Operations Research

Publication details, including instructions for authors and subscription information:
http://pubsonline.informs.org

Tractable Equilibria in Sponsored Search with Endogenous Budgets

Dragos Florin Ciocan, Krishnamurthy Iyer

To cite this article:
Dragos Florin Ciocan, Krishnamurthy Iyer (2020) Tractable Equilibria in Sponsored Search with Endogenous Budgets. Operations Research

Published online in Articles in Advance 24 Dec 2020

. https://doi.org/10.1287/opre.2020.2052

informs
http://pubsonline.informs.org/journal/opre

OPERATIONS RESEARCH
Articles in Advance, pp. 1–18
ISSN 0030-364X (print), ISSN 1526-5463 (online)

Crosscutting Areas

Tractable Equilibria in Sponsored Search with Endogenous Budgets

Dragos Florin Ciocan,[a] Krishnamurthy Iyer[b]

[a] INSEAD, Technology and Operations Management, Fontainebleau 77300, France; [b] Industrial and Systems Engineering, University of Minnesota, Minneapolis, Minnesota 55455

Contact: florin.ciocan@insead.edu, https://orcid.org/0000-0003-2246-6066 (DFC); kriyer@umn.edu, https://orcid.org/0000-0002-5538-1432 (KI)

Received: February 11, 2020
Accepted: April 23, 2020
Published Online in Articles in Advance: December 24, 2020

Subject Classifications: games/group decisions; bidding/auctions
Area of Review: Market Analytics and Revenue Management

https://doi.org/10.1287/opre.2020.2052

Abstract. We consider an ad network's problem of allocating the auction for each individual impression to an optimal subset of advertisers with the goal of revenue maximization. This is a variant of bipartite matching except that advertisers may strategize by choosing their bidding profiles and their total budget. Because the ad network's allocation rule affects the bidders' strategies, equilibrium analysis is challenging. We show that this analysis is tractable when advertisers face a linear budget cost r_j. In particular, we show that the strategy in which advertisers bid their valuations shaded by a factor of $1 + r_j$ is an approximate equilibrium with the error decreasing with market size. This equilibrium can be interpreted as one in which a bidder facing an opportunity cost r_j is guaranteed a return on investment of at least r_j per dollar spent. Furthermore, in this equilibrium, the optimal allocation for the ad network, as determined from a linear program (LP), is greedy with high probability. This is in contrast with the exogenous budgets case, in which the LP optimization is challenging at practical scales. These results are evidence that, although in general such bipartite matching problems may be challenging to solve because of their high dimensionality, the optimal solution is remarkably simple at equilibrium.

Funding: K. Iyer gratefully acknowledges support from the National Science Foundation [Grant CMMI-2002156].
Supplemental Material: The online appendix is available at https://doi.org/10.1287/opre.2020.2052.

Keywords: bipartite matching • endogenous budgets • truthful equilibrium

1. Introduction

In the sponsored search advertising market, ad-serving systems (henceforth referred to as *ad networks*), such as Google's AdWords, monetize millions of search engine keyword queries (*impressions*) every day. Furthermore, these markets involve a large number of advertisers (*bidders*), that compete for the opportunity to display their ads. Given the immense scale, it is practically important yet technically quite challenging to understand the interplay between the behavior of the bidders and that of the ad network and its effect on the long-run state of the market.

Taking a deeper look into the market operation, for each keyword query, the ad network must make two decisions: first, the ad network must choose the set of advertisers that will compete for the impression and, second, within that set of advertisers, determine one (or more) advertiser(s) to which the impression is allocated. As a solution to the latter decision problem, the generalized second price (GSP) (Edelman et al. 2007, Varian 2007) auction has emerged as the gold standard allocation mechanism.

The former problem, namely restricting the set of bidders that participate in the auction for an impression, is known as *throttling* (Goel et al. 2010) and has received a lot of recent attention. A number of factors play an important role in the ad networks' throttling policy, such as the bidders' (remaining) budgets, targeting requirements, and campaign durations (see, e.g., Balseiro et al. 2017, Conitzer et al. 2017). Given the scale of the market, simple throttling policies, such as the *greedy* mechanism that always picks the maximal set of bidders eligible for each impression, have a practical appeal. However, Abrams et al. (2007) and Goel et al. (2010) show that, through optimal throttling, the ad network can ensure substantial revenue gains as compared with simple allocation mechanisms.

Given the substantial revenue gains from throttling, the considerable literature on constrained resource allocation problems has been brought to bear on this problem. In particular, given the bids and the budgets of the bidders and assuming that the ad network's objective is to maximize its own aggregate revenues, the problem of finding the optimal throttling decision can be cast as a linear program (LP); this LP finds the optimal allocation of impressions to sets of advertisers that will participate in the GSP auction. Although appealing, the dimensions of this problem in practice render solving the corresponding LP to optimality a computationally daunting task. This suggests that the ad network faces a trade-off between adopting a computationally simple mechanism, such as the greedy mechanism, and incurring a revenue loss or implementing the optimal throttling mechanism but facing a computationally challenging resource-allocation problem.

However, the preceding discussion crucially ignores the fact that, in the long run, the bidders not only strategize over the individual bids they submit to the auctions for the impressions, but also strategically set their campaign budgets. Furthermore, with the advent of automated bidding agents, a reasonable case can be made that budgets are the primary lever many advertisers use to optimize their campaigns. For example, in its Ads Help Center, Facebook (2019) features advice on how to set budgets as part of advertising fundamentals, whereas advice on setting bid strategies features less prominently.[1]

There are many reasons for this strategic choice of budgets. For one, the bidders may have cost of capital or opportunity costs for the budgets that preclude setting budgets substantially higher than their expected expenditures. Another related effect is that bidders may have a return-on-investment (ROI) constraint and set budgets and bids so as to remain above this ROI constraint (Borgs et al. 2007, Auerbach et al. 2008, Wilkens et al. 2017). But, more importantly, because the ad network's throttling policy takes the bidders' budgets into account, these budgets provide another lever the bidders may use to affect their (or their competitions') ad allocations.

If the bidder behavior is endogenized, the analysis of suboptimality of the greedy mechanism present in Abrams et al. (2007) and Goel et al. (2010) is invalidated, and the question remains: how does the interplay between the ad network's throttling policy and the bidders' bidding and budget-setting behavior affect the market operation? We seek to address this question in this paper and to obtain valuable insights about the market and its operation when bidders' budgets arise endogenously as a response to an ad network's allocation policy.

1.1. Model Description and Results

Formally, we consider a sponsored search market with n impressions and m bidders, indexed by i and, respectively, j. For simplicity, we assume there is a single ad slot available for each impression. The market operates in the following way:

- The bidding profile of each advertiser j is simultaneously announced to the ad network in the form of an n-dimensional bid vector $\mathbf{b}_j$ for the n different impressions and a budget B_j. We think of B_j as a hard constraint on j's total spend for impressions.
- Upon receiving the bidding profiles $\{\mathbf{b}_j, B_j\}_{j\in[m]}$, the ad network decides, for each impression i, a subset of bidders that will participate in the second price auction for i. The outcome of the auction is to award the impression to the highest bidder but at a cost equal to the second highest bid.

We assume that each bidder j faces a linear cost $r_j \cdot B_j$ for the budgets they declare to the ad network: r_j can be interpreted as an interest rate or opportunity cost of the budget that is committed to an ad campaign.

We focus on the setting in which the ad network seeks to maximize its (expected) revenues. Finally, as we make precise in Section 2, we assume that the ad network commits to solving a linear optimization problem to produce the optimal subset of advertisers that are allocated to each impression i. Informally, this linear program maximizes the ad network's own aggregate revenues subject to not exceeding advertiser budget constraints or the supply of impressions.

Having defined the primitives of the problem, we give a concrete example that there are substantial revenue gains from optimal throttling when budgets are set exogenously. Consider the following example adapted from Abrams et al. (2007):

Example 1. Consider an instance with $n = 2$ impressions and $m = 3$ bidders. The configuration of budgets and bids is given in Table 1. Let us assume that impressions 1 and 2 arrive as a fluid over some time interval normalized to $[0, 1]$. We examine two policies:

- The ad network uses a greedy policy to allocate every infinitesimal amount of impressions 1 and 2. This means that the entire set $\{1, 2, 3\}$ of bidders participates in the auction, and bidder 1 wins both $i = 1$ and 2, up to the time $\tau = 1/2$ when bidder 1 runs out of budget and effectively drops out of the system. Although bidder 2 still has positive budget, the greedy policy can only allocate to bidder 2 impression 1 for the remaining time horizon; bidder 3 remains the only bidder eligible to win the remaining amount of impression $i = 2$ and wins it at 0 (we assume there is no reserve). Total ad network revenues are $\frac{3}{2} - \frac{\delta}{2}$.
- The ad network chooses the set of bidders $\{2, 3\}$ for $i = 1$ and $\{1, 3\}$ for $i = 2$. Then, bidder 2 wins $i = 1$ at a spend of $1 - \delta$, and 1 wins $i = 2$ at a spend of $1 - \delta$. The ad network's revenues are now $2 - 2\delta$. We note the fact that this allocation of sets of advertisers to the

two impressions can indeed be obtained by the ad network via a linear program that optimizes over X^i_{jk} the fraction of impression i that is throttled to a pair (j,k) of advertisers:

$$\max_{\mathbf{X}\geq 0}\ X^1_{12}+(1-\delta)X^1_{13}+(1-\delta)X^1_{23}+(1-\delta)X^2_{13}$$

$$\text{subject to } X^1_{12}+(1-\delta)X^1_{13}+(1-\delta)X^2_{13}\leq 1-\frac{\delta}{2},$$

$$X^1_{23}\leq 1 \quad \text{(bidder budget constraints)}$$

$$X^1_{12}+X^1_{13}+X^1_{23}\leq 1,$$

$$X^2_{13}\leq 1. \qquad \text{(item supply constraints).}$$

Thus, as we let $\delta\to 0$ in this example, the greedy policy only garners 3/4 of the optimal revenues. Although the example instance we consider is small, the suboptimality of the greedy mechanism does not hinge on the size of the market in the sense that it is possible to modify this example to arbitrarily large dimensions of m and n and maintain the same 3/4 gap. Finally, although here we consider a fluid model for the arrival of impressions, a similar argument can be made in the setting in which impressions arrive discreetly when the optimality gap can be shown to be 1/2 (Abrams et al. 2007).

Crucially, in the preceding example, we have assumed that bidder behavior (in terms of bid and budget profiles) does not respond to the ad network changing its throttling policy. Our contribution in this paper is to endogenize this behavior; surprisingly, this reveals appealing structural properties of the market in equilibrium, which would not hold true in the exogenous case. We discuss these results in more detail next.

1.1.1. Existence of a Tractable Equilibrium. Showing the existence of an equilibrium in our model turns out to be considerably difficult. To highlight the challenge, consider that bidder j's best response function depends on the ad network's choice of allocation (or throttling policy), which, in turn, is the optimal assignment of a linear program. It is typically difficult to perform sensitivity analysis on the (high-dimensional) solution to a linear program, which may fail to have the continuity and convexity properties needed by the machinery of fixed-point theorems to which game-theoretic analyses typically appeal.

Table 1. Example of Suboptimality of Greedy Throttling with $n=2$ and $m=3$

Bidder	Bid for $i=1$	Bid for $i=2$	Budget
1	$1+\delta$	1	$1-\frac{\delta}{2}$
2	1	0	1
3	$1-\delta$	$1-\delta$	2

To circumvent this difficulty, we define a notion of ϵ-approximate equilibrium, which guarantees that the expected ex ante gain of an arbitrary bidder j's deviation from the strategy σ_j is at most a $1+\epsilon$ factor away from j's profit under (σ_j,σ_{-j}). We show that a bidding strategy in which advertisers bid their true valuation shaded down by a factor of $1+r_j$ and declare a budget equal to their expected spend multiplied by a factor of $1+o(\epsilon)$, is such an ϵ-approximate equilibrium when the distributions of bidder valuations are heterogeneous exponentials. We give a precise characterization of ϵ in terms of the market size parameters m and n.

An interesting interpretation of the bid shading we observe in the approximate equilibrium is that it effectively implements an ROI guarantee for the advertisers, which shade their bids exactly to the point at which they are guaranteed a return on their spend that exceeds r_j. We believe this interplay is new in the literature on sponsored search markets. Moreover, there is evidence that such an ROI-driven paradigm for the operation of online advertising markets, rather than one that is built around hard budgets, can better model real advertiser behavior (Auerbach et al. 2008).

1.1.2. Optimality of the Greedy Throttling Policy. In addition, we show that, when bidders play this approximate equilibrium, the ad network's optimal throttling policy is, with exponentially high probability, greedy. In other words, assuming that the market plays out into this equilibrium, the ad network does not need to solve the large-scale linear optimization problems that are typically required for generic budgeted bipartite matching settings. We find this result quite surprising, particularly because the practical necessity to solve such massive matching problems has motivated some of the recent research on large-scale optimization. We additionally remark that, thus, the resulting strategies and equilibrium landscape have a simple structure that is easy to interpret and rationalize by the bidders participating in the market.

1.1.3. Extension to Reserve Prices and General Valuation Distributions. We also explore two extensions to our model. First, we show that, if the ad network is allowed to run a reserve price, then our approximate equilibrium remains valid with the slight modification that budget declarations increase to account for this reserve. Second, we consider symmetric bidders with valuations coming from general distributions with hazard rates that are simultaneously bounded from above and below and show that our strategy still constitutes an equilibrium.

1.1.4. Technical Contributions. The main technical contribution of our paper is in showing that, under endogenous budgets, both the equilibrium behavior of the bidders and the ad network's optimization problem drastically simplify. In particular, we provide an exact characterization of the large market regime under which truthful reporting of values, shaded down by a $1 + r_j$ factor, and a greedy allocation by the ad network constitute an approximate equilibrium and obtain bounds on the rate of convergence. Our analysis proceeds in two steps:

1. In broad terms, to show that this response is approximately optimal for a single bidder given the strategies of other bidders and the allocation policy of the ad network, we identify a corresponding event for each candidate deviating strategy. Then, we prove that, conditioned on that event, the strategy under consideration obtains a lower payoff than the equilibrium strategy. As a step toward showing this result, we use LP duality to characterize the structure of the ad network's allocation as a response to this deviation.

2. The second step of our analysis involves showing that the event identified in the first step occurs with very high probability as the number of items and agents increases. This step is fairly technical and involves obtaining probabilistic bounds using Bernstein-type concentration results for subexponential random variables.

Although our approximation bounds on the equilibrium require a large market with $n \gg m$, we emphasize that our analysis and results differ from the other notions of equilibria in large markets, such as mean field equilibrium (MFE) (Gummadi et al. 2011, Iyer et al. 2014, Balseiro et al. 2015). In particular, in an MFE, one typically assumes that the market is populated by an infinite number of bidders, and each bidder takes the allocation and bid distribution in the market to be fixed and independent of the bidder's actions in the market. In contrast, our model explicitly takes into account the effect each bidder's strategy has on the ad network's allocation. In particular, in our model, a single bidder could, in principle, increase the spend of other bidders and exhaust their budgets, thereby reducing the competition.

On the other hand, MFE models explicitly include market dynamics, allowing for bidder strategies to depend on their current states, such as remaining budgets, whereas our model assumes that bidder behavior is static, and each submits a single fixed bid profile and budget at the start of the campaign. Nevertheless, previous literature, such as Balseiro et al. (2015), shows that static analysis provides a good approximation to dynamic bidding behavior in large markets. Specifically, via a "fluid approximation" of the market, any variability that could induce bidders to change their bids dynamically is smoothed out. Moreover, in a dynamic setting, a bidder must model other bidders' dynamic responses in order to influence their spending and exhaust their budgets; thus, a bidder's deviations may need to be substantially more complex than those in the static setting in order to be profitable, making them less practically plausible. Finally, we note that our static model allows for more deviations for a bidder, and the bidder can base bids on the entire valuation vector. In contrast, in a dynamic model, the deviations are more restricted as they cannot depend on future valuations. Thus, we believe our guarantees are fairly conservative in a dynamic setting.

1.2. Literature Review

Several variants of the sponsored search matching problem described have been considered in the literature with most of the work falling into two distinct streams. The first stream treats the bidding profiles $(\mathbf{b}_j, B_j)$ of the advertisers in the system as exogenous and formulates the core problem as one of optimal resource allocation or budgeted bipartite matching to compute an optimal allocation of impressions to advertisers. Our throttling policies are akin to bipartite matching policies with the added complication that prices depend not only on the bidder that wins the impression, but rather on the set of bidders that participate in the auction; thus, a throttling policy can be thought of as bipartite matching between impressions and sets of bidders.

The other stream fixes the ad network's throttling policy as exogenous and typically simple (such as greedy) and examines the game-theoretic behavior of advertisers bidding in the resulting system. Our current work is an attempt to bridge these two streams of literature on sponsored search ad matching: the model is an instance of the optimal budgeted bipartite matching problem but one in which we endogenize the bids and budgets as coming from the bidders' self-interested behavior. We organize our literature review around these two distinct research streams.

1.2.1. Optimal Bipartite Matching. A large body of research deals with the following prototypical matching problem: we are given a bipartite graph with fixed rewards p_{ij} on its edges and a budget on each right-side vertex. Any time we match a right-side vertex i to a left-side vertex j, we receive the reward tied to the edge but consume some amount c_{ij} from j's budget. In addition, $\mathbf{p}$ and $\mathbf{c}$ could be given by a second price mechanism as in our sponsored search example or by a first price mechanism that is more common in display advertising settings but are exogenous parameters. The objective is to find a matching of maximal aggregate reward subject to constraints.

When the supply of left-side vertices is deterministically known, the problem is solvable by a linear program. Thus, most of the attention is devoted to models in which this supply is uncertain, either stochastic or adversarial. In adversarial settings and for particular choices of $\mathbf{p}$ and $\mathbf{c}$, such as the AdWords structure described, there are algorithms that achieve $1 - 1/e$ approximations of the off-line optimal, such as Karp et al. (1990), Mehta et al. (2007), and Buchbinder et al. (2007); this is the best achievable competitive ratio against an adversary. On the other hand, if the left-side vertices arrive in a random permutation or if they are independent and identically distributed (i.i.d.) samples from a fixed distribution, the competitive ratios can be refined to yield essentially optimal performance as in Devanur and Hayes (2009), Feldman et al. (2009, 2010), and Agrawal and Devanur (2015). We defer to the survey of Mehta et al. (2013) for a more comprehensive coverage of this literature.

The bipartite matching problem is also reminiscent of the network revenue management problem from operations research. There, in the case that vertices arrive as a known-rate Poisson process, one can employ a fluid analysis to show that the dynamic matching problem can be well approximated by a static linear program that uses mean arrival rates as estimates of actual arrivals (Gallego and Van Ryzin 1997, Talluri and Van Ryzin 1998). Ciocan and Farias (2012) use a similar LP approach when the rates of the Poisson processes are themselves stochastic.

We mention that, in practical situations, ad networks match tens of millions of impressions per day with tens of thousands of advertisers; the large scale of the problems makes even basic optimization methods, such as linear programming, intractable computationally. This has been a considerable hurdle to implementing throttling in practice.

1.2.2. Strategic Bidders and Mechanism Design. There is a vast body of literature that studies the strategic behavior of bidders in auctions and analyzes the mechanism design issues therein (Edelman et al. 2007; Varian 2007, 2009; Babaioff et al. 2009; Devanur and Kakade 2009). A number of these papers study strategic behavior and mechanism design in dynamic settings (Bergemann and Välimäki 2010, Gummadi et al. 2011, Athey and Segal 2013, Nazerzadeh et al. 2013, Iyer et al. 2014, Balseiro et al. 2015, Mirrokni and Nazerzadeh 2017, Dütting et al. 2016, Balseiro and Gur 2019, Kanoria and Nazerzadeh 2020).

Closer to our work, Balseiro et al. (2015) analyze bidding behavior of agents under budget constraints in a large market using the methodology of a mean field equilibrium. The authors show that, in equilibrium, each advertiser shades the bid dynamically, and the shading factor depends on the remaining budget.

Finally, a number of recent papers (Zhou et al. 2008, Charles et al. 2013, Karande et al. 2013, Asadpour et al. 2014, Balseiro et al. 2017, Conitzer et al. 2017) study budget management strategies adopted by advertisers to effectively manage their budget depletion. Balseiro et al. (2017) study a problem close to our model, in which they characterize, for a number of different budget-management strategies of the ad network, the "system equilibrium." A system equilibrium ignores strategic considerations of the bidders, and assuming they report their values truthfully, constitutes the long-run market outcome. In contrast, in our model, the bidders' strategic considerations are central; however, our focus is only on the optimal allocation of the ad network. Furthermore, they consider a setting in which budget constraints only need to hold in expectation, and the bidders incur a cost of the expected spend and not on the budget allocation.

2. Model

We consider a setting in which there are m bidders and n items: in a sponsored search market, bidders correspond to advertisers and items to impressions. Each bidder is endowed with a vector of valuations $\mathbf{V}_j = (V_{ij} : i \in [n])$ for the items. Each individual valuation V_{ij} is drawn i.i.d. (across bidders and items) from an exponential distribution with parameter λ_j. We make this distributional choice for $\mathbf{V}$ mainly for technical convenience; in Section 4, we discuss extensions to valuation distributions with symmetric and bounded hazard rates and examine the precise distributional properties needed for our analysis. From a practical perspective, such a bounded hazard-rate condition is justified under sufficient heterogeneity in the impressions and if the ad network can use tools, such as targeting, to identify and differentiate among the impressions.

2.1. Market Operation

The sequence of events in the sponsored search auction market is as follows:

1. Each bidder j's valuation $\mathbf{V}_j$ is realized and remains private information to j.
2. Based on the bidder's own realization of $\mathbf{V}_j$, each bidder j submits a budget $B_j(\mathbf{V}_j)$ and bids $\mathbf{b}_j(\mathbf{V}_j) = (b_{ij}(\mathbf{V}_j) : i \in [n])$ to the ad network.
3. The ad network observes the submitted bids $\mathbf{b} = (\mathbf{b}_j : j \in [m])$ and the budgets $\mathbf{B} = (B_j : j \in [m])$ and decides on the allocation $\mathbf{X}(\mathbf{B}, \mathbf{b})$ of items to auctions among subsets of bidders. Thus, we let $X(\mathbf{B}, \mathbf{b}) = (X_S^i(\mathbf{B}, \mathbf{b}) : i \in [n], S \subseteq [m], |S| \geq 2)$, where $X_S^i(\mathbf{B}, \mathbf{b})$ denotes the proportion of item i auctioned off among bidders in subset S. Because of the second price auction mechanism, we require that at least two bidders form S.

Our formulation allows for fractional allocations of items; because we are dealing with a regime in which $n \gg m$ and, as we see, an optimal allocation uses only few bidder subsets S, this is a reasonable simplification.

Furthermore, methods to convert such fractional allocations to integral ones are common in the literature, such as interpreting the fractions X_S^i as probabilities of allocation (Jasin and Kumar 2012).

Additionally, we denote by $\delta > 0$ the outside value that the ad network can garner for any item not allocated to an auction. The main purpose of this outside value is technical: in degenerate instances in which many allocations $\mathbf{X}$ yield the same revenue to the ad network, the parameter δ acts as a means to select allocations that are most effective. We let δ be arbitrarily small (but positive), minimizing its impact on the ad network's revenue.

2.2. Budget Costs

We assume that each bidder faces a linear cost $r_j \cdot B_j$ for the budget it commits to the ad network. We interpret r_j as capturing the opportunity cost of allocating a budget to the advertising campaign. The linear assumption is a reasonable model for budget costs in practice, in which the advertisers' budget decisions do not exhibit extreme scale variation. More broadly, the linear assumption serves as a good first-order approximation to general budget cost structures and is a first step toward tackling this challenging setting in its full generality. Moreover, a merit of the linear assumption is that, as we show, it leads to a simple equilibrium that one can expect agents in a sponsored search system could implement in practice.

We also highlight that, here, bidders incur cost for the unspent portion of their budgets. An alternative model is one in which bidders borrow money in real time as they win items. However, we believe a more accurate description of real-life operations is that advertisers commit a budget to a sponsored search market at the expense of using it for other outside opportunities; thus, it is more realistic to account for the cost of capital incurred on the entire commitment of budget B_j.

2.3. Strategy, Allocation, and Payoffs

A strategy for bidder j, denoted by σ_j, specifies the pair $(B_j(\mathbf{V}_j), \mathbf{b}_j(\mathbf{V}_j))$ for all values of $\mathbf{V}_j$; we denote the space of all strategies by Σ. Note that, in general, an agent's bids may be correlated across items through its dependence on the entire valuation profile $\mathbf{V}_j$.

For each $(\mathbf{B}, \mathbf{b})$, the ad network chooses the allocation $\mathbf{X}(\mathbf{B}, \mathbf{b})$ to maximize its revenue, subject to bidders' budget constraints being satisfied. Before formally describing this allocation policy, we introduce a notation: given a bid vector $\mathbf{b}$, for any item $i \in [n]$ and bidder j, we let $p^{\mathbf{b}}(i,j)$ denote the highest bid less than the bid of bidder j:

$$p^{\mathbf{b}}(i,j) = \begin{cases} \max_{k \neq j, b_{ik} \le b_{ij}} \{b_{ik}\}, & \text{if } b_{ij} \neq \min_k \{b_{ik}\}; \\ 0, & \text{otherwise.} \end{cases} \quad (1)$$

Recall that the ad network's basic decision variable is X_S^i, the proportion of item i that is allocated to an auction among bidders in subset $S \subseteq [m]$. With the goal toward revenue maximization, it turns out that it suffices to consider only the auctions between a bidder j and the bidder k with the highest bid below j's, for a payment of $p^{\mathbf{b}}(i,j)$. Intuitively, this is because

- Allocating item i to the bidder set $\{j,k\}$ attains the same outcome as allocating it to the larger set $\{j,k\} \cup S_0$, where S_0 includes other bidders with bids below j's and k's. Thus, the ad network can achieve optimal revenues if it restricts the allocation to only pairs of bidders.
- The ad network prefers allocating i to the bidder set $\{j,k\}$ instead of $\{j,\ell\}$ for some bidder ℓ with $b_{i\ell} < b_{ik} = p^{\mathbf{b}}(i,j)$ as that always uses item i more efficiently. Thus, we can further restrict the allocations to only these bidder pairs.

We show this equivalence formally in Online Appendix C.1.

This observation allows us consider alternative decision variables Y_{ij} for each item i and bidder j that capture the proportion of item i that is auctioned between bidder j and the bidder with bid $p^{\mathbf{b}}(i,j)$. In other words, Y_{ij} captures the proportion of item i that is won by bidder j. With this change of variables, the ad network commits to choosing an allocation policy $\mathbf{Y}(\mathbf{B}, \mathbf{b})$ that is a solution to the following LP:

$$\begin{aligned} \mathbf{Y}(\mathbf{B}, \mathbf{b}) \in \arg\max_{\mathbf{Y} \ge 0} & \sum_{i=1}^n \sum_{j=1}^m p^{\mathbf{b}}(i,j) Y_{ij} + \delta \sum_{i=1}^n \left(1 - \sum_{j=1}^m Y_{ij}\right) \\ \text{subject to } & \sum_{i=1}^n p^{\mathbf{b}}(i,j) Y_{ij} \le B_j, \quad \text{for each } j \in [m], \\ & \sum_{j=1}^m Y_{ij} \le 1, \quad \text{for each } i \in [n]. \end{aligned} \quad (2)$$

Here, the first constraint ensures that the payments of bidder j do not exceed the bidder's declared budget, and the second constraint ensures that the total allocation of item i does not exceed its available quantity of one. Finally, the objective captures the revenue across all allocations together with the outside value of unallocated items.

Having described the ad network's allocation policy, we are now ready to describe the payoff of bidder j. Given a strategy profile $\sigma = (\sigma_j, \sigma_{-j})$, the realized spend of bidder j is given by

$$\Psi_j(\sigma_j, \sigma_{-j}) \triangleq \sum_{i=1}^n p^{\mathbf{b}}(i,j) Y_{ij}(\mathbf{B}, \mathbf{b}),$$

where $\mathbf{Y}(\mathbf{B}, \mathbf{b})$ optimizes (2). Similarly, we define the payoff of bidder j as

$$\pi_j(\sigma_j, \sigma_{-j}) \triangleq \sum_{i=1}^n \left(V_{ij} - p^{\mathbf{b}}(i,j)\right) Y_{ij}(\mathbf{B}, \mathbf{b}) - r_j B_j.$$

Note that this is a net payoff because it subtracts the cost of budget $r_j B_j$ from the utility gained by j from winning items. In our analysis, we also refer to bidder j's payoff minus the cost incurred on the unspent (slack) portion of its budget, which we denote by

$$\tilde{\pi}_j(\sigma_j, \sigma_{-j}) \triangleq \pi_j(\sigma_j, \sigma_{-j}) + r_j \cdot (B_j - \Psi_j(\sigma_j, \sigma_{-j})).$$

We call the quantity $\tilde{\pi}_j$ the *modified profit* of bidder j. An interpretation of the modified profit $\tilde{\pi}_j$ is that it is equal to the bidder's payoff had it set its budget exactly equal to realized spend.

2.4. Other Notation

Given a bid vector $\mathbf{b}$, we employ the following notation to identify the set of items for which bidder k has the qth highest bid:

$$U^{\mathbf{b}}_{k,q} \triangleq \{i \in [n] \text{s.t. bidder } k \text{ has } q\text{-th highest bid for item } i \text{ across all bidders } l \in [m]\}.$$

Recall that we focus on the setting in which the valuations are distributed continuously. Consequently, we ignore issues of tie breaking among bids because ties have measure zero for the bidder strategies we consider. Therefore, in the rest of the paper, we assume that, for each i and q, there exists a unique k such that $i \in U^{\mathbf{b}}_{k,q}$. Finally, for a vector of bids across one item $\mathbf{b}_i$, we define $b^{(q)}_i$ to be the qth highest valuation among $\mathbf{b}_i$ and $b^{(q)}_{i,-j}$ to be the qth highest valuation among all $\mathbf{b}_i$ excluding b_{ij}.

With this notation, we obtain the following definition of a greedy allocation policy:

Definition 1. Given a bid vector $\mathbf{b}$, an allocation $\mathbf{Y}$ is *greedy* if $Y_{ij} = \mathbf{I}\{i \in U^{\mathbf{b}}_{j,1}, b^{(2)}_i \geq \delta\}$.

Note that, in the greedy allocation, each item i is allocated to the bidder j with the highest bid (i.e., $i \in U^{\mathbf{b}}_{j,1}$) as long as the bidder's payment $p^{\mathbf{b}}(i,j) = b^{(2)}_i$ is greater than the value of outside option δ. Recall that we allow the outside option δ to be arbitrarily small (but positive). Thus, in the regime in which $\delta \to 0$, each item is won by the highest bidder under the greedy allocation.

2.5. Equilibrium Concept

Summarizing our preceding description of the model, observe that the bidders compete by submitting bidding profiles and budgets given that the ad network sets the subsequent allocation according to the solution to the linear program (2). One natural candidate to describe the resulting outcome of the competition is a Bayes–Nash equilibrium, in which each bidder sets a bidding profile and budgets optimally given the competitions' choice. For the unbudgeted setting, in which the budget cost r_j for each bidder j equals zero, the following result characterizes the Bayes–Nash equilibrium. Let σ^0_j denote the strategy in which bidder j truthfully bids $\mathbf{b}_j = \mathbf{V}_j$ and sets a budget equal to ∞. We have the following lemma:

Lemma 1 (Unbudgeted Case). *Assume $r_j = 0$ for all j. Then, the strategy profile $\sigma^0 = (\sigma^0_j : j \in [m])$ constitutes a Bayes–Nash equilibrium. Given the bidders follow the strategy profile σ^0, it is optimal for the auctioneer to set $\mathbf{Y}(\infty, \mathbf{V})$ to be the greedy allocation.*

Proof. In this setting, the auctioneer essentially implements a second price auction between the top two bidders for each item, independently across the items. This implies that each bidder's dominant strategy is to bid truthfully. To prove the greediness of the optimal allocation, note that, because budgets are infinite, the auctioneer (i.e., ad network) obtains the highest revenue if it runs an auction between the top two bidders for each item, independently across the items. Infinite budgets ensure the allocation is feasible. □

Observe that the equilibrium in the costless budgets case has two highly desirable properties. First, the equilibrium is *truthful*, meaning the bid profile submitted by each bidder is equal to the bidder's value profile. This result, although not entirely surprising, follows from the fact that, because the bidders set high budgets, there is effectively no budget constraint to link the individual item auctions together. The intuition is that large enough budgets decouple the problem into a set of unbudgeted, independent, single-item second price auctions. Second, in the equilibrium, the ad network's optimal allocation has a *simple* solution structure; in particular, the ad network's decision problem (2) can be solved greedily.

There are two challenges to generalizing the preceding result to the case in which the budget costs are nonzero. First, given the complexity of the actions and the noncompactness of the action space, it is unclear whether an equilibrium exists in general. Second, and more importantly, even if such an equilibrium exists, it is unlikely to be simple. This is because, under positive budget costs, a bidder's budget is unlikely to be significantly larger than its expected spend, implying that the bidder's bids across different auctions are highly linked. Thus, when bidder j's budget constraint is binding, the corresponding lack of slack provides strong incentives for the bidder's competitors to underbid on items they expect to win and overbid on other items; the overbidding increases j's spend per item won and restricts it from competing for some other items that can then be won by other bidders for lower payments. This suggests that truthful bidding with *tight* budgets, that is, budgets that are set equal to expected spending, cannot be an equilibrium. We formally show this via an example in Proposition EC.1 in Online Appendix C.2.

Furthermore, this complexity is not primarily driven by the stochastic nature of the problem, such as the underlying mismatch between the realized spending and the expected spending budgeted by the bidders. In fact, even in a large market regime in which the number of both items and bidders are scaled proportionally, which essentially "fluidizes" both the bidders' problems and washes out stochastic effects, one can show that truthful bidding with tight budgets is not an equilibrium. Thus, a "large markets" assumption on its own is not sufficient to yield a natural equilibrium.

In the face of this difficulty, one possible approach forward could be to study mixed equilibria of the underlying game between the bidders and the ad network. However, this presents many challenges of its own: first, we have the technical challenge of working with distributions over a continuum of pure strategies and showing existence; second, and more importantly, it is not immediately clear whether such an approach would yield itself to obtaining a structural understanding of the bidders' equilibrium strategies and budget reports.

Instead, we focus on approximate notions of equilibrium, in which we relax the requirement that each bidder responds optimally to the ad network's and other bidders' strategies. In contrast, we only require each bidder's strategy to be an approximate best response. We define such an equilibrium concept. In the following, we let $\mathbf{E}[\cdot]$ denote the expectation over the valuations $\mathbf{V}$ as well as any randomizations in the bidders' strategy profile.

Definition 2. For a given profile σ_{-j} of strategies of all bidders different from j, we call a strategy σ_j of bidder j an ϵ-approximate best response, denoted by $\sigma_j \in BR_j^{\epsilon}(\sigma_{-j})$, if

$$\sup_{\sigma_j'}\left\{\mathbf{E}\left[\pi_j\left(\sigma_j', \sigma_{-j}\right)\right]\right\} \le \mathbf{E}\left[\pi_j(\sigma_j, \sigma_{-j})\right] \cdot (1+\epsilon).$$

Given the best response definition, our definition of an approximate equilibrium is the following:

Definition 3. A strategy profile $\sigma = (\sigma_j : j \in [m])$ is an ϵ-equilibrium if $\sigma_j \in BR_j^{\epsilon}(\sigma_{-j})$ for each $j \in [m]$.

Our notion of approximate best response differs from standard notions in two distinct ways: (i) We consider approximate best response strategies that are ϵ-close to the best response strategy in a multiplicative sense. In contrast, standard notions of approximate equilibria focus on additive error terms, and (ii) the approximation holds ex ante with respect to the expectation taken over all possible realizations of $\mathbf{V}_j$. Under our definition, there may be values of $\mathbf{V}_j$ for which bidder j can gain substantially from a unilateral deviation from an ϵ-best response. However, our analysis implicitly shows that the probability j receives such a $\mathbf{V}_j$ is exponentially decreasing as m, n grow appropriately large; this yields an alternative equilibrium definition, namely ex post with high probability, for which our results still hold.

In the following section, we show that, indeed, the concept of an approximate equilibrium allows us to obtain the two desirable properties of the equilibrium in the unbudgeted case as obtained in Lemma 1. In particular, we show that there exists an approximate equilibrium, in which the bid profile and the resulting allocation are simple. To state our results, we use the notation $\theta_j \triangleq \lambda_j(1+r_j)$. Furthermore, we assume that each θ_j belongs to some fixed interval $[\theta_{\min}, \theta_{\max}]$ independent of n and m. We let $\vartheta \triangleq \theta_{\max}/\theta_{\min}$ signify the degree of heterogeneity among the bidders. Moreover, we let $\mathbf{E}_{-j}[\cdot] = \mathbf{E}[\cdot|\mathbf{V}_j]$ denote the expectation over $\mathbf{V}_{-j}$, the vector of valuations excluding j's, for a fixed valuation $\mathbf{V}_j$ for bidder j. For example, we write $\mathbf{E}_{-j}[\pi_j(\sigma)]$ to denote the expected payoff of bidder j under strategy profile σ, conditioned on the realization $\mathbf{V}_j$ of its own valuations. Finally, we assume that the outside value δ satisfies $0 < \delta \le \log(4/3)/\theta_{\max}$.

3. Equilibrium Characterization

We begin by describing the strategy profile that we show constitutes an approximate equilibrium as defined in Definition 3. Informally, the bid profile in the strategy is obtained by shading the valuation profile, whereas the budgets are set so that each bidder adds a small multiplicative slack over its expected spend in its budget declaration. The challenge here is balancing two tensions: finding an equilibrium in which the induced slack budgets are relatively small and, thus, are not too costly for the advertisers yet large enough to essentially decouple the problem as in Lemma 1.

3.1. A Tractable Equilibrium

For a given $\epsilon > 0$, we define the strategy profile $\sigma^* = (\sigma_j^* : j \in [m])$, where, for each $j \in [m]$, the strategy $\sigma_j^* \triangleq (B_j^*, \mathbf{b}_j^*)$ is given by

$$\mathbf{b}_j^* \triangleq \mathbf{W}_j = \frac{\mathbf{V}_j}{1+r_j},$$

$$B_j^* \triangleq \mathbf{E}_{-j}[\bar{\Psi}_j] \cdot (1+\beta(m)) = \mathbf{E}_{-j}[\bar{\Psi}_j] \cdot \left(1 + \frac{2\vartheta}{m^{1/\vartheta}}\right), \tag{3}$$

where $\vartheta = \theta_{\max}/\theta_{\min}$, and for each j, $\bar{\Psi}_j$ is defined as

$$\bar{\Psi}_j \triangleq \sum_{i=1}^{n} W_i^{(2)} \mathbf{I}\left\{i \in U_{j,1}^{\mathbf{W}}\right\}. \tag{4}$$

Note that $\bar{\Psi}_j$ corresponds to the realized spend that j incurs if budgets are set equal to infinity and the value of the outside option δ is set to zero. We make a few remarks about this strategy:

i. Under σ^*, the bid profile has an appealing minimum ROI interpretation. To see this, consider the simplest case in which all $r_j = r > 0$; if this is indeed an equilibrium in which the ad network's optimal allocation is greedy as in the unbudgeted case, then any given bidder j wins items i in the set $U^{\mathbf{b}}_{j,1} \cap \{i \text{ s.t. } b^{(2)}_i > \delta\} \approx U^{\mathbf{b}}_{j,1} = U^{\mathbf{V}}_{j,1}$, at a profit of

$$\begin{aligned} &V_{ij} - \text{payment} - \text{opportunity cost} \\ &= V_{ij} - p^{\mathbf{b}}(i,j) - rp^{\mathbf{b}}(i,j) \\ &= V^{(1)}_i - (1+r)\frac{V^{(2)}_i}{1+r} = V^{(1)}_i - V^{(2)}_i, \end{aligned}$$

which is exactly the payoff in the costless case. Moreover, this equilibrium could be thought of as one in which, for every item i won,

$$V_{ij}/p^{\mathbf{b}}(i,j) \geq 1 + r_j.$$

In other words, bidders bid in such a way that the return on every dollar spent (ROI) is guaranteed to exceed their budget cost r.

ii. The precise multiplicative slack that j sets on top of $\mathbf{E}_{-j}[\bar{\Psi}_j]$, which is the expected spend in the unbudgeted case with $\delta = 0$, is $\beta(m) = 2\vartheta/m^{1/\vartheta}$. This multiplicative slack is consistent with typical advertiser behavior, in which advertisers typically adjust budgets over time so as to match their spends while, at the same time, erring on the side of having some small amount of leftover budget.

iii. Under σ^*, each bidder j bids by discounting all its valuations by a common factor of $1 + r_j$. Thus, under σ^*, the knowledge of r_j is enough to back out bidder j's valuations from its bids.

iv. The shading factor here is of a different nature from those encountered in the previous literature, such as in (Balseiro et al. 2015); here, the shading depends on a physical opportunity cost of budget rather than a Lagrangian penalty on the bidder's budget constraint, which, in fact, is equal to zero with high probability in our model.

Our main result, which we prove in the next section, is that σ^* does indeed constitute an ϵ-approximate equilibrium:

Theorem 1. *Let $m \geq \max\{(2\vartheta)^\vartheta, \exp(3\vartheta + 5\vartheta\log(\vartheta) + 1)\}$ and $n \geq 6\vartheta m^{2(\vartheta+1/\vartheta)}\log(m)/C$, where C is a fixed constant independent of m. Then, for any bidder j, we have $\sigma^*_j \in BR^{\epsilon}_j(\sigma^*_{-j})$ for $\epsilon = 11\vartheta^2\log(m)/m^{1/\vartheta}$. In other words, we have, for any deviation σ_j of any bidder j,*

$$\mathbf{E}\Big[\pi_j\Big(\sigma_j, \sigma^*_{-j}\Big)\Big] \leq \mathbf{E}[\pi_j(\sigma^*)] \cdot \left(1 + \frac{11\vartheta^2\log(m)}{m^{1/\vartheta}}\right).$$

Consequently, under the same condition on n and m and for the same value of ϵ, the strategy σ^* is an ϵ-approximate equilibrium.

Theorem 1 can be interpreted as a large-markets guarantee for the validity of the equilibrium: namely, for any choice of $\epsilon > 0$, there exist market parameters n and m that are large enough for σ^* to become an approximate equilibrium with approximation error parameter ϵ. Informally, the theorem highlights the benefits of thick markets in terms of enforcing simple equilibria, such as the one considered here. Specifically, we show that, for any given $\epsilon > 0$, a market with $m > \Omega(1/\epsilon^\vartheta \cdot \log(1/\epsilon)^\vartheta)$ bidders with a corresponding number of items suffices to achieve the stated approximate equilibrium.

Moreover, the following theorem establishes that, given that all bidders are playing σ^*, the ad network's LP optimization problem has a simple greedy solution with high probability. Its proof, given at the end of this section, shows that the bidder budget constraints are slack in our proposed equilibrium with high probability.

Theorem 2. *Let $m \geq \max\{(2\vartheta)^\vartheta, \exp(3\vartheta + 5\vartheta\log(\vartheta) + 1)\}$ and $n \geq 6\vartheta m^{2(\vartheta+1/\vartheta)+1}\log(m)/C$, where C is a fixed constant independent of m. Then, under the strategy profile σ^* with probability at least $1 - \exp(-m)$, the unique optimal allocation the ad network chooses is $\mathbf{Y}(\mathbf{B}^*, \mathbf{b}^*) = \mathbf{Y}(\infty, \mathbf{b}^*)$. Thus, the optimal allocation induced by the strategy profile σ^* is greedy.*

Here, we remark that, although $\mathbf{Y}$ is not necessarily greedy on *all* sample paths, it is greedy with large probability as $m \to \infty$.[2] This is a desirable property because, practically, for most realizations $\mathbf{V}$, the auctioneer can simply check if the greedy solution is, in fact, optimal. This is a significant computational advantage of this equilibrium versus one in which the ad network would have to solve a high-dimensional linear optimization problem given *any* realization of the bidder valuations.

Finally, we note that our analysis does not rule out the existence of other equilibria. However, other equilibria that are not small perturbations of the equilibrium we consider are unlikely to share its simplicity, in which bidders uniformly shade their valuations by $1 + r_j$, and furthermore, their declared budgets depend on simple aggregate markets statistics (such as bid distribution). Such equilibria are likely to either involve correlation in the bidders' bids across different items or strategies that depend on more complicated market statistics. This complexity invariably renders these equilibria implausible, especially in the asymptotic regime we consider. The simplicity of σ^* also suggests its *focality*, making it easier for the bidders to coordinate on the equilibrium and plausible for it to be played by them.

3.2. Characterization of the Best Response Strategy to σ^*

The goal of this section is to prove that the strategy σ^* derived from the truthful bidding strategy is bidder j's best response in the approximate sense of Definition 2, that is, $\sup_{\sigma_j}\{\mathbf{E}[\pi_j(\sigma_j,\sigma^*_{-j})]\} \le \mathbf{E}[\pi_j(\sigma^*)]\cdot(1+\epsilon)$. We first outline the main building blocks of our proof:

i. The key step in deriving our result is to show that deviating from σ^* cannot improve any bidder's *modified profits* $\tilde{\pi}_j$. In other words, we would like to prove that

$$\tilde{\pi}_j\left(\sigma_j,\sigma^*_{-j}\right) \le \tilde{\pi}_j\left(\sigma^*_j,\sigma^*_{-j}\right), \text{ for all } \sigma_j, j.$$

Unfortunately, we cannot prove this uniformly for all realizations of $\mathbf{V}$, so instead, our proof of this inequality holds only for "typical" realizations. In particular, we show in Lemma 3 that, for any j and σ_j, when conditioned on a high-probability event $E_j(\sigma_j)$, it must be that

$$\mathbf{P}\left(\tilde{\pi}_j\left(\sigma_j,\sigma^*_{-j}\right) \le \tilde{\pi}_j\left(\sigma^*_j,\sigma^*_{-j}\right)\middle|E_j(\sigma_j)\right) = 1. \tag{5}$$

Essentially, the event $E_j(\sigma_j)$ that we construct enforces the condition that the additional $\beta(m)=\frac{2\vartheta}{m^{1/\vartheta}}$ fraction of expected spend that bidders declare under σ^* provides sufficient slack to prevent j from playing some alternative strategy $\sigma_j \neq \sigma^*_j$ and succeeding in increasing its payoff.

ii. We further prove in Online Appendix A that $E_j(\sigma_j)$ is typical in that it occurs with high probability for any σ_j as n and m grow large. Then, Lemma 5 uses this high-probability bound to show that

$$\sup_{\sigma_j}\mathbf{E}\left[\tilde{\pi}_j\left(\sigma_j,\sigma^*_{-j}\right)\right] \le \mathbf{E}\left[\tilde{\pi}_j(\sigma^*)\right]\cdot(1+\epsilon). \tag{6}$$

iii. Because the modified profit $\tilde{\pi}_j$ does not account for the expected cost of unspent budget, which is equal to $r_j\cdot\mathbf{E}[B^*_j-\Psi_j(\sigma^*)]$, inequality (6) is not sufficient to prove that σ^* is a best response. To conclude, we additionally show that the cost of this slack is dominated in scale by j's modified profit, that is, $r_j\cdot\mathbf{E}[B^*_j-\Psi_j(\sigma^*)] = o(\mathbf{E}[\tilde{\pi}_j(\sigma^*)])$, and thus, it does not lead to a high enough diminution of profits to break the approximate best response condition.

In the following, we provide more in-depth insights into the mechanics of our technique together with formal proofs of some of the key results. We organize this discussion around the three road map steps outlined here.

3.2.1. Step (i). We begin by introducing the following "slack" term $\Xi_k(\sigma_j)$ for any deviation σ_j of bidder j and for any bidder $k\neq j$:

$$\begin{aligned}\Xi_k(\sigma_j) \triangleq &\sum_i\left(b_{ij}-W_i^{(2)}\right)\mathbf{I}\left\{i\in U^{\mathbf{W}}_{k,1}\cap U^{\mathbf{b}}_{j,2}\right\}\\ &-\sum_i W^{(2)}_{i,-j}\mathbf{I}\left\{i\in U^{\mathbf{W}}_{k,1}\cap U^{\mathbf{b}}_{j,1}\right\}\\ &+\sum_i W^{(2)}_i\mathbf{I}\left\{i\in U^{\mathbf{W}}_{j,1}\cap U^{\mathbf{W}}_{k,2}\right\}.\end{aligned} \tag{7}$$

Intuitively, $\Xi_k(\sigma_j)$ is interpretable as excess budget that bidder k would have had to declare on top of $\bar{\Psi}_k$ in order to prevent the deviation σ_j from being profitable to j. In the following, we show that $\Xi_k(\sigma_j)$ provides a sufficient amount of slack above $\bar{\Psi}_k$ to achieve this.

We construct $\Xi_k(\sigma_j)$ via the following thought experiment: let us consider any arbitrary deviation σ_j, and furthermore, let us assume bidder k wins all the items for which it is the highest bidder (which would be the case under the greedy allocation for small enough δ). Under this scenario, the following lemma (proved in Online Appendix B) shows that bidder k's maximum possible spend is still below the sum of its spend under the truthful equilibrium, $\bar{\Psi}_k$, plus the slack term $\Xi_k(\sigma_j)$:

Lemma 2. *For any deviation σ_j of bidder j and for any $k\neq j$, we have*

$$\sum_{i\in[n]} p^{\mathbf{b}}(i,k)\mathbf{I}\left\{i\in U^{\mathbf{b}}_{k,1}\right\} \le \bar{\Psi}_k + \Xi_k(\sigma_j),$$

where $\mathbf{b}$ is the bid vector induced by the strategy (σ_j,σ^*_{-j}).

One can heuristically break $\Xi_k(\sigma_j)$ up into two distinct components that qualitatively highlight the types of deviations against which this slack protects bidder $k\neq j$:

- The term $\sum_i W^{(2)}_i\mathbf{I}\{i\in U^{\mathbf{W}}_{j,1}\cap U^{\mathbf{W}}_{k,2}\}$ is budget slack preventing underbidding behavior, in which j understates its true valuations. This can be profitable when bidder k runs out of budget by winning other items and, therefore, cannot compete for those items bidder j underbid, which j can then win with a high profit. Qualitatively, this term ensures that, for an item i that j would win under σ^*, another bidder possessing the second highest valuation (that can become the highest bidder if j under-reports) has enough slack budget to win i away from j.
- On the other hand, the $(b_{ij}-W^{(2)}_i)\mathbf{I}\{i\in U^{\mathbf{W}}_{k,1}\cap U^{\mathbf{b}}_{j,2}\}$ term corresponds to the externality that j can impose on some other bidder k by successfully overbidding and increasing the spend on an item k was going to win. Such strategies deplete competitor k's budget

beyond its projections and, in turn, orphan some items that k would have won in equilibrium; j can then win these orphaned items with low bids. In the specification of our slack, this term is balanced by $\sum_i W_{i,-j}^{(2)}\mathbf{I}\{i \in U_{k,1}^{\mathbf{W}} \cap U_{j,1}^{\mathbf{b}}\}$, which corresponds to overbid items for which j becomes the highest bidder.

A challenging part of our analysis is that the impact of bidder j's deviation strategies, which factors into the magnitude of $\Xi_k(\sigma_j)$, can be large: unlike with other approaches, such as mean field equilibrium, in which a bidder's effect on the market is infinitesimal, here a deviating bidder can sizably impact the entire market. This can be seen with overbidding, for which there exist bidding strategies $\mathbf{b}_j$ such that $\mathbf{E}[\sum_i(b_{ij} - W_i^{(2)})\mathbf{I}\{i \in U_{k,1}^{\mathbf{W}} \cap U_{j,2}^{\mathbf{b}}\}] = \Theta(\mathbf{E}[\sum_i(W_i^{(1)} - W_i^{(2)})\mathbf{I}\{i \in U_{k,1}^{\mathbf{W}}\}]) = \Theta(\tilde{\pi}_k(\sigma^*))$; that is, bidding $\mathbf{b}_j$ can succeed in artificially increasing the spend of almost all other bidders. More importantly, it does so by a quantity that washes out their potential profits under σ^*. Were it not for the additional term $\mathbf{E}[W_{i,-j}^{(2)}\mathbf{I}\{i \in U_{k,1}^{\mathbf{W}} \cap U_{j,1}^{\mathbf{b}}\}]$ that we show counterbalances $\mathbf{E}[(b_{ij} - W_i^{(2)})\mathbf{I}\{i \in U_{k,1}^{\mathbf{W}} \cap U_{j,2}^{\mathbf{b}}\}]$ and reduces the necessary $\mathbf{E}\Xi_k(\sigma_j)$, it would be too costly for bidders to sustain such an equilibrium that requires exceedingly large slack budgets.

As shown in Lemma 2, if the network's optimal allocation were greedy, bidder k could bound its spending by declaring a budget that exceeds $\bar{\Psi}_k + \Xi_k(\sigma_j)$. Hence, for a deviation σ_j, we consider the event $E_{jk}(\sigma_j)$, in which bidder k's declared budget B_k^* satisfies this condition:

$$E_{jk}(\sigma_j) \triangleq \{\bar{\Psi}_k + \Xi_k(\sigma_j) < B_k^*\}. \tag{8}$$

Finally, we define $E_j(\sigma_j)$ as the union of the events $E_{jk}(\sigma_j)$ for all bidders $k \neq j$ together with an additional condition that $\bar{\Psi}_j < B_j^*$:

$$E_j(\sigma_j) \triangleq \{\bar{\Psi}_k + \Xi_k(\sigma_j) < B_k^*, \text{ for all } k \neq j\} \cap \left\{\bar{\Psi}_j < B_j^*\right\}, \tag{9}$$

An important remark is that Lemma 2 together with conditioning on the event $E_j(\sigma_j)$ do *not* ensure that the deviating bidder j's budget constraint is slack under $(\sigma_j, \sigma_{-j}^*)$. Thus, we cannot conclude that the greedy allocation is optimal under $(\sigma_j, \sigma_{-j}^*)$. This is problematic because j's best deviations are likely to be ones in which its budget constraint is tight and for which the optimal allocation is not greedy. For example, j would not want any slack budget as that would lead to it winning unprofitable items $i \in U_{k,1}^{\mathbf{W}} \cap U_{j,1}^{\mathbf{b}}$.

The following key lemma shows that, in the event $E_j(\sigma_j)$, bidder j's modified profit under the deviation σ_j is lower than that under the strategy σ_j^*. In proving it, we close the remaining difficulty discussed in the previous paragraph and characterize how the allocation $\mathbf{Y}$ behaves even when j's budget constraint is tight.

Lemma 3. *Fix a bidder j. Suppose all other bidders $k \neq j$ play the strategy σ_k^*, whereas bidder j unilaterally deviates to a strategy σ_j. In the event $E_j(\sigma_j)$, we have, almost surely, $\tilde{\pi}_j(\sigma_j, \sigma_{-j}^*) \leq \tilde{\pi}_j(\sigma_j^*, \sigma_{-j}^*)$.*

The proof of the lemma involves a duality-based argument that exploits the LP structure of the auctioneer's optimization problem (2). To elaborate, let μ_j be the dual variable associated with each bidder's budget constraint and ν_i be the dual variable associated with each item constraint. Then, the dual of (2) is given by

$$\begin{aligned} \min_{\mu,\nu} \quad & \sum_{i\in[n]} \nu_i + \sum_{k\in[m]} B_k\mu_k + \delta n \\ \text{subject to} \quad & \nu_i + \mu_k p^{\mathbf{b}}(i,k) \geq p^{\mathbf{b}}(i,k) - \delta, \\ & \text{for each } i \in [n], k \in [m], \\ & \mu \geq 0, \nu \geq 0. \end{aligned} \tag{10}$$

Let $(\bar{\mu}, \bar{\nu})$ denote an optimal solution to the dual (10). We make two quick observations: first, the presence of a positive outside option, that is, $\delta > 0$, guarantees that $\bar{\mu}_k < 1$ for all $k \in [m]$. This is the technical reason why we incorporate the outside option δ into the ad network's optimization problem. Second, we have $\bar{\nu}_i = (\max_{\ell\in[m]}\{(1-\bar{\mu}_\ell)p^{\mathbf{b}}(i,\ell)\} - \delta)^+$ for all $i \in [n]$. Thus, the complementary slackness conditions imply that, if $Y_{ik} > 0$, then k must belong to $\arg\max_\ell\{(1-\bar{\mu}_\ell)p^{\mathbf{b}}(i,\ell)\}$.

As a step toward proving our key lemma, we first show that, in the event $E_j(\sigma_j)$, the optimal dual variable $\bar{\mu}_j$ corresponding to bidder j's budget constraint is the largest among all $\bar{\mu}_\ell$. Thus, informally, in the event $E_j(\sigma_j)$, it is bidder j's budget constraint that is most active.

Lemma 4. *Let $(\bar{\mu}, \bar{\nu})$ denote the optimal solution to the dual (10). Then, in the event $E_j(\sigma_j)$, we have $\bar{\mu}_j \geq \max_{k\neq j} \bar{\mu}_k$.*

Proof. For any fixed optimal allocation $\mathbf{Y}$ to the primal problem (2), consider the following definition: we say there is a chain from bidder k to bidder j if there exists a pair of sequences $\{k_1, \ldots, k_s\}$ and $\{i_1, \ldots, i_{s-1}\}$ such that

1. $k = k_1, j = k_s$ and $k_t \notin \{k, j\}$ for all $1 < t < s$.
2. $Y_{i_t,k_t} > 0$ for all $1 \leq t < s$.
3. $i_t \in U_{k_{t+1},1}^{\mathbf{b}}$ for all $1 \leq t < s$.

First, consider a bidder k such that there exists a chain from bidder k to bidder j. Let $\{k_1, \ldots, k_s\}$ and $\{i_1, \ldots, i_{s-1}\}$ be the sequences that form a chain from k to j. By definition, for any $1 \leq t < s$, we have $Y_{i_t,k_t} > 0$, implying by complementary slackness that

$$(1 - \bar{\mu}_{k_t})p^{\mathbf{b}}(i_t, k_t) = \bar{\nu}_{i_t} + \delta \geq (1 - \bar{\mu}_{k_{t+1}})p^{\mathbf{b}}(i_t, k_{t+1}).$$

Moreover, because $i_t \in U^{\mathbf{b}}_{k_{t+1},1}$, the payment of k_{t+1} for item i_t is at least as high as that of k_t for the same item as well as strictly positive, so we have

$$p^{\mathbf{b}}(i_t, k_t) \le p^{\mathbf{b}}(i_t, k_{t+1}) \text{ and } 0 < p^{\mathbf{b}}(i_t, k_{t+1}).$$

These two inequalities, taken together with the complementary slackness condition, imply that $\bar{\mu}_{k_t} \le \bar{\mu}_{k_t+1}$ for all $1 \le t < s$, and hence, $\bar{\mu}_k \le \bar{\mu}_j$ for any k such that there is a chain from k to j.

Next, let $\mathcal{J}_0 \triangleq \{\ell \ne j :$ there does not exist a chain from ℓ to $j\}$, and consider a bidder $k \in \arg\max\{\mu_\ell : \ell \in \mathcal{J}_0\}$. First, if $Y_{ik} = 0$ for all $i \in [n]$, then bidder k's spend equals zero. Complementary slackness then implies that $\bar{\mu}_k = 0$, and hence, $\bar{\mu}_\ell = 0 \le \bar{\mu}_j$ for all $\ell \in \mathcal{J}_0$.

Second, suppose there exists an $i \in [n]$ such that $Y_{ik} > 0$. We claim that $i \in U^{\mathbf{b}}_{k,1}$. To see this, for the sake of contradiction, suppose instead that $i \in U^{\mathbf{b}}_{q,1}$ for some $q \ne k$. Because there is no chain from k to j, it must be the case that $q \in \mathcal{J}_0$; otherwise, one can create a chain from k to j by adjoining i and k to a chain from q to j. Because $i \in U^{\mathbf{b}}_{q,1}$, we must have $p^{\mathbf{b}}(i, q) > \max\{p^{\mathbf{b}}(i, k), 0\}$. Finally, complementary slackness yields that $(1 - \bar{\mu}_k)p^{\mathbf{b}}(i, k) \ge (1 - \bar{\mu}_q)p^{\mathbf{b}}(i, q)$. From these inequalities and the fact that $\bar{\mu}_q < 1$, we obtain that $\bar{\mu}_q > \bar{\mu}_k$, contradicting the fact that $k \in \arg\max_{\ell \in \mathcal{J}_0} \bar{\mu}_\ell$. Thus, for any $i \in [m]$ such that $Y_{ik} > 0$, we have $i \in U^{\mathbf{b}}_{k,1}$. Hence, we obtain

$$\sum_{i\in[n]} p^{\mathbf{b}}(i,k) Y_{ik} \le \sum_i p^{\mathbf{b}}(i,k)\mathbf{I}\{i \in U^{\mathbf{b}}_{k,1}\} \le \bar{\Psi}_k + \Xi_k(\sigma_j), \quad (11)$$

where for the second inequality, we have used Lemma 2. Now, in the event $E_j(\sigma_j)$, we obtain that the right-hand side expression in (11) is strictly less than B^*_k, implying that the spend of bidder k is strictly below its budget. This implies, by complementary slackness, that $\bar{\mu}_k = 0$, and hence, $\bar{\mu}_\ell = 0 \le \bar{\mu}_j$ for all $\ell \in \mathcal{J}_0$. $\square$

Armed with this result, we are now ready to prove our key lemma.

Proof of Lemma 3. Consider $i \in [n]$ such that $Y_{ij} > 0$. By complementary slackness, we obtain that $(1 - \bar{\mu}_j)p^{\mathbf{b}}(i, j) \ge (1 - \bar{\mu}_k)p^{\mathbf{b}}(i, k)$ for all $k \in [m]$. By Lemma 4, $\bar{\mu}_j \ge \bar{\mu}_k$ for all $k \in [m]$ in the event $E_j(\sigma_j)$. Together with the fact that $\bar{\mu}_j < 1$, we obtain that, in the event $E_j(\sigma_j)$, if $Y_{ij} > 0$, then $p^{\mathbf{b}}(i, j) \ge p^{\mathbf{b}}(i, k)$ for all $k \in [m]$. This, in turn, implies that, in the event $E_j(\sigma_j)$, for any $i \in [n]$ with $Y_{ij} > 0$, bidder j's bid b_{ij} must be at least as high as the highest bid among all other bidders, that is, $b_{ij} \ge W^{(1)}_{i,-j}$. There are now two possibilities, depending on how the truthful bid of bidder j compares with the bids of other bidders:

a. If $W_{ij} \le W^{(1)}_{i,-j}$ or, equivalently, $i \notin U^{\mathbf{W}}_{j,1}$, then j gains $V_{ij} - W^{(1)}_{i,-j} - r_j W^{(1)}_{i,-j} = (1 + r_j)(W_{ij} - W^{(1)}_{i,-j}) \le 0$ on the item i when accounting for the cost of j's spend to win i. Thus, j does not gain from winning such items. For such items, bidding b_{ij} is worse than bidding W_{ij}.

b. If $W_{ij} > W^{(1)}_{i,-j}$, then $i \in U^{\mathbf{W}}_{j,1}$ and $p^{\mathbf{b}}(i, j) = p^{\mathbf{W}}(i, j)$. Because we condition on $E_j(\sigma_j)$ and, thus, $\bar{\Psi}_j \le B^*_j$, we have bidder j winning these same items under σ^*. This deviation is not profitable because j is not lowering the prices at which it wins these items.

This argument shows that the strategy σ_j cannot obtain more payoff for bidder j than σ^*_j, conditional on the event $E_j(\sigma_j)$, implying $\mathbf{P}(\tilde{\pi}_j(\sigma_j, \sigma^*_{-j}) \le \tilde{\pi}_j(\sigma^*_j, \sigma^*_{-j}) | E_j(\sigma_j)) = 1$. $\square$

3.2.2. Step (ii). Having stated the key lemma for the first part of our road map, which hinges on the event $E_j(\sigma_j)$ happening, we give some insight into step (ii) of the road map, in which we prove this event happens with high probability as the size of the market increases. This requires that $\bar{\Psi}_k + \Xi_k(\sigma_j) < B^*_k$ happens with high probability for all $k \ne j$. Showing this involves two steps:

a. First, we make use of the subexponential tails of the valuation distribution to show that $B^*_k - (\bar{\Psi}_k + \Xi_k(\sigma_j))$ satisfies Bernstein's concentration bound. This allows us to argue that this quantity approaches its expectation $\mathbf{E}[B^*_k] - (\mathbf{E}[\bar{\Psi}_k] + \mathbf{E}[\Xi_k(\sigma_j)])$ as n grows large, and thus, so long as $\mathbf{E}[B^*_k] - (\mathbf{E}[\bar{\Psi}_k] + \mathbf{E}[\Xi_k(\sigma_j)])$ is sufficiently positive, the required inequality holds with high probability.

b. Next, given the budget of the form $B^*_k = \mathbf{E}_{-k}[\bar{\Psi}_k] \cdot (1 + \beta(m))$, we obtain

$$\begin{aligned}\mathbf{E}[B^*_k] &- \left(\mathbf{E}[\bar{\Psi}_k] + \mathbf{E}[\Xi_k(\sigma_j)]\right) \\ &= \mathbf{E}[\bar{\Psi}_k] \cdot \beta(m) - \mathbf{E}[\Xi_k(\sigma_j)].\end{aligned}$$

Thus, for the left-hand side to be positive, we require

$$\beta(m) > \frac{\mathbf{E}\Xi_k(\sigma_j)}{\mathbf{E}\bar{\Psi}_k} = \underbrace{\frac{\mathbf{E}\left[\left(b_{ij} - W^{(2)}_i\right)\mathbf{I}\left\{i \in U^{\mathbf{W}}_{k,1} \cap U^{\mathbf{b}}_{j,2}\right\}\right] - \mathbf{E}\left[W^{(2)}_{i,-j}\mathbf{I}\left\{i \in U^{\mathbf{W}}_{k,1} \cap U^{\mathbf{b}}_{j,1}\right\}\right]}{\mathbf{E}\left[W^{(2)}_i \mathbf{I}\left\{i \in U^{\mathbf{W}}_{k,1}\right\}\right]}}_{\text{overbidding term}} + \underbrace{\frac{\mathbf{E}\left[W^{(2)}_i \mathbf{I}\left\{i \in U^{\mathbf{W}}_{j,1} \cap U^{\mathbf{W}}_{k,2}\right\}\right]}{\mathbf{E}\left[W^{(2)}_i \mathbf{I}\left\{i \in U^{\mathbf{W}}_{k,1}\right\}\right]}}_{\text{underbidding term}}. \quad (12)$$

Using the subexponentiality of the valuation function, we show that the preceding inequality is indeed satisfied by choosing a multiplicative factor $\beta(m) = 2\vartheta/m^{1/\vartheta}$ for k's slack budget as in (3). The precise details of this argument can be found in Online Appendix A. Note that this argument suggests an incentive to set a high slack factor $\beta(m)$, thus increasing the likelihood of $E_j(\sigma_j)$. However, as we show in the next step, using too large a slack eats into a bidder's profits via its cost of unspent budget.

Having established our key technical result in Lemma 3, and together with the result that $\mathbf{P}[E_j(\sigma_j)]$ is large, it is straightforward to convert a guarantee on modified profits that only holds over $E_j(\sigma_j)$ into a guarantee that holds in expectation over all $\mathbf{V}$ for modified profits $\tilde{\pi}_j$.

Lemma 5. *Let* $m \geq \max\{(2\vartheta)^{\vartheta}, \exp(3\vartheta + 1)\}$ *and* $n \geq 6\vartheta m^{2(\vartheta+1/\vartheta)}\log(m)/C$ *(where C is a universal constant independent of m). For any bidder j and for any deviation* σ_j*, we have*

$$\mathbf{E}\left[\tilde{\pi}_j\left(\sigma_j, \sigma^*_{-j}\right)\right] \leq \mathbf{E}\left[\tilde{\pi}_j(\sigma^*)\right] \cdot \left(1 + \frac{3\vartheta}{m^{3\vartheta/2}}\right).$$

3.2.3. Step (iii). To conclude with a best response guarantee, step (iii) argues that k's expected modified profit outweighs the expected cost it incurs on unspent budget given by $r_k \cdot \mathbf{E}[B^*_k - \Psi_k(\sigma^*)]$. Formally, the expected cost of unspent budget, as a fraction of the expected modified profit, is given by

$$\begin{aligned}\frac{r_k \cdot \mathbf{E}\left[B^*_k - \Psi_k(\sigma^*)\right]}{\mathbf{E}[\tilde{\pi}_k(\sigma^*)]} &\approx \frac{n \cdot \beta(m) \cdot r_k \cdot \mathbf{E}\left[W_i^{(2)} \mathbf{I}\left\{i \in U^{\mathbf{W}}_{k,1}\right\}\right]}{n \cdot (1 + r_k) \cdot \mathbf{E}\left[\left(W_i^{(1)} - W_i^{(2)}\right)\mathbf{I}\left\{i \in U^{\mathbf{W}}_{k,1}\right\}\right]} \\ &= \frac{r_k}{1 + r_k} \cdot \frac{\beta(m)}{\alpha(m)},\end{aligned} \tag{13}$$

where $\alpha(m) \triangleq \frac{\mathbf{E}[(W_i^{(1)} - W_i^{(2)})\mathbf{I}\{i \in U^{\mathbf{W}}_{k,1}\}]}{\mathbf{E}[W_i^{(2)}\mathbf{I}\{i \in U^{\mathbf{W}}_{k,1}\}]}$ depends only on the valuation distribution.

As becomes apparent in the proof of Theorem 1, this ratio also drives the approximation error of our equilibrium. Thus, the ratio $\beta(m)/\alpha(m)$ must be small for this approximation to be good. Because $\alpha(m)$ is fixed by the valuation distribution, this imposes the condition that the multiplicative factor $\beta(m)$ is not too large, competing with the incentive for larger $\beta(m)$ from step (ii). As our analysis in Online Appendix B shows, these two competing incentives are perfectly counterbalanced by the choice, as in (3), of $\beta(m) = 2\vartheta/m^{1/\vartheta}$.

Given the preceding intuition and results, we can now prove our main result, which translates the bound on $\tilde{\pi}_j$ to a bound for π_j by accounting for the additional cost of $r_j \cdot \mathbf{E}[B^*_j - \Psi_j(\sigma^*)]$ and implicitly showing that $\beta(m)/\alpha(m)$ is small enough to sustain the equilibrium:

Proof of Theorem 1. Throughout the proof, let $m \geq \max\{(2\vartheta)^{\vartheta}, \exp(3\vartheta + 5\vartheta\log(\vartheta) + 1)\}$ and $n \geq 6\vartheta m^{2(\vartheta+1/\vartheta)}\log(m)/C$. As proved in Lemma EC.11,

$$\mathbf{E}\left[\tilde{\pi}_j(\sigma^*)\right] \geq \frac{9(1 + r_j)}{10\vartheta\log(m)} \cdot \mathbf{E}\left[\bar{\Psi}_j\right], \tag{14}$$

and at the same time,

$$\begin{aligned}\mathbf{E}\left[B^*_j - \Psi_j(\sigma^*)\right] &= \mathbf{E}\left[\mathbf{E}_{-j}\left[\bar{\Psi}_j\right] \cdot \left(1 + \frac{2\vartheta}{m^{1/\vartheta}}\right) - \Psi_j(\sigma^*)\right] \\ &= \mathbf{E}\left[\bar{\Psi}_j\right] \cdot \left(1 + \frac{2\vartheta}{m^{1/\vartheta}}\right) - \mathbf{E}\left[\Psi_j(\sigma^*)\right] \\ &\leq \left(\frac{2\vartheta}{m^{1/\vartheta}} + \frac{1}{m^{1/\vartheta}}\right) \cdot \mathbf{E}\left[\bar{\Psi}_j\right] \\ &\leq \frac{3\vartheta}{m^{1/\vartheta}} \cdot \mathbf{E}\left[\bar{\Psi}_j\right],\end{aligned} \tag{15}$$

where, in the first equality, we have used the definition (3) of B^*_j, the tower property, to remove the conditioning on $\mathbf{V}_j$ in the second equality and Lemma EC.10 for the third inequality.

Because, by definition, $\mathbf{E}[\pi_j(\sigma^*)] = \mathbf{E}[\tilde{\pi}_j(\sigma^*)] - r_j\mathbf{E}[B^*_j - \Psi_j(\sigma^*)]$, we can combine inequalities (14) and (15) to obtain

$$\begin{aligned}\frac{\mathbf{E}\left[\pi_j(\sigma^*)\right]}{\mathbf{E}\left[\tilde{\pi}_j(\sigma^*)\right]} &= 1 - r_j\frac{\mathbf{E}\left[B^*_j - \Psi_j(\sigma^*)\right]}{\mathbf{E}\left[\tilde{\pi}_j(\sigma^*)\right]} \geq 1 - r_j\frac{\frac{3\vartheta}{m^{1/\vartheta}}\mathbf{E}\left[\bar{\Psi}_j\right]}{\frac{9(1+r_j)}{10\vartheta\log(m)}\mathbf{E}\left[\bar{\Psi}_j\right]} \\ &\geq 1 - \frac{10 r_j\vartheta^2\log(m)}{3(1 + r_j)m^{1/\vartheta}} \geq 1 - \frac{10\vartheta^2\log(m)}{3m^{1/\vartheta}}.\end{aligned} \tag{16}$$

Now, for $m \geq \exp(3\vartheta + 5\vartheta\log(\vartheta))$, we have that the final expression is greater than $1/3$. Thus, we have, for any strategy σ_j,

$$\begin{aligned}\mathbf{E}\left[\pi_j\left(\sigma_j, \sigma^*_{-j}\right)\right] &\leq \mathbf{E}\left[\tilde{\pi}_j\left(\sigma_j, \sigma^*_{-j}\right)\right] \\ &\leq \mathbf{E}\left[\tilde{\pi}_j(\sigma^*)\right] \cdot \left(1 + \frac{3\vartheta}{m^{3\vartheta/2}}\right) \\ &\leq \mathbf{E}\left[\pi_j(\sigma^*)\right] \cdot \frac{\left(1 + \frac{3\vartheta}{m^{3\vartheta/2}}\right)}{\left(1 - \frac{10\vartheta^2\log(m)}{3m^{1/\vartheta}}\right)} \\ &\leq \mathbf{E}\left[\pi_j(\sigma^*)\right] \cdot \left(1 + \frac{11\vartheta^2\log(m)}{m^{1/\vartheta}}\right),\end{aligned}$$

where the second inequality follows from Lemma 5, the third from (16), and the fourth inequality follows from the fact that, for $\epsilon \leq 2/3$, we have $\frac{1}{1-\epsilon} \leq 1 + 3\epsilon$. This completes the proof. □

To end this section, we provide a proof of Theorem 2, which shows that, in equilibrium, the allocation policy is greedy with high probability. We begin with the following lemma, which is a consequence of our key technical lemma, Lemma 4, and which states that, for the particular case when $\sigma_j = \sigma^*_j$ and all other bidders follow σ^*, the optimal allocation $\mathbf{Y}$ is greedy.

Lemma 6. *Fix a bidder j, and suppose all bidders k play the strategy σ_k^*. Then, in the event $E_j(\sigma_j^*)$, the optimal allocation is greedy.*

Proof. Fix the strategy profile σ^*. For a given realization of valuations $\mathbf{V}$ belonging to the event $E_j(\sigma_j^*)$, let $\mathbf{Y}$ be an optimal allocation, and let $(\bar{\mu}, \bar{\nu})$ denote an optimal solution to the dual (10). By Lemma 4, in the event $E_j(\sigma_j^*)$, we have $1 > \bar{\mu}_j \geq \max_{k \neq j} \bar{\mu}_k$. Hence, in this event, by complementary slackness, for any $i \in [n]$ with $Y_{ij} > 0$, we have $p^{\mathbf{W}}(i,j) \geq p^{\mathbf{W}}(i,k)$ for all $k \in [m]$. Thus, we obtain that, in the event $E_j(\sigma_j^*)$, if $Y_{ij} > 0$, then $i \in U_{j,1}^{\mathbf{W}}$. Hence, we obtain

$$\sum_{i \in [n]} Y_{i,j} p^{\mathbf{W}}(i,j) \leq \sum_{i \in [n]} W_i^{(2)} \mathbf{I}\left\{i \in U_{j,1}^{\mathbf{W}}\right\} = \bar{\Psi}_j < B_j^*.$$

Thus, in the event $E_j(\sigma_j^*)$, we observe that bidder j's budget constraint is not tight, and hence, by complementary slackness, we obtain $\bar{\mu}_j = 0$. Because, in this event, we have $\bar{\mu}_j \geq \bar{\mu}_k$ for all $k \in [m]$, we conclude that $\bar{\mu}_k = 0$ for all $k \in [m]$. Moreover, for any bidder k, if $Y_{ik} > 0$, then $p^{\mathbf{W}}(i,k) \geq p^{\mathbf{W}}(i,\ell)$ for all ℓ by complementary slackness. Because the budget constraints are not tight, we conclude that the optimal allocation is unique and greedy. □

Theorem 2 now follows directly from this and is formally proved in Online Appendix B.

4. Extensions

In this section, we explore two extensions to our base model from Section 2. Specifically, we show that our equilibrium can be sustained when the ad network sets reserve prices as well as if we allow more general valuation distributions. We discuss these new model features in more detail:

4.1. Reserve Prices

In our baseline model, we assume that the allocation mechanism implemented by the ad network is a second price auction. Our results can be extended to accommodate reserve prices, a common feature of modern ad-serving systems. Specifically, we allow the ad network to set a reserve price ρ, which places a floor on the price at which an item can be sold to a bidder; in such a mechanism, only bidders who bid over this reserve can be awarded the item, for a payment equal to the maximum of the next highest bid and the reserve price. Thus, the definition (1) of $p^{\mathbf{b}}(i,j)$ changes to

$$p^{\mathbf{b}}(i,j) = \begin{cases} \max\left\{\max_{k \neq j, b_{ik} \leq b_{ij}} \{b_{ik}\}, \rho\right\} & \text{if } b_{ij} \geq \rho; \\ 0 & \text{otherwise.} \end{cases} \tag{17}$$

We place two bounds on the magnitude of the reserve prices that we allow the network to set:

a. We require that ρ satisfies

$$\mathbf{P}\left(W_i^{(2)} \geq \rho \,\middle|\, i \in U_{k,1}^{\mathbf{W}}\right) \geq \eta, \quad \text{for any } i \in [n] \text{ and } k \in [m]. \tag{18}$$

This is for some fixed $\eta > 0$. The interpretation of this condition is that reserve prices should not be so high that the ROIs of bidders are effectively nullified under σ^*, and who then do not have an incentive to participate in the market. This could happen because, when we introduce a reserve, the analog to the quantity $\alpha(m)$ introduced in (13) becomes

$$\alpha(m) = \frac{\mathbf{E}\left[W_i^{(1)} - \max\left\{W_i^{(2)}, \rho\right\} \,\middle|\, i \in U_{k,1}^{\mathbf{W}}, W_i^{(1)} \geq \rho\right]}{\mathbf{E}\left[\max\left\{W_i^{(2)}, \rho\right\} \,\middle|\, i \in U_{k,1}^{\mathbf{W}}, W_i^{(1)} \geq \rho\right]}.$$

One can show that $\alpha(m)$, in fact, vanishes as $\rho \to \infty$, but at the same time, the slack term $\beta(m)$ required for equilibrium to hold remains the same up to constant factors. Thus, the ratio $\beta(m)/\alpha(m)$ increases as $\rho \to \infty$, and the natural counterpart of σ^* to the reserve setting is unlikely to be sustained as an equilibrium.

We remark that this condition does not, however, remove reasonable reserve values. For example, if $\eta = 1/2$ and bidders are symmetric, the condition allows $\rho = \text{median}\,(W_i^{(2)}) \approx \mathbf{E}W_i^{(2)}$, which is the expected payment one would see in equilibrium.

b. We also assume that the reserve price ρ exceeds the outside option value δ, which is a mild technical assumption. Recall that the presence of the δ terms in problem (2) is required to remove some degenerate allocations by ensuring that $\mu_\ell < 1$ for all bidders ℓ and can be set to an arbitrarily small but positive value. Thus, requiring that $\rho > \delta$ does not substantively restrict the space of allowable reserves. At the same time, this simplifies the analysis: whereas in the previous section, we had to account for the event that a bidder's payment could be positive but below δ, in which case the bidder would not be allocated the item, here a positive payment automatically satisfies that $p^{\mathbf{b}}(i,j) \geq \rho > \delta$.

Under these two conditions, our main theorems continue to hold with a minor modification. Specifically, the bidders' strategies in the (approximate) equilibrium $\sigma^{*,\text{res}} = (B_j^{*,\text{res}}, \mathbf{b}_j^{*,\text{res}})$ now reflect the change in their expected spending as follows:

$$\mathbf{b}_j^{*,\text{res}} \triangleq \mathbf{W}_j = \frac{\mathbf{V}_j}{1 + r_j},$$
$$B_j^{*,\text{res}} \triangleq \mathbf{E}_{-j}\left[\bar{\Psi}_j^{\text{res}}\right] \cdot \left(1 + \frac{2\vartheta}{\eta m^{1/\vartheta}}\right), \tag{19}$$

where the quantity $\bar{\Psi}_j^{\text{res}}$ once again describes the total realized spending of bidder j under infinite budgets:

$$\bar{\Psi}_j^{\text{res}} \triangleq \sum_{i=1}^{n} \max\left\{W_i^{(2)}, \rho\right\} \mathbf{I}\left\{i \in U_{j,1}^{\mathbf{W}}, W_i^{(1)} \geq \rho\right\}. \quad (20)$$

The theorems characterizing this equilibrium are qualitatively identical to Theorems 1 and 2 and are, thus, stated in Online Appendix F. However, we comment on the changes we need to make to the proofs to accommodate reserves. Mainly, the definitions of the corresponding event $E_j^{\text{res}}(\sigma_j)$ and the slack $\Xi_{j,k}^{\text{res}}(\sigma_j)$ change to incorporate the changes in the definitions of $p^{\mathbf{b}}(i,j)$ and the allocation mechanism:

$$\begin{aligned} E_j^{\text{res}}(\sigma_j) \triangleq &\left\{\bar{\Psi}_k^{\text{res}} + \Xi_k^{\text{res}}(\sigma_j) < B_k^{*,\text{res}}, \text{ for all } k \neq j\right\} \\ &\cap \left\{\bar{\Psi}_j^{\text{res}} < B_j^{*,\text{res}}\right\}, \end{aligned} \quad (21)$$

where

$$\begin{aligned} \Xi_k^{\text{res}}(\sigma_j) \triangleq &\sum_i \left(\max\{b_{ij}, \rho\} - \max\left\{W_i^{(2)}, \rho\right\}\right) \\ &\cdot \mathbf{I}\left\{i \in U_{k,1}^{\mathbf{W}} \cap U_{j,2}^{\mathbf{b}}, W_i^{(1)} \geq \rho\right\} \\ &- \sum_i \max\left\{W_{i,-j}^{(2)}, \rho\right\} \mathbf{I}\left\{i \in U_{k,1}^{\mathbf{W}} \cap U_{j,1}^{\mathbf{b}}, W_i^{(1)} \geq \rho\right\} \\ &+ \sum_i \max\left\{W_i^{(2)}, \rho\right\} \mathbf{I}\left\{i \in U_{j,1}^{\mathbf{W}} \cap U_{k,2}^{\mathbf{W}}, W_i^{(1)} \geq \rho\right\}, \end{aligned} \quad (22)$$

and note that $\Xi_k^{\text{res}}(\sigma_j)$ differs from $\Xi_k(\sigma_j)$ in two respects: First, to reflect the change in allocation mechanism, $\Xi_k^{\text{res}}(\sigma_j)$ includes the indicator $\mathbf{I}\{W_i^{(1)} \geq \rho\}$ to capture the fact that item is allocated under truthful bidding. Second, to reflect the change in bidders' payments, the quantities $W_i^{(2)}$, $W_{i,-j}^{(2)}$, and b_{ij} are now replaced by $\max\{W_i^{(2)}, \rho\}$, $\max\{W_{i,-j}^{(2)}, \rho\}$, and $\max\{b_{ij}, \rho\}$, respectively. With these changes in the place, our concentration bounds can once again be shown to hold with appropriate changes. We summarize the main changes in Online Appendix F.

4.2. Bounded Hazard Rate Valuation Distributions

Although our existing analysis has relied on the condition that bidder valuations are exponentially distributed, we expect that our results carry over for a significantly wider family of distributions. Here, we prove a result in this direction by extending our equilibrium prescription to bidder valuations with bounded hazard rate distributions.

To highlight the main differences from the baseline analysis from Section 3, we make a simplification by assuming that bidders are *symmetric* in the sense that bidders' discounted valuations $W_{ij} = V_{ij}/(1 + r_j)$ are independently and identically distributed and drawn from a distribution with cumulative distribution function $F(\cdot)$ with continuous hazard rate $h(\cdot)$ that satisfies the bounds

$$\theta_{\min} \leq h(v) \leq \theta_{\max}, \ \forall v \in [0, \infty). \quad (23)$$

In a slight abuse of notation, we continue to define $\vartheta \triangleq \theta_{\max}/\theta_{\min}$. The symmetry assumption allows us to focus on the most salient changes while keeping the analysis tractable.

For bounded hazard rate valuations, our main result that σ^* is an approximate equilibrium carries through as is, whereas our proof needs some minor modifications. We describe these modifications in some detail in Online Appendix G.

4.3. Generalizing Our Results to Other Distributions

Because we conjecture that our results would hold for more general families of valuation distributions, we also discuss how one could potentially further relax our distributional assumptions.

Interestingly, an artifact of our analysis is that it requires a *lower* bound on the hazard rates of the valuations. This is because, as Equations (12) and (13) together with the related discussion imply, the valuation distributions should allow that the ratio $\beta(m)/\alpha(m)$ is small, where

$$\frac{\beta(m)}{\alpha(m)} \approx \underbrace{\frac{\mathbf{E}\left[\left(b_{ij} - W_i^{(2)}\right)\mathbf{I}\left\{i \in U_{k,1}^{\mathbf{W}} \cap U_{j,2}^{\mathbf{b}}\right\}\right] - \mathbf{E}\left[W_{i,-j}^{(2)}\mathbf{I}\left\{i \in U_{k,1}^{\mathbf{W}} \cap U_{j,1}^{\mathbf{b}}\right\}\right]}{\mathbf{E}\left[\left(W_i^{(1)} - W_i^{(2)}\right)\mathbf{I}\left\{i \in U_{k,1}^{\mathbf{W}}\right\}\right]}}_{\text{overbidding term}} + \underbrace{\frac{\mathbf{E}\left[W_i^{(2)}\mathbf{I}\left\{i \in U_{j,1}^{\mathbf{W}} \cap U_{k,2}^{\mathbf{W}}\right\}\right]}{\mathbf{E}\left[\left(W_i^{(1)} - W_i^{(2)}\right)\mathbf{I}\left\{i \in U_{k,1}^{\mathbf{W}}\right\}\right]}}_{\text{underbidding term}}.$$

We remark that, as discussed in Section 3.2, the first term in the right-hand side can be interpreted as coming from overbidding-type behavior, whereas the second is attributed to underbidding. In order to bound the numerator of the overbidding term over all possible deviations $\mathbf{b}_j$, our proof technique requires that $h_j(v) \geq \theta_{\min}$, where $\theta_{\min} \geq 0$ and $h_j(\cdot)$ is the hazard rate of W_{ij} (note that, here, the distributions of W_{ij} could be asymmetric). This is a property we use in Lemma EC.8 as we discuss following its proof. We observe that this lower bound is also necessary for the Bernstein concentration bounds that are crucial to our proofs.

As we briefly mention in Section 2, in an ad network, the bounded hazard rate condition implies a large degree of heterogeneity in each advertiser's value for the impressions. This may arise when the ad network and advertisers have access to fine-grained covariates (such as user demographics, purchase history, etc.) for each impression, which are used for displaying highly targeted ads. Conversely, without a minimal level of targeting, an advertiser's value for an impression is likely to be more concentrated. Thus, our analysis and the results suggest that better targeting may give rise to simpler equilibrium outcomes.

In addition to the hazard rate lower bound, bounding the magnitude of the overbidding term also requires the following tail bound condition:

$$\frac{1}{\theta_{\min}} \cdot \mathbf{P}\left(W_{i,-j}^{(2)} \leq \frac{1}{\theta_{\min}}, i \in U_{k,1}^{\mathbf{W}}\right) = o(1) \cdot \mathbf{E}\left[\left(W_i^{(1)} - W_i^{(2)}\right)\mathbf{I}\left\{i \in U_{k,1}^{\mathbf{W}}\right\}\right],$$

which, for exponential valuations, we prove via Lemmas EC.6 and EC.12.

Finally, bounding the underbidding term requires that

$$\mathbf{E}\left[W_i^{(2)}\mathbf{I}\left\{i \in U_{j,1}^{\mathbf{W}} \cap U_{k,2}^{\mathbf{W}}\right\}\right] = o(1) \cdot \mathbf{E}\left[\left(W_i^{(1)} - W_i^{(2)}\right)\mathbf{I}\left\{i \in U_{k,1}^{\mathbf{W}}\right\}\right].$$

For the case of exponential distributions, we establish this requirement in Lemmas EC.5 and EC.12. We note that, for the bounded hazard rate case, the assumption of symmetry of W_{ij}s simplifies the analysis of both these last two properties.

Before concluding, we emphasize that the two preceding conditions, although difficult to characterize or check analytically, are easy to check via simulation. Thus, a decision maker could use samples from the valuation distributions to verify whether our proposed equilibrium holds.

5. Discussion and Conclusions

In this paper, we examine the strategic bidding behavior of advertisers in a sponsored search market. Although this problem appears to be quite challenging from both game-theoretic and optimal matching perspectives, we find that, under the assumption that budgets arise out of endogenous bidder behavior, simple tractable equilibria emerge. Moreover, these equilibria are attractive because they also simplify the ad network's decision problem into a highly tractable greedy policy.

An interesting direction to pursue would be to study the market equilibrium in a dynamic setting in which the items arrive over time and bidders can adjust their bids dynamically. There are two main considerations regarding porting our results to this dynamic setting: first, whether our equilibrium strategies, which assume the knowledge of the entire valuation vector $\mathbf{V}_i$, are implementable in a setting in which valuations are revealed as and when impressions arrive and, second, even if such implementation were feasible, whether our equilibrium guarantees carry over in settings in which bidders can adjust their bids over time.

To address the first consideration, note that, as the equilibrium bids are shaded versions of the valuations, a bidder can equivalently specify a bidding function that outputs the bid as and when the impressions arrive; this is akin to contingent bidding in which bidders submit a vector of bids from which a true bid is selected upon the resolution of the contingency. Similarly, each bidder can set its equilibrium budget prior to seeing the valuations through a slight modification to our equilibrium in which the budget is set as the unconditional expectation of the bidder's spend (plus a slack); because the spend concentrates in the large market limit we consider, this modification does not affect our analysis.

For the second consideration, although formalizing the dynamic setting is more intricate than the static version studied here, we expect that such adjustments would not impact our equilibrium guarantees. There are a few reasons for this: First, our existing analysis already provides ex ante guarantees for the equilibrium; this makes use of the fact that in a fluid regime, variability in a bidder's valuation realizations washes away. In this setting, one can show that, if other bidders follow static strategies, a single bidder's utility does not substantially improve by adopting a dynamic over a static strategy. Second, we note that our static model allows for more deviations for a bidder, in which it can base its bids on the entire valuation vector. In contrast, in a dynamic model, the deviations are more restricted as they cannot depend on future valuations. Thus, we believe our guarantees are fairly conservative in a dynamic setting.

Another question that our work opens up is whether our equilibrium, although simple and plausible, is unique. Although this question is already significant in the static setting, it becomes even more pressing when we consider dynamics. To highlight the concern, the ability of bidders to dynamically change bids can potentially reduce the efficacy of the bidders' deviations; this leaves open the possibility that other dynamic bidding equilibria may emerge in such a setting. If other equilibria exist, further justifications are needed that the equilibrium we study would arise given it leads to a highly competitive market. Thus, an interesting open problem is to identify conditions for uniqueness for notions of approximate equilibria or to pose strong criteria for equilibrium selection that select equilibria such as ours.

Acknowledgments

The authors thank Gustavo Vulcano (the department editor handling this paper), the anonymous associate editor, and reviewers for comments, which helped to substantially improve both the content and exposition of this work. Portions of this work were completed when Krishnamurthy Iyer was with Cornell University.

Endnotes

[1] We thank the anonymous referee for this example about Facebook's ad system.

[2] As stated, the minimum number of items n needed for Theorem 2 to hold is larger than that in Theorem 1 by a factor of m. This additional factor allows us to establish the optimality of the greedy solution except with an *exponentially* small probability. With the same bound on n as in Theorem 1, our proof method can be used to obtain optimality of greedy except for a *polynomially* small probability.

References

Abrams Z, Mendelevitch O, Tomlin J (2007) Optimal delivery of sponsored search advertisements subject to budget constraints. *Proc. 8th ACM Conf. Electronic Commerce* (ACM, New York), 272–278.

Agrawal S, Devanur NR (2015) Fast algorithms for online stochastic convex programming. *Proc. 26th Annual ACM-SIAM Sympos. Discrete Algorithms* (Society for Industrial and Applied Mathematics, Philadelphia), 1405–1424.

Asadpour A, Bateni MH, Bhawalkar K, Mirrokni V (2014) Concise bid optimization strategies with multiple budget constraints. *Internat. Conf. Web Internet Econom.* (Springer, Cham), 263–276.

Athey S, Segal I (2013) An efficient dynamic mechanism. *Econometrica* 81(6):2463–2485.

Auerbach J, Galenson J, Sundararajan M (2008) An empirical analysis of return on investment maximization in sponsored search auctions. *Proc. 2nd Internat. Workshop Data Mining Audience Intelligence Advertising* (ACM, New York), 1–9.

Babaioff M, Sharma Y, Slivkins A (2009) Characterizing truthful multi-armed bandit mechanisms: Extended abstract. *Proc. 10th ACM Conf. Electronic Commerce* (ACM, New York), 79–88.

Balseiro S, Gur Y (2019) Learning in repeated auctions with budgets: Regret minimization and equilibrium. *Management Sci.* 65(9): 3952–3968.

Balseiro S, Kim A, Mahdian M, Mirrokni V (2017) Budget management strategies in repeated auctions. *Proc. 26th Internat. Conf. World Wide Web* (International World Wide Web Conferences Steering Committee, Republic and Canton of Geneva, Switzerland), 15–23.

Balseiro SR, Besbes O, Weintraub GY (2015) Repeated auctions with budgets in ad exchanges: Approximations and design. *Management Sci.* 61(4):864–884.

Bergemann D, Välimäki J (2010) The dynamic pivot mechanism. *Econometrica* 78(2):771–789.

Borgs C, Chayes J, Immorlica N, Jain K, Etesami O, Mahdian M (2007) Dynamics of bid optimization in online advertisement auctions. *Proc. 16th Internat. Conf. World Wide Web* (ACM, New York), 531–540.

Buchbinder N, Jain K, Naor JS (2007) Online primal-dual algorithms for maximizing ad-auctions revenue. Arge L, Hoffmann M, Welzl E, eds. *Algorithms – ESA 2007* (Springer, Berlin), 253–264.

Charles D, Chakrabarty D, Chickering M, Devanur NR, Wang L (2013) Budget smoothing for internet ad auctions: A game theoretic approach. *Proc. 14th ACM Conf. Electronic Commerce* (ACM, New York), 163–180.

Ciocan DF, Farias V (2012) Model predictive control for dynamic resource allocation. *Math. Oper. Res.* 37(3):501–525.

Conitzer V, Kroer C, Sodomka E, Stier-Moses NE (2017) Multiplicative pacing equilibria in auction markets. Preprint, submitted February 8, https://arxiv.org/abs/1706.07151.

Devanur NR, Hayes TP (2009) The adwords problem: Online keyword matching with budgeted bidders under random permutations. *Proc. 10th ACM Conf. Electronic Commerce* (ACM, New York), 71–78.

Devanur NR, Kakade SM (2009) The price of truthfulness for pay-per-click auctions. *Proc. 10th ACM Conf. Electronic Commerce* (ACM, New York), 99–106.

Dütting P, Fischer F, Parkes DC (2016) Non-truthful position auctions are more robust to misspecification. Preprint, submitted June 21, https://arxiv.org/abs/1602.07593.

Edelman B, Ostrovsky M, Schwarz M (2007) Internet advertising and the generalized second-price auction: Selling billions of dollars worth of keywords. *Amer. Econom. Rev.* 97(1):242–259.

Facebook (2019) Facebook ads help center. Accessed August 25, 2019, https://www.facebook.com/business/help.

Feldman J, Mehta A, Mirrokni V, Muthukrishnan S (2009) Online stochastic matching: Beating 1-1/e. *50th Annual IEEE Sympos. Foundations Comput. Sci.* (The Institute of Electrical and Electronics Engineers, Washington, DC), 117–126.

Feldman J, Henzinger M, Korula N, Mirrokni VS, Stein C (2010) Online stochastic packing applied to display ad allocation. *Eur. Sympos.* Algorithms (Springer, Berlin), 182–194.

Gallego G, Van Ryzin G (1997) A multiproduct dynamic pricing problem and its applications to network yield management. *Oper. Res.* 45(1):24–41.

Goel A, Mahdian M, Nazerzadeh H, Saberi A (2010) Advertisement allocation for generalized second-pricing schemes. *Oper. Res. Lett.* 38(6):571–576.

Gummadi R, Key PB, Proutiere A (2011) Optimal bidding strategies in dynamic auctions with budget constraints. *49th Annual Allerton Conf. Comm., Control, Comput.* (The Institute of Electrical and Electronics Engineers, Washington, DC), 588.

Iyer K, Johari R, Sundararajan M (2014) Mean field equilibria of dynamic auctions with learning. *Management Sci.* 60(12): 2949–2970.

Jasin S, Kumar S (2012) A re-solving heuristic with bounded revenue loss for network revenue management with customer choice. *Math. Oper. Res.* 37(2):313–345.

Kanoria Y, Nazerzadeh H (2020) Dynamic reserve prices for repeated auctions: Learning from bids. Preprint, submitted February 18, https://arxiv.org/abs/2002.07331.

Karande C, Mehta A, Srikant R (2013) Optimizing budget constrained spend in search advertising. *Proc. 6th ACM Internat. Conf. Web Search Data Mining* (ACM, New York), 697–706.

Karp RM, Vazirani UV, Vazirani VV (1990) An optimal algorithm for on-line bipartite matching. *Proc. 22nd Annual ACM Sympos. Theory Comput.* (ACM, New York), 352–358.

Mehta A (2013) Online matching and ad allocation. *Foundations and Trends® in Theoretical Computer Science*, vol 8, no. 4 (now Publishers Inc., Hanover, MA), 265–368.

Mehta A, Saberi A, Vazirani U, Vazirani V (2007) Adwords and generalized online matching. *J. ACM* 54(5):22-es.

Mirrokni V, Nazerzadeh H (2017) Deals or no deals: Contract design for online advertising. *Proc. 26th Intern. Conf. World Wide Web (WWW '17)*. International World Wide Web Conferences Steering Committee (Republic and Canton of Geneva, CHE), 7–14.

Nazerzadeh H, Saberi A, Vohra R (2013) Dynamic pay-per-action mechanisms and applications to online advertising. *Oper. Res.* 61(1):98–111.

Talluri K, Van Ryzin G (1998) An analysis of bid-price controls for network revenue management. *Management Sci.* 44(11): 1577–1593.

Varian HR (2007) Position auctions. *Internat. J. Indust. Organ.* 25(6):1163–1178.

Varian HR (2009) Online ad auctions. *Amer. Econom. Rev.* 99(2):430–434.

Wilkens CA, Cavallo R, Niazadeh R (2017) GSP: The Cinderella of mechanism design. *Proc. 26th Internat. Conf. World Wide Web*, (ACM, New York) 25–32.
Zhou Y, Chakrabarty D, Lukose R (2008) *Budget Constrained Bidding in Keyword Auctions and Online Knapsack Problems* (Springer, Berlin), 566–576.

Dragos Florin Ciocan is an assistant professor of Technology and Operations Management at INSEAD. His research interests are in stochastic optimization, large-scale distributed optimization, and nontraditional applications of revenue management.

Krishnamurthy Iyer is an associate professor in the Department of Industrial and Systems Engineering at the University of Minnesota. His research interests include game theory, mechanism design, and stochastic modeling with applications in markets and service systems.