresponding component. Note that in this case, the campaigns  $i \in \mathcal{C}_c(p^*, \mathbf{b}^*, \gamma^*)$  that belong to this component can only get impressions from targeting groups that belong to the set  $\mathcal{C}_t(p^*, \mathbf{b}^*, \gamma^*)$ , from targeting groups that belong to a component with a higher bid value  $q^* \in \mathcal{P}(\mathbf{b}^*, \gamma^*)$ , or from the set of targeting groups  $\mathcal{T}_0(p^*)$  given by

$$\mathcal{T}_0(p^*) = \mathcal{T}_0 \cap \cup_{i \in \mathcal{C}_c(p^*, \mathbf{b}^*, \gamma^*)} A_i,$$

i.e. the set  $\mathcal{T}_0(p^*)$  contains all the targeting groups  $j \in \mathcal{T}_0$  from which at least one campaign  $i \in \mathcal{C}_c(p^*, \mathbf{b}^*, \gamma^*)$  can obtain impressions.

As we have that for  $D_0(b)$  given by

$$D_0(b) = \sum_{j \in \mathcal{T}_0(p^*)} D_j(b),$$

that

$$D_0(p^*) = 0,$$

it follows from the same argument as given for the case of the component with the highest bid value, that it is not possible to reduce the the bound given by Eq. (11) for obtaining the impressions for the components corresponding to the  $(n + 1)$  highest bid values in  $\mathcal{P}(\mathbf{b}^*, \gamma^*)$  by shifting some impressions from the component  $\mathcal{C}(p^*, \mathbf{b}^*, \gamma^*)$  to the set  $\mathcal{T}_0(p^*)$ .

There it remains to consider the case where we shift some impressions from the component  $\mathcal{C}(p^*, \mathbf{b}^*, \gamma^*)$  to a component  $\mathcal{C}(q^*, \mathbf{b}^*, \gamma^*)$  with a higher bid value  $q^*$ .

For this case, let  $p'$  be such that

$$D_{p^*}(p') \geq I_{p^*} - \Delta_I$$

and

$$D_{p^*}(p') < I_{p^*} - \Delta_I, \quad 0 \leq b < p',$$

and let  $q'$  be such that

$$D_{q^*}(q') \geq I_{q^*} + \Delta_I$$

and

$$D_{q^*}(q') < I_{q^*} + \Delta_I, \quad 0 \leq b < q'.$$

Under these definitions, by shifting an amount of  $\Delta_I$  impressions from component  $\mathcal{C}(p^*, \mathbf{b}^*, \gamma^*)$  to  $\mathcal{C}(q^*, \mathbf{b}^*, \gamma^*)$ , we can reduce the bound given by Eq. (11) by at most an amount given by

$$\begin{aligned} & I_{p^*} p - \int_0^{p^*} D_p(b) db \\ & - (I_{p^*} - \Delta_I) p' + \int_0^{p'} D_p(b) db \\ = & I_{p^*} (p^* - p') + \Delta_I p' \\ & - \int_{p'}^{p^*} D_p(b) db \\ \leq & I_{p^*} (p^* - p') + \Delta_I p' \\ & - D_p(p')(p^* - p') \\ = & (I_{p^*} - D_p(p'))(p^* - p') + \Delta_I p' \\ \leq & \Delta_I (p^* - p') + \Delta_I p' \\ = & \Delta_I p^*. \end{aligned}$$

Similarly, by shifting an amount of  $\Delta_I$  impressions from component  $\mathcal{C}(p^*, \mathbf{b}^*, \gamma^*)$  to  $\mathcal{C}(q^*, \mathbf{b}^*, \gamma^*)$ , we increase the lower bound by at least an amount given by

$$\begin{aligned} & -I_{q^*} q^* + \int_0^{q^*} D_q(b) db \\ & (I_{q^*} + \Delta_I) q' - \int_0^{q'} D_q(b) db \\ = & I_{q^*} (q' - q^*) + \Delta_I q' \\ & - \int_{q^*}^{q'} D_q(b) db \\ \geq & I_{q^*} (q' - q^*) + \Delta_I q' \\ & - D_q^{(-)}(q')(q' - q^*) \\ = & (I_{q^*} + \Delta_I - D_q^{(-)}(q'))(q' - q^*) + \Delta_I q^*, \\ \geq & \Delta_I q^*. \end{aligned}$$

As we have that

$$p^* < q^*,$$

it follows that we can not lower the bound for getting the impressions for component  $\mathcal{C}(p^*, \mathbf{b}^*, \gamma^*)$  by shifting some impressions to component  $\mathcal{C}(q^*, \mathbf{b}^*, \gamma^*)$ .

This establishes the result that the bound given by Eq. (11) provides a lower bound for obtaining the impressions for the campaigns in the components with the  $(n+1)$  highest bid values  $p \in \mathcal{P}(\mathbf{b}^*, \gamma^*)$ . This completes the induction step, and the proof of the proposition.  $\square$

Next we prove Proposition 8. Recall the statement of the proposition.

**Proposition 21.** *Let  $(\mathbf{b}^*, \gamma^*)$  with components  $\mathcal{C}(p^*, \mathbf{b}^*, \gamma^*)$ ,  $p^* \in \mathcal{P}(\mathbf{b}^*, \gamma^*)$ , be the allocation obtained by the algorithm in Section 7. Then for every component  $\mathcal{C}(p^*, \mathbf{b}^*, \gamma^*)$ ,  $p^* \in \mathcal{P}(\mathbf{b}^*, \gamma^*)$ , the following is true. Let  $j \in \mathcal{C}_t(p^*, \mathbf{b}^*, \gamma^*)$  be a targeting group that belongs to the component  $\mathcal{C}(p^*, \mathbf{b}^*, \gamma^*)$ , and for all campaigns  $i \in \mathcal{B}_j(\mathbf{b}^*, \gamma^*)$  let*

$$\gamma_{i,j} = \frac{\gamma_{i,j}^*}{\sum_{m \in \mathcal{B}_j(\mathbf{b}^*, \gamma^*)} \gamma_{m,j}^*}$$

and

$$I_{i,j} = \gamma_{i,j}^* D_j(p^*).$$

Furthermore, let  $s_{p^*}(b_1)$  be a mixed bidding strategy for the component  $\mathcal{C}(p^*, \mathbf{b}^*, \gamma^*)$  such that for all campaigns  $i \in \mathcal{C}_c(p^*, \mathbf{b}^*, \gamma^*)$  and  $j \in \mathcal{C}_t(p^*, \mathbf{b}^*, \gamma^*)$  we have that

$$\mathcal{P}_{i,j} = \{b_1, p^*\}, \quad j \in \mathcal{A}_i(\mathbf{b}^*, \gamma^*),$$

as well as

$$\gamma_{i,j}(b_1) = \frac{\gamma_{i,j} D_1(p^*) - I_{i,j}}{D_1(p^*) - D_1(b_1)}, \quad j \in \mathcal{A}_i(\mathbf{b}^*, \gamma^*),$$

and

$$\gamma_{i,j}(p^*) = \gamma_{i,j} - \gamma_{1,1}(b_1), \quad j \in \mathcal{A}_i(\mathbf{b}^*, \gamma^*).$$

Finally, let  $C_{p^*}(s_{p^*}(b_1))$  be the cost of strategy  $s_{p^*}(b_1)$  on component  $\mathcal{C}(p^*, \mathbf{b}^*, \gamma^*)$ .

If  $D_{p^*}^{(-)}(p^*) > 0$ , then we have that

$$\lim_{b_1 \uparrow b^*} C_{p^*}(s_{p^*}(b_1)) = I_{p^*} p^* - \int_0^{p^*} D_{p^*}(b) db.$$

*Proof.* To prove the proposition, we proceed as follows. Note that by definition we have for every targeting group  $j \in \mathcal{C}_t(p^*, \mathbf{b}^*, \gamma^*)$  that

$$\sum_{i \in \mathcal{B}_j(\mathbf{b}^*, \gamma^*)} [\gamma_{i,j}(b_1) + \gamma_{i,j}(b^*)] = \sum_{i \in \mathcal{B}_j(\mathbf{b}^*, \gamma^*)} \gamma_{i,j} = 1.$$

Furthermore, for every targeting group  $j \in \mathcal{C}_t(p^*, \mathbf{b}^*, \gamma^*)$  we have for every campaign  $i \in \mathcal{B}_j(\mathbf{b}^*, \gamma^*)$ ,

$$\begin{aligned} & \gamma_{i,j}(b_1) D_j(b_1) + \gamma_{i,j}(b^*) D_j(p^*) \\ &= \gamma_{i,j}(b_1) D_j(b_1) + [\gamma_{i,j} - \gamma_{1,1}(b_1)] D_j(p^*) \\ &= \gamma_{i,j} D_j(p^*) - \gamma_{i,j}(b_1) (D_j(p^*) - D_j(b_1)) \\ &= \gamma_{i,j} D_j(p^*) - \gamma_{i,j} D_j(p^*) + I_{i,j} \\ &= \gamma_{i,j}^* D_j(p^*) = I_{i,j}. \end{aligned}$$

Using this result, it follows that for each campaign  $i \in \mathcal{C}_c(p^*, \mathbf{b}^*, \gamma^*)$  we have that

$$\begin{aligned} & \sum_{j \in \mathcal{A}_i(\mathbf{b}^*, \gamma^*)} [\gamma_{i,j}(b_1) D_j(b_1) + \gamma_{i,j}(b^*) D_j(p^*)] \\ &= \sum_{j \in \mathcal{A}_i(\mathbf{b}^*, \gamma^*)} \gamma_{i,j}^* D_j(p^*) \\ &= I_i. \end{aligned}$$

This means that the mixed strategy  $s_{p^*}(b_1)$  obtains for each campaign  $i \in \mathcal{C}_c(p^*, \mathbf{b}^*, \gamma^*)$  the required number of impressions.

Finally, under the bidding strategy  $s(b_1)$  we have for each campaign  $i \in \mathcal{C}_c(p^*, \mathbf{b}^*, \gamma^*)$  the cost for obtaining the

$$I_{i,j} = \gamma_{i,j}^* D_j(p^*)$$

impressions from targeting group  $j \in \mathcal{A}_i(\mathbf{b}^*, \gamma^*)$  is given by

$$\begin{aligned} & \gamma_{i,j}(b_1) \left[ D_j(b_1)b_1 - \int_0^{b_1} D_j(b)db \right] \\ & + (\gamma_{i,j} - \gamma_{i,j}(b_1)) \left[ D_j(p^*)p^* - \int_0^{p^*} D_j(b)db \right] \\ = & \gamma_{i,j}(b_1)D_1(b_1)b_1 + (\gamma_{i,j} - \gamma_{i,j}(b_1))D_1(p^*)p^* \\ & - \gamma_{i,j} \int_0^{p^*} D_j(b)db \\ & + \gamma_{i,j}(b_1) \int_{b_1}^{p^*} D_j(b)db \\ = & \gamma_{i,j}(b_1)D_1(b_1)p^* + (\gamma_{i,j} - \gamma_{i,j}(b_1))D_1(p^*)p^* \\ & + \gamma_{i,j}(b_1)D_1(b_1)(b_1 - p^*) \\ & - \gamma_{i,j} \int_0^{p^*} D_j(b)db \\ & + \gamma_{i,j}(b_1) \int_{b_1}^{p^*} D_j(b)db \\ = & I_{i,j} - \gamma_{i,j} \int_0^{p^*} D_j(b)db \\ & + \gamma_{i,j}(b_1)D_1(b_1)(b_1 - p^*) \\ & + \gamma_{i,j}(b_1) \int_{b_1}^{p^*} D_j(b)db. \end{aligned}$$

Combining the above results, it then follows that

$$\begin{aligned} & \lim_{b_1 \uparrow p^*} C_{p^*}(s_{p^*}(b_1)) \\ = & \sum_{i \in \mathcal{C}_c(p^*, \mathbf{b}^*, \gamma^*)} \sum_{j \in \mathcal{A}_i(\mathbf{b}^*, \gamma^*)} \left[ I_{i,j}p^* - \gamma_{i,j} \int_0^{p^*} D_j(b)db \right] \\ & + \lim_{b_1 \uparrow p^*} \sum_{i \in \mathcal{C}_c(p^*, \mathbf{b}^*, \gamma^*)} \sum_{j \in \mathcal{A}_i(\mathbf{b}^*, \gamma^*)} \gamma_{i,j}(b_1)D_1(b_1)(b_1 - p^*) \\ & + \lim_{b_1 \uparrow p^*} \sum_{i \in \mathcal{C}_c(p^*, \mathbf{b}^*, \gamma^*)} \gamma_{i,j}(b_1) \int_{b_1}^{p^*} D_j(b)db \\ = & \sum_{i \in \mathcal{C}_c(p^*, \mathbf{b}^*, \gamma^*)} \sum_{j \in \mathcal{A}_i(\mathbf{b}^*, \gamma^*)} \left[ I_{i,j}p^* - \gamma_{i,j} \int_0^{p^*} D_j(b)db \right] \end{aligned}$$

It then follows that

$$\begin{aligned}
& \lim_{b_1 \uparrow p^*} C_{p^*}(s_{p^*}(b_1)) \\
&= \sum_{i \in \mathcal{C}_c(p^*, \mathbf{b}^*, \gamma^*)} \sum_{j \in \mathcal{A}_i(\mathbf{b}^*, \gamma^*)} I_{i,j} p^* \\
&\quad - \sum_{i \in \mathcal{C}_c(p^*, \mathbf{b}^*, \gamma^*)} \sum_{j \in \mathcal{A}_i(\mathbf{b}^*, \gamma^*)} \gamma_{i,j} \int_0^{p^*} D_j(b) db
\end{aligned}$$

Using the fact that

$$\begin{aligned}
& \sum_{i \in \mathcal{C}_c(p^*, \mathbf{b}^*, \gamma^*)} \sum_{j \in \mathcal{A}_i(\mathbf{b}^*, \gamma^*)} \gamma_{i,j} \int_0^{p^*} D_j(b) db \\
&= \sum_{j \in \mathcal{C}_t(p^*, \mathbf{b}^*, \gamma^*)} \sum_{i \in \mathcal{B}_j(\mathbf{b}^*, \gamma^*)} \gamma_{i,j} \int_0^{p^*} D_j(b) db,
\end{aligned}$$

we obtain that

$$\begin{aligned}
&= I_{p^*} p^* - \sum_{j \in \mathcal{C}_t(p^*, \mathbf{b}^*, \gamma^*)} \sum_{i \in \mathcal{B}_j(\mathbf{b}^*, \gamma^*)} \gamma_{i,j} \int_0^{p^*} D_j(b) db \\
&= I_{p^*} p^* - \int_0^{p^*} \left( \sum_{j \in \mathcal{C}_t(p^*, \mathbf{b}^*, \gamma^*)} D_j(b) \right) db \\
&= I_{p^*} p^* - \int_0^{p^*} D_{p^*}(b) db.
\end{aligned}$$

The result of the proposition then follows. □