MANAGEMENT SCIENCE
Vol. 61, No. 4, April 2015, pp. 864–884
ISSN 0025-1909 (print) | ISSN 1526-5501 (online)
http://dx.doi.org/10.1287/mnsc.2014.2022

Repeated Auctions with Budgets in Ad Exchanges: Approximations and Design

Santiago R. Balseiro
Fuqua School of Business, Duke University, Durham, North Carolina 27708, srb43@duke.edu

Omar Besbes, Gabriel Y. Weintraub
Graduate School of Business, Columbia University, New York, New York 10027
{ob2105@columbia.edu, gyw2105@columbia.edu}

Ad exchanges are emerging Internet markets where advertisers may purchase display ad placements, in real time and based on specific viewer information, directly from publishers via a simple auction mechanism. Advertisers join these markets with a prespecified budget and participate in multiple second-price auctions over the length of a campaign. This paper studies the competitive landscape that arises in ad exchanges and the implications for publishers' decisions. The presence of budgets introduces dynamic interactions among advertisers that need to be taken into account when attempting to characterize the bidding landscape or the impact of changes in the auction design. To this end, we introduce the notion of a fluid mean-field equilibrium (FMFE) that is behaviorally appealing and computationally tractable, and in some important cases, it yields a closed-form characterization. We establish that an FMFE approximates well the rational behavior of advertisers in these markets. We then show how this framework may be used to provide sharp prescriptions for key auction design decisions that publishers face in these markets. In particular, we show that ignoring budgets, a common practice in this literature, can result in significant profit losses for the publisher when setting the reserve price.

Keywords: auction design; revenue management; ad exchange; display advertising; Internet; budget constraints; dynamic games; mean field; fluid approximation
History: Received September 19, 2012; accepted June 6, 2014, by Dimitris Bertsimas, optimization. Published online in *Articles in Advance* January 27, 2015.

1. Introduction

The market for display ads on the Internet, consisting of graphical content such as banners and videos on Web pages, has grown significantly in the last decade, generating about $11 billion in the United States in 2012 (Internet Advertising Bureau 2013). This growth has been accompanied by the emergence of alternative channels for the purchase of display ads. Whereas traditionally, advertisers would purchase display ad placements by negotiating long-term contracts directly with publishers (Web page owners), spot markets for ad slots, called ad exchanges, have emerged, and the ad spending through these continues to grow (Vranica 2013). Google's DoubleClick, OpenX, and Yahoo!'s Right Media are examples of such exchanges.

An ad exchange is a platform that operates as an intermediary between online publishers and advertisers. When a user visits a Web page (e.g., the *New York Times* online), the publisher may post an ad slot in the exchange potentially together with some user information known to her, such as the user's geographical location and her cookies. Based on this latter information, and in conjunction with their targeting criteria, interested advertisers (or bidders) post bids. Then, an auction is run to determine the winner and the ad to be shown to the user. The latter process happens in milliseconds, between the time a user requests a page and the time the page is displayed. As viewers visit her website, the publisher repeatedly offers slots to display advertisements; typically, a given publisher runs millions of these auctions per day. For their part, advertisers participate in the exchange with the objective of fulfilling marketing campaigns. In practice, such campaigns are commonly based on a predetermined budget and extend for a fixed amount of time over which advertisers participate in a large volume of auctions. Given the large number opportunities and the time scale on which decisions are made, bidding is fully automated. See Muthukrishnan (2009) for a more detailed description of ad exchanges.

The prevalence of advertisers' budget constraints in these markets links the different auctions over time, and therefore, traditional equilibrium and revenue optimization analysis for static auctions do not apply in this setting. Thus motivated, this paper introduces a new approach to study the key auction design decisions that publishers face while considering the strategic response of budget-constrained advertisers. In particular, the framework captures some key characteristics of an exchange and allows us to start quantifying some central trade-offs faced by publishers and advertisers in this new channel.

1.1. Main Contributions

Advertisers participate in repeated auctions subject to budget constraints, and therefore they typically require dynamic bidding strategies to optimize the allocation of budget to incoming impressions in order to maximize cumulated profits over the length of the campaign. In many cases, advertisers may have similar targeting criteria and bid for the same inventory of ads. Thus, the dynamic bidding strategy an advertiser adopts impacts the competitive landscape for other advertisers in the market. Moreover, the publisher's auction design decisions, such as the reserve price, also impact these interactions. Thus motivated, we formulate our ad exchange model as a game among advertisers and the publisher.[1] First, the publisher defines the parameters of a second-price auction that become common knowledge. Then, given the auction format, advertisers compete in a *dynamic* game. To quantify the impact of auction design parameters, the first question pertains to the competitive landscape that emerges for fixed auction decisions.

An important challenge in our analysis is solving for the equilibrium of the dynamic game among advertisers induced by the auction rules. At one extreme of agent sophistication, one may consider traditional game-theoretic notions of equilibrium such as perfect Bayesian equilibrium (PBE), in which advertisers maintain priors on the states of all other bidders and update them accordingly using Bayes' rule. Such an approach presents two main drawbacks. First, the analysis of the resulting game is, in most cases, intractable from both analytical and computational standpoints. Second, such sophistication and informational requirements on the part of agents is highly unrealistic.

1.1.1. Fluid Mean-Field Equilibrium. The main contribution of this paper is the introduction of an equilibrium notion that is tractable and provides a good approximation to the strategic interactions among budget-constrained bidders in an ad exchange. Our notion of equilibrium combines in a novel way two different approximations to address the limitations in PBE. First, we consider a *mean-field* approximation to relax the informational requirements of agents. The motivation behind the latter is that, when the number of players is large, there is little value in tracking the specific actions of all agents, and one may rely on some aggregate and stationary representation of the competitors' bids. The mean-field approximation assumes that, even when the overall number of advertisers in the market is large, only a small fraction participates in every auction, which closely reflects the existing competitive landscape in today's ad exchanges. This type of approximation has appeared in other auction and industrial organization applications (see, e.g., Adlakha et al. 2015, Iyer et al. 2014, Weintraub et al. 2008). Second, borrowing techniques from the revenue management literature (see, e.g., Gallego and van Ryzin 1994), we consider a stochastic *fluid* approximation to handle the complex dynamics of the advertisers' control problem. Such approximations are suitable when the number of opportunities is large and the payment per opportunity is small compared with the budget; hence, these are well motivated in the context of ad exchanges (see, e.g., Netmining 2011).

Using the two approximations above, we define the notion of a *fluid mean-field equilibrium* (FMFE).[2] We show that FMFE provides a good approximation to the rational behavior of agents as markets become large, yielding theoretical support for the use of FMFE as an equilibrium concept in this setting. Moreover, we show through a combination of theoretical and numerical results that the FMFE strategy is typically close to being a best response among a large class of strategies that keep track of all available information in the market, even in small markets with few advertisers (e.g., 5 to 10), providing further practical support to the concept. Specifically, in small markets a bidder may have incentives to overbid and deplete competitors' budgets to decrease competition in the future. We show, however, that the incentives to exercise such strategic behavior are low relative to playing FMFE even in small markets.

From a structural perspective, when a second-price auction is conducted, remarkably, the resulting FMFE strategy has a simple yet appealing form: an advertiser needs to shade her values by a *constant factor.* Furthermore, in equilibrium, advertisers will deplete their budget at an essentially constant rate, a typical practical requirement known as *smooth budget depletion*. Intuitively, when budgets are tight, advertisers shade their bids because there is an option value for future good opportunities. In addition, we show that an FMFE always exists and provide a set of sufficient conditions that guarantee its uniqueness. We also provide a characterization for FMFE that suggests a simple and efficient algorithm for its computation. Finally, we derive a closed-form characterization of FMFE in the case of homogeneous bidders. These succinct characterizations of equilibria are remarkably rare, and

[1] In practice, ad exchanges may be operated by third parties; for simplification, in this paper we assume that the publisher and the party running the exchange constitute a single entity.

[2] In §1.2 we compare and contrast FMFE with related notions of equilibria introduced in previous work.

one may significantly leverage them when studying the publisher's problem.

1.1.2. Auction Design. We show how a publisher that maximizes expected profits can use FMFE as a tool for backtesting different auction designs while accounting for the strategic response of budget-constrained advertisers. In particular, we focus on optimally setting the reserve price. When solving her optimization problem, the publisher trades off the revenues extracted from the auction with the opportunity cost of selling the impressions through an alternative channel. In addition, she needs to consider that changing the auction parameters may change the FMFE strategies played by advertisers. In particular, we show through a combination of theoretical and numerical results that ignoring budgets typically results in reserve prices that are lower than optimal and may result in significant profit losses for the publisher. We believe these results are particularly relevant because budgets are typically ignored in the literature when setting optimal reserve prices in ad exchanges, despite their prevalence in practice (see the related literature below). We further highlight that other levers may be optimized through the proposed framework, such as the allocation of impressions to the exchange or the extent of user information to disclose to the advertisers.

Overall, this paper is among the first in the literature (with the exception of Gummadi et al. 2011, which we discuss below) to provide a framework for profit optimization in repeated auctions, considering the strategic response of budget-constrained bidders. As such, we believe this work can have a practical impact on the design of ad exchange auctions. More broadly, we expect that FMFE may have additional applications beyond the one presented in this paper.

1.2. Related Work

This work contributes to various streams of literature. By accounting for advertisers' budget constraints and the resulting intertemporal dependencies and dynamic bidding strategies they induce, we contribute to the Internet advertising literature in particular and, more generally, to the literature on auction design in dynamic settings. To gain tractability, some papers have also used mean-field approximations in these settings. In this vein, Iyer et al. (2014) study repeated auctions in which bidders learn about their private value over time. Our mean-field approximation builds on theirs. However, in our setting dynamics are driven by budget constraints as opposed to learning, resulting in a different model. Moreover, in addition to the mean-field approximation, we impose a fluid approximation on the bidders' control problem. Relative to Iyer et al., this yields a more succinct characterization of equilibria based on shading factors that (1) brings computational advantages, (2) provides closed-form solutions in some settings even for the optimal auction decisions, and (3) allows using univalence theorems to provide broad sets of conditions under which FMFE is unique. Furthermore, for special cases, we provide approximation results under a sharper scaling, more in line with the typical scales observed in practice. In summary, the *combination* of the two approximations yields results that are extremely rare in the analysis of dynamic games, even after imposing (only) a mean-field approximation.

Closest to our paper is the study of Gummadi et al. (2011), who, in simultaneous and independent work, also study budget-constrained bidders in repeated auctions and introduce a similar equilibrium concept to FMFE. However, the studies differ along several important dimensions. Gummadi et al. study a more general class of online budgeting problems in an asymptotic regime in which the spending per interaction is small relative to the budget; a particular case of this general formulation is repeated second-price auctions with budget constraints. The present paper, in contrast, focuses on second-price auctions and provides the following sharper results for them that are not present in Gummadi et al. First, we rigorously justify FMFE as a solution concept through an asymptotic result for large markets and numerical results for finite markets, and we provide sufficient conditions for uniqueness of FMFE. Furthermore, we also study various auction design decisions of the publisher, providing important insights on, e.g., reserve price optimization.

More broadly, our work contributes to the growing literature on display advertising and, in particular, on that with ad exchanges. From the publisher's perspective, various studies analyze display ad allocation with both guaranteed contracts and spot markets; see, e.g., Alaei et al. (2009), Ghosh et al. (2009a), Yang et al. (2012), and Balseiro et al. (2014). These papers, however, take the actions of the advertisers as exogenous in the auction design. Chen (2011) employs a mechanism design approach to characterize the optimal dynamic auction for the publisher in the presence of guaranteed contract constraints. In this work, however, the publisher faces short-lived advertisers and budget constraints are also ignored. Vulcano et al. (2002) consider a related problem in the context of a single-leg revenue management problem. Celis et al. (2014) introduce a new randomized auction mechanism that experimentally performs better than an optimized second-price auction in markets that become thin as a result of targeting. They consider, however, a one-shot auction and do not take into account the dynamics introduced by budget constraints. Arnon and Mansour (2011) consider an abstraction of a repeated budgeted second-price auction in which the repeated interactions are collapsed into a single-shot auction with multiple identical copies of the same item and study the pure Nash equilibrium of this game. They do not provide, however, a rigorous justification of the solution concept. From the advertiser's perspective, Ghosh et al. (2009b) study the design of a bidding agent for a campaign in the presence of an exogenous market.

There is some body of literature on display advertising from a revenue management angle that focuses exclusively on guaranteed contracts (see, e.g., Araman and Fridgeirsdottir 2011, Najafi-Asadolahi and Fridgeirsdottir 2014, Roels and Fridgeirsdottir 2009, Turner 2012). In the related area of TV broadcasting, Araman and Popescu (2010) study the allocation of advertising space between forward contracts and the spot market when the planner faces supply uncertainty. From a methodological standpoint, our work also relates to a stream of work in revenue management. The single-agent fluid approximation we use and some of the intuition underlying it is related to that of, e.g., Gallego and van Ryzin (1994). Building on the latter and focusing on price competition, Gallego and Hu (2014) use a related notion of fluid, or open-loop, equilibrium. Other papers studying dynamic games in revenue management (all focusing on price competition) include Dudey (1992), Farias et al. (2011), and Martínez-de-Albéniz and Talluri (2011).

Our work is related to various streams of literature in auctions. First, previous work has studied auctions with financially constrained bidders in static one-shot settings (see, e.g., Che and Gale 1998, 2000; Laffont and Robert 1996; Maskin 2000; Pai and Vohra 2014). In §5 we show that in our dynamic model we obtain drastically different results to some of the main results in that stream. In addition, whereas our focus is on the impact of budget constraints on second-price auctions, our work is somewhat related to the recent literature in optimal dynamic mechanism design. (See Bergemann and Said 2010 for a survey.) Finally, our work relates to previous papers in repeated auctions, such as Jofre-Bonet and Pesendorfer (2003), in which, similarly to our model, bidders shade their bids to incorporate the option value of future auctions. However, in contrast with our work, the latter paper assumes Markov perfect equilibrium behavior in an empirical setting.

2. Model Description

We study a continuous-time infinite horizon setting in which users arrive to an online publisher's Web page according to a Poisson process $\{N(t)\}_{t\geq 0}$ with intensity η. We index the sequence of arriving users by $n \geq 1$, and we denote the sequence of arrival times by $\{t_n\}_{n\geq 1}$. When a user requests the Web page, the publisher may display one advertisement; this event is referred to as an *impression*. The publisher may decide to send the impression to an *ad exchange*, where an auction among potentially interested advertisers is run to decide which ad to show to the user. The exchange determines the winning bid via a second-price auction with a reserve price and returns a payment to the publisher. The rules of the auction and the characteristics of the users' arrival process are common knowledge.

2.1. Advertisers

Advertisers arrive to the exchange according to a Poisson process $\{K(t)\}_{t\geq 0}$ with intensity λ. We index the sequence of arriving advertisers by $k \geq 1$ and denote the arrival times by $\{\tau_k\}_{k\geq 1}$.[3]

Advertiser k is characterized by a type vector $\theta_k = (b_k, s_k, \alpha_k, \gamma_k) \in \mathbb{R}^4$. The first component of the type, b_k, denotes the budget, and the second component, s_k, denotes the campaign length. That is, the kth advertiser's campaign takes place over the time horizon $[\tau_k, \tau_k + s_k)$, and her total expenditure cannot exceed b_k. Once the advertiser leaves the exchange, she never comes back.

When the publisher contacts the exchange, she submits some partial information about the user visiting the website that, for example, could include *cookies*. This information, in turn, may heterogeneously affect the targeting criteria and the value an advertiser perceives for the impression, which are captured by α_k and γ_k, as we now explain. When the nth user arrives, the advertisers in the exchange observe the user information disclosed by the publisher and determine whether they will participate or not in the auction based on their targeting criteria. We assume that the kth advertiser matches a user with probability α_k independently and at random (across both impressions and advertisers). Conditional on a match, advertisers have independent private values for an impression. In particular, all values for advertiser k are independent and identically distributed random variables with a continuous cumulative distribution $F_v(\cdot\,; \gamma_k)$, parameterized by $\gamma_k \in \mathbb{R}$. The distributions have compact support $[\underline{V}, \overline{V}] \subset \mathbb{R}_+$ and continuously differentiable density.[4]

At the moment of arrival, an advertiser's type is drawn independently from a common knowledge distribution with support $\mathbf{\Theta}$, a finite subset of the strictly positive orthant $\mathbb{R}^4_{++}$. This distribution characterizes the heterogeneity among advertisers in the market.

[3] We note that our approach does not rely in the assumption of Poisson arrivals. In fact, our framework is general, and as shown in §6, it also applies, for example, to the case of synchronous campaigns, when all campaigns start and end at the same time (e.g., weekly or monthly campaigns). In reality, arrivals may lie in a middle ground with a combination of some campaigns repeating over time through a regular schedule, a random inflow of new advertisers (launching, for example, a one-off campaign), and exits of existing advertisers. Our framework could be applied to this setting as well.

[4] By assuming private values, we will ignore the effects of adverse selection and cherry-picking in common value auctions when some advertisers have superior information. See Levin and Milgrom (2010) and Abraham et al. (2013) for work that discusses and analyzes this setting.

Advertisers have a quasilinear utility function given by the difference between the sum of the valuations generated by the impressions won minus the expenditures corresponding to the second-price rule over all auctions in which they participate during the length of their campaign. The objective of each advertiser is to maximize her expected utility subject to her budget constraint.

2.2. Publisher

On the supply side, the publisher has an opportunity cost for selling her inventory of impressions in the exchange; that is, the publisher obtains some fixed amount $c > 0$ for each impression not won by some advertiser in the exchange. The publisher's payoff is given by the long-run average profit rate generated by the auctions, where the profit is measured as the difference between the payment from the auction and the lost opportunity cost c when the impression is won by an advertiser in the exchange. The publisher's objective is to maximize her payoff by adjusting the reserve price r to set for the auctions.

Notation. Given a random variable X, we denote a realization x with lowercase, its sample space $\mathbf{X}$ with bold capitals, the cumulative distribution function by $F_x(\cdot)$, and the law by $\mathbb{P}_x\{\cdot\}$.[5]

3. Equilibrium Concept

Given the auction design decisions of the publisher, the advertisers participate in a game of incomplete information. Moreover, because the budget constraints couple advertisers' decisions across periods, the game is dynamic and does not reduce to a sequence of static auctions.

A standard solution concept used for dynamic games of incomplete information is that of weak perfect Bayesian equilibrium (WPBE) (Mas-Colell et al. 1995). Roughly speaking, in such a game, a pure strategy for advertiser k is a mapping from histories to bids, where the histories represent past observations. A strategy specifies, given a history and assuming the advertiser participates in an auction at time t, an amount to bid. A strategy profile in conjunction with a belief system constitutes a WPBE if the following holds. First, given a belief system and the competitors' strategies, an advertiser's bidding strategy maximizes expected future payoffs. Second, beliefs must be consistent with the equilibrium strategies and Bayes' rule whenever possible.

WPBE and commonly used refinements, such as perfect Bayesian equilibrium and sequential equilibrium, require advertisers to hold beliefs about the entire future dynamics of the market, including the future market states. With more than a few competitors in the market, this imposes a very strong rationality assumption over advertisers because these belief distributions are high-dimensional. Moreover, to find a best response, advertisers need to solve a dynamic programming problem that optimizes over history-dependent strategies. This optimization problem can be highly dimensional and intractable both analytically and computationally. Hence, solving for WPBE for most markets of interest is not possible. More importantly, WPBE imposes informational requirements and a level of sophistication on the part of agents that seems unrealistic. This motivates the introduction of alternative equilibrium concepts. After some background in §3.1, we introduce such an alternative in §3.2.

3.1. Mean-Field and Fluid Approximation

When selecting an amount to bid, an advertiser needs to form some expectation of the distribution of bids she will compete against. There are various possible bases for such an expectation as a function of the sophistication of the advertiser and the type of information she would have access to. In practice, the number of advertisers in an exchange is often large, on the order of hundreds or even thousands. The first approximation we make is based on the premise that, given a large number of advertisers in the market, the distribution of competitors' bids is stationary and that these random quantities are uncorrelated among periods. Moreover, the bids of any particular advertiser do not affect this distribution. In these markets, it is common that auctioneers provide a "bid landscape" based on aggregated historical data that inherently assumes stationarity, at least for some significant time horizon. This information is commonly used by advertisers to set their bids, and therefore, our assumption about the distribution of competitors' bids may naturally arise in practice (Ghosh et al. 2009b, Iyer et al. 2014). In the present paper, although our approximation is predicated on the overall number of advertisers in the market being large, the average number of bidders *per auction* need not be large. For this reason, running auctions remains useful in this regime; a small number of bidders with heterogeneous valuations participate in each one of them.

To win an auction, an advertiser competes against all other bidders and against the reserve price r. We denote by D the steady-state maximum of the competitors' bids, where we assume that the publisher is a competitor that submits a bid equal to r. Assume for a moment that D is independent and identically distributed (i.i.d.) across different auctions and distributed according to a cumulative distribution function $F_d(\cdot)$. (Note that $F_d(\cdot)$ will be endogenously determined in equilibrium in §3.2.)

[5] For space considerations, only selected proofs are presented in the main appendix. All other proofs are presented in a supplementary appendix (available as supplemental material at http://dx.doi.org/10.1287/mnsc.2014.2022).

In this setting, the advertiser's dynamic bidding problem in the repeated auctions can be cast as a revenue-management-type stochastic dynamic programming problem in which bidding decisions across periods are coupled through the budget constraint. However, the resulting Hamilton-Jacobi-Bellman equation is a partial differential equation that, in general, does not have a closed-form solution. To get a better handle on the bidder's dynamic optimization problem, we introduce a second level of approximation motivated by the fact that a given advertiser has a large number of bidding opportunities. (Campaigns span for weeks or months, and thousands of impressions arrive per day.) In such an environment, the advertiser's stochastic dynamic programming problem can be well approximated through a stochastic fluid model. In particular, the approximation we focus on is predicated on the assumption that bidders solve a control problem in which the budget constraint need only be satisfied in expectation. Under the latter assumption, it is possible to show that one can restrict attention to stationary bidding strategies that ignore the individual state and are only dependent on the actual realization of the bidder's value without loss of optimality. We emphasize here that the budget constraint is imposed almost surely when we conduct performance analysis in §6. The main point is that the stationary bidding strategies derived above can be shown to provide advertisers with provably good policies in the real system (with constraints imposed almost surely) when the number of impressions and budgets are large, so the number of bidding opportunities over the campaign length also grows large.

Now, the control problem, for a bidder of type $\theta = (b, s, \alpha, \gamma)$, is one of finding a *fluid-based bidding strategy* $\beta^{\mathrm{F}}_\theta(v; F_d)$ that bids depending solely on her value v for the impression. A bidder with total campaign length s observes, in expectation, a total number of $\alpha\eta s$ impressions during her stay in the exchange. By conditioning on the impressions' arrival process, and using our assumption of the stationarity of the maximum bids and the valuations, the bidder's optimization problem is given by

$$J^{\mathrm{F}}_\theta(F_d) = \max_{w(\cdot)} \quad \alpha\eta s \mathbb{E}[\mathbf{1}\{D \le w(V)\}(V - D)] \tag{1a}$$

$$\text{s.t.} \quad \alpha\eta s \mathbb{E}[\mathbf{1}\{D \le w(V)\} D] \le b, \tag{1b}$$

where the expectation is taken over both the maximum bids D and the valuations V, which are independently distributed according to $F_d(\cdot)$ and $F_v(\cdot\,; \gamma)$, respectively. Note that the payments in the bidders' problem are consistent with a second-price rule. The bidder optimizes over a bidding strategy that maps its own valuation to a bid; hence, the resulting problem is an infinite-dimensional optimization problem. The next result provides, however, a succinct characterization of an optimal fluid-based bidding strategy.

Proposition 3.1. *Suppose that $\mathbb{E}[D] < \infty$. Let μ^*_θ be an optimal solution of the dual problem $\inf_{\mu \ge 0} \Psi_\theta(\mu; F_d)$, with $\Psi_\theta(\mu; F_d) = \alpha\eta s \mathbb{E}[V - (1+\mu)D]^+ + \mu b$. Then, an optimal bidding strategy that solves* (1) *for type θ is given by*

$$\beta^{\mathrm{F}}_\theta(v; F_d) = \frac{v}{1 + \mu^*_\theta}.$$

The optimal bidding strategy has a simple form: an advertiser of type θ needs to *shade* her values by the *constant factor* $1 + \mu^*_\theta$, and this factor guarantees that the advertiser's expenditure does not exceed the budget. In the previous expression, μ^*_θ is the optimal dual multiplier of the budget constraint and gives the marginal utility in the advertiser's campaign of one extra unit of budget. Intuitively, when budgets are tight, advertisers shade their bids because there is an option value for future good opportunities. When budgets are not tight, the optimal dual multiplier is equal to zero, and advertisers bid truthfully as in a static second-price auction. The proof of the result relies on an analysis of the dual of problem (1). Although the latter is not a convex program, the proof establishes from first principles that no duality gap exists in the present case.

3.2. Fluid Mean-Field Equilibrium

We now define the dynamics of the market as a prelude to introducing the equilibrium concept we focus on. At any point in time, there can be an arbitrary number of advertisers in the exchange, and these dynamics are governed by the patterns of arrivals and departures. In particular, the number of advertisers in the exchange behaves as an $M/G/\infty$ queue. We denote by $\mathcal{Q}(t)$ the set of indices of the advertisers in the exchange at time t and by $Q(t) = |\mathcal{Q}(t)|$ the total number of advertisers in the system. The *market state* at time t is given by the set of bidders in the exchange, together with their individual states and types, $\Omega(t) = \{\mathcal{Q}(t), \{b_k(t), s_k(t), \theta_k\}_{k \in \mathcal{Q}(t)}\}$, where we denote by $b_k(t)$ and $s_k(t)$ the kth advertiser remaining budget and residual time in the market by time t, respectively. When advertisers implement fluid-based strategies, the market state encodes all the information relevant to determine the evolution of the market, and the process $\Omega = \{\Omega(t)\}_{t \ge 0}$ is Markov.

In our equilibrium concept, we will require the consistency of the distribution of the maximum bid that bidders conjecture they compete against with the bidding strategies they use. A difficulty with this consistency check is that the number of *active* bidders, those that match the target criteria and have remaining budgets, depends on the market dynamics. In particular, the budget dynamics depend on who wins and how much the winner pays in each auction. Hence, in principle, characterizing the resulting steady-state distribution of the maximum bid of the active competitors (that have remaining budgets) is complex. However, it is reasonable to conjecture that, when the number of opportunities during the campaign length is large, rational advertisers would deplete their budgets close to the end of their campaign with high probability. For analytical tractability we impose that, in our equilibrium concept, any bidder currently in the exchange that matches the targeting criteria, without regard of her budget, gets to bid. Under this assumption, the number of bidders in an auction equals the number of advertisers matching the targeting criteria, denoted by $M(t)$, which is just an independent sampling from the process $Q(t)$.[6] In the proof of Theorem 6.1 and in the technical report by Balseiro et al. (2012), we show this layer of approximation is in fact asymptotically correct. Indeed, the performance analysis in §6 takes into account that, when advertisers implement the FMFE strategies, stochastic fluctuations in their expenditure may actually induce them to run out of their budgets before the end of the campaign, at which point they cannot continue to participate in any auction.

Because arrival and departures of advertisers are governed by an $M/G/\infty$ queue and campaign lengths are bounded, it is not hard to show that under fluid-based strategies, the market process Ω is Harris recurrent, so it is ergodic and admits a unique invariant steady-state distribution (see, e.g., Asmussen 2003, p. 203). Let $(\hat{M}, \{\hat{\Theta}_k\}_{k=1}^{\hat{M}})$ be a random vector that describes the number of matching bidders, together with their respective types when sampling a market state according to the invariant distribution. Notice that advertisers with longer campaign lengths and higher matching probability are more likely to participate in an auction, and thus the law of a type sampled from the invariant distribution does not coincide with the law of the types in the population. Indeed, by exploiting the fact that arrival-time and service-time pairs constitute a Poisson random measure on the plane (see, e.g., Eick et al. 1993), one can show that $\hat{M}$ is Poisson with parameter $\mathbb{E}[\alpha_\Theta \lambda s_\Theta]$ and that each component of the vector of types is independently and identically distributed as $\mathbb{P}_{\hat{\Theta}}\{\theta\} = (\alpha_\theta s_\theta / \mathbb{E}[\alpha_\Theta s_\Theta])\, \mathbb{P}_\Theta\{\theta\}$ for each type $\theta \in \boldsymbol{\Theta}$ and independent of $\hat{M}$.[7]

For a fluid-based strategy profile $\boldsymbol{\beta} = \{\beta_\theta(\cdot)\colon \theta \in \boldsymbol{\Theta}\}$ with $\beta_\theta\colon [\underline{V}, \overline{V}] \to \mathbb{R}$, we denote by $F_d(\boldsymbol{\beta})$ the distribution of the following random variable:

$$\max\left(\{\beta_{\hat{\Theta}_k}(V_{\hat{\Theta}_k})\}_{k=1}^{\hat{M}}, r\right), \tag{2}$$

which represents the steady-state maximum bid. Note that, here, V_θ are independent valuations sampled according to $F_v(\cdot\,; \gamma_\theta)$. We are now in a position to formally define the notion of an FMFE.

Definition 3.1 (Fluid Mean-Field Equilibrium). A fluid-based strategy profile $\boldsymbol{\beta}$ constitutes an FMFE if for every advertiser's type $\theta \in \Theta$, the bidding function β_θ is optimal for problem (1) given that the distribution of the maximum bid of other advertisers is given by $F_d(\boldsymbol{\beta})$ (Equation (2)).

In essence, an FMFE is a set of bidding strategies such that (i) these strategies induce a given competitive landscape as represented by the steady-state distribution of the maximum bid, and (ii) given this landscape, advertisers' optimal fluid-based bidding strategies coincide with the initial ones. We focus on symmetric equilibria in the sense that all bidders of a given type adopt the same strategy. Note that in the FMFE, upon arrival to the system, an advertiser is assumed to compete against the market steady-state maximum bid D.[8]

3.2.1. Remarks. We introduced the FMFE by heuristically arguing that it should be a sensible equilibrium concept for large markets when the number of bidding opportunities per advertiser are also large. In Theorem 6.1, we show that when all advertisers implement the FMFE strategy, the relative profit increase of any unilateral deviation to a strategy that keeps track of all information available to the advertiser becomes negligible as the scale of the market increases. This provides a theoretical justification for using the FMFE as an approximation of advertisers' behavior.

In the asymptotic regime described above, the matching probabilities are decreased so that the number of bidders per auction remains constant, and therefore, the probability that two advertisers participate repeatedly in the same auctions becomes negligible. In real-world markets, it might be the case that similar advertisers compete repeatedly in the same auctions to advertise to the same users. Nonetheless, in §6.2 we show through a combination of theoretical and numerical results that even with a moderate number of advertisers (e.g., 5 to 10) FMFE strategies are typically close to being a best response. Naturally, in these markets two advertisers may interact repeatedly over time, and our results show that the FMFE provides a good approximation to the rational behavior of agents even in these cases.

4. FMEE Characterization

In this section we prove the existence of an FMFE, provide conditions for uniqueness, and characterize the

[6] We note that an important difference between our FMFE and the related equilibrium concept proposed in parallel by Gummadi et al. (2011) is that they do not impose this additional layer of approximation. This plays a key role to obtain tractability in our analysis.

[7] For a type $\theta \in \boldsymbol{\Theta}$, we denote, with some abuse of notation, the corresponding budget by b_θ, the campaign length by s_θ, the matching probability by α_θ, and the valuation parameter by γ_θ. Additionally, we denote by Θ a random variable distributed according to the law of types in the population.

[8] Note that by the Poisson arrivals see time averages (PASTA) property of a Poisson arrival process this assumption is in fact correct.

FMFE. Proposition 3.1 will significantly simplify our analysis because it allows one to formulate the equilibrium conditions in terms of a vector of multipliers instead of bidding functions. By doing so, the problem of finding the equilibrium strategy function for a given type will be reduced to finding a single multiplier.

4.1. Equilibrium Existence and Sufficient Conditions for Uniqueness

We first prove the existence of an FMFE for a fixed reserve price. Recall from Proposition 3.1 that, in an optimal fluid bidding strategy, advertisers of type θ shade their bids using a fixed multiplier μ_θ. In the following, we denote by $\boldsymbol{\mu} = \{\mu_\theta\}_{\theta\in\boldsymbol{\Theta}}$ a vector of multipliers in $\mathbb{R}_+^{|\Theta|}$ for the different advertisers' types. Given a postulated profile of multipliers $\boldsymbol{\mu}$, let $F_d(\boldsymbol{\mu})$ denote the steady-state distribution of the maximum bid and let $\Psi_\theta(\mu;\boldsymbol{\mu}) \triangleq \Psi_\theta(\mu; F_d(\boldsymbol{\mu}))$ be the dual objective for one θ-type advertiser (as defined in Proposition 3.1) when all other bidders adopt a strategy given by the vector $\boldsymbol{\mu}$ (including those of the same type). In the dual formulation, a vector of multipliers $\boldsymbol{\mu}^*$ constitutes an FMFE if and only if it satisfies the best-response condition

$$\mu_\theta^* \in \arg\min_{\mu\geq 0} \Psi_\theta(\mu;\boldsymbol{\mu}^*), \quad \text{for all types } \theta\in\Theta. \tag{3}$$

One may establish that the system of equations (3) always admits a solution to obtain the following.

Theorem 4.1. *There always exists an FMFE.*

The proof shows that the dual strategy space can be reduced to a compact set and that the dual objective function is jointly continuous in its arguments and convex in the first argument. Then, a standard result that relies on Kakutani's fixed-point theorem implies the existence of an FMFE.

We now turn to sufficient conditions for uniqueness. Let $\mathbf{G}$: $\mathbb{R}_+^{|\boldsymbol{\Theta}|} \times \mathbb{R}_+ \to \mathbb{R}_+^{|\boldsymbol{\Theta}|}$ be a vector-valued function that maps a profile of multipliers and a reserve price to the steady-state expected expenditures per auction of each type. The expected expenditure of a θ-type bidder in a second-price auction when advertisers implement a profile of multipliers $\boldsymbol{\mu}$ is given by $G_\theta(\boldsymbol{\mu}, r) \triangleq \mathbb{E}[\mathbf{1}\{(1+\mu_\theta)D \leq V\}D]$, where the maximum competing bid is given by $D = \max((V_{\hat{\Theta}}/(1+\mu_{\hat{\Theta}}))_{1:\hat{M}}, r)$.[9] In the next assumption, we say that a matrix is a P-matrix if all its principal minors are positive (Horn and Johnson 1991, p. 120).

Assumption 4.1 (P-Matrix). *The Jacobian of $-\mathbf{G}$ with respect to $\boldsymbol{\mu}$ is a P-matrix for all $\boldsymbol{\mu}$ in $\mathbb{R}_+^{|\boldsymbol{\Theta}|}$.*

[9] Note that, consistent with the FMFE assumption and the PASTA property, the bidder competes against the *market* steady-state maximum bid.

Assumption 4.1 can be shown to hold for various cases of interest. For example, it is easy to see that it always holds for the case of homogeneous advertisers, i.e., when the space of types $\boldsymbol{\Theta}$ is a singleton. In §2 of the supplementary appendix, we provide an important class of settings with two types of bidders in which it also holds. The P-matrix condition can be interpreted as a monotonicity condition on the expected expenditures. Namely, if a group of types increases its multipliers simultaneously, then the expenditures cannot increase for every type in the group. The next theorem shows that the equilibrium is unique under the P-matrix assumption.

Theorem 4.2. *Suppose Assumption* 4.1 *holds. Then, there is a unique FMFE of the form $\beta_\theta(v) = v/(1+\mu_\theta)$, θ in Θ.*

We prove the result by formulating the FMFE conditions as a nonlinear complementarity problem (NCP), as presented in Corollary 4.1 below, and employing a univalence theorem to show that the expected expenditure mapping is injective (Facchinei and Pang 2003a). We note that results regarding the uniqueness of equilibria in dynamic games are extremely rare (Doraszelski and Pakes 2007).

Providing conditions for which Assumption 4.1 holds is challenging for more than two types of bidders. In our numerical experiments, we use a myopic best-response algorithm, presented in detail in §5.2.1, that could naturally describe how agents learn to play the game and reach an FMFE. It is reassuring that in our computational experience, for a given model instance with two or more types, this algorithm always found the same FMFE even when starting from different initial points.

We finish this subsection by noting that, under further mild regularity conditions, one can establish that any set of continuous increasing bidding functions that constitute an FMFE necessarily yields the same outcome (in terms of auctions' allocations and payments) as that of the FMFE in Theorem 4.2. In the rest of the paper, we focus on the simple and intuitive FMFE strategies that can be described by a vector of dual multipliers.

4.2. Equilibrium Characterization

A direct corollary of the earlier results and their proofs yields the following succinct characterization.

Corollary 4.1. *Any FMFE characterized by a vector of multipliers $\boldsymbol{\mu}^*$, such that $\beta_\theta(v) = v/(1+\mu_\theta^*)$ for all $v\in[\underline{V}, \overline{V}]$ and $\theta\in\Theta$, solves*

$$\mu_\theta^* \geq 0 \perp \alpha_\theta \eta s_\theta G_\theta(\boldsymbol{\mu}^*, r) \leq b_\theta \quad \forall\,\theta\in\Theta,$$

where $\perp$ indicates a complementarity condition between the multiplier and the expenditure—that is, at least one condition should be met with equality.

The expected expenditure for a bidder of type θ over its campaign when bidders use a vector of multipliers $\boldsymbol{\mu}$ is given by $\alpha_\theta \eta s_\theta G_\theta(\boldsymbol{\mu}, r)$ because, on average, she faces ηs_θ auctions and participates in a fraction α_θ of them. Intuitively, the result states that, in equilibrium, advertisers of a given type may only shade their bids if their total expenditure over the course of the campaign (in expectation) is equal to their budget. If, in expectation, advertisers have excess budget at the end of a campaign, then their multiplier is equal to zero and they should bid truthfully. This equilibrium characterization lends itself to tractable algorithms to compute FMFE because the strategy of each advertiser type is determined by *a single number* that satisfies the complementary conditions above. See, for example, Chapter 9 of Facchinei and Pang (2003b) for a discussion of numerical algorithms for this kind of NCP.

We conclude this subsection by refining the result for the case of homogeneous bidders, in which one can provide a quasi-closed form characterization for FMFE. Suppose that Θ is a singleton. Let $G_0(r) = G_\theta(0, r)$ denote the steady-state unconstrained expected expenditure per auction of a single bidder for a second-price auction with reserve price r when all advertisers (including the bidder herself) bid their own values. Note that the expected expenditure for a bidder over its campaign when all bidders are truthful is given by $\alpha\eta s G_0(r)$. This quantity plays a key role in the FMFE characterization.

Proposition 4.1. *Suppose $\boldsymbol{\Theta}$ is a singleton. Then an FMEE exists and is unique. In addition, the equilibrium may be characterized as follows: $\beta_\theta(v) = v/(1+\mu^*)$ for all $v \in [\underline{V}, \overline{V}]$, where $\mu^* = 0$ if $\alpha\eta s G_0(r) < b$, and μ^* is the unique solution to $\alpha\eta s G_0(r(1+\mu)) = b(1+\mu)$ if $\alpha\eta s G_0(r) \ge b$.*

The result provides a complete characterization of the unique FMFE. In particular, it states that if budgets are large (i.e., $\alpha\eta s G_0(r) < b$), then, in equilibrium, advertisers will bid truthfully. However, if budgets are tight (i.e., $\alpha\eta s G_0(r) \ge b$), then advertisers will shade their bids, in equilibrium, considering the option value of future opportunities. We also further note here that in the case in which the reserve price is equal to zero ($r = 0$), the equilibrium multiplier may be characterized in closed form by $\mu^* = (\alpha\eta s G_0(0)/b - 1)^+$.

5. Auction Design: Reserve Price Optimization

In this section, we study the publisher's profit maximization problem. First, we use the framework developed in the previous sections to formulate the problem. Then, we study the resulting optimization problem and derive insights into how to account for budgets when setting the reserve price.

We model the grand game played between the publisher and advertisers as a Stackelberg game in which the publisher is the leader and the advertisers are the followers. In particular, the publisher first selects the reserve price in the second-price auction r, and then the advertisers react and play the induced dynamic game among them. In our analysis, we assume that the solution concept for the game played between advertisers is that of an FMFE. The publisher's objective is to maximize her long-run average profit rate from the auctions while considering the opportunity cost c of the alternative channel.

To mathematically formulate the problem, we define $I(\boldsymbol{\mu}, r) = 1 - F_d(r; \boldsymbol{\mu})$ as the probability that the impression is won by some advertiser in the exchange when advertisers shade according to the profile $\boldsymbol{\mu}$ and the publisher sets a reserve price r. Using the characterization of an FMFE in Corollary 4.1, we can write the publisher's problem in terms of multipliers and obtain the following mathematical program with equilibrium constraints (MPEC):

$$\begin{aligned} \max_r \quad & \lambda \sum_{\theta \in \boldsymbol{\Theta}} \{ p_\theta \alpha_\theta \eta s_\theta G_\theta(\boldsymbol{\mu}, r) \} - \eta c I(\boldsymbol{\mu}, r) \\ \text{s.t.} \quad & \mu_\theta \ge 0 \perp \alpha_\theta \eta s_\theta G_\theta(\boldsymbol{\mu}, r) \le b_\theta \quad \forall\, \theta \in \Theta, \end{aligned} \tag{4}$$

where $p_\theta \triangleq \mathbb{P}_\Theta\{\theta\}$ is the probability that an arriving advertiser is of type θ. We denote by $\Pi(\boldsymbol{\mu}, r)$ the objective function of the MPEC. The first term in the objective is the publisher's revenue rate obtained from all bidders' types in the auctions, which is equal to the average expenditure of the advertisers. Note that the revenue rate obtained from a given type is equal to the bidders' average expenditure over their campaign times the arrival rate of the bidders. The second term is the opportunity cost by unit of time, which is incurred whenever an impression is won by some advertiser in the exchange and, therefore, cannot be sold in the alternative channel.

Note that the MPEC above considers that, when the publisher changes the reserve price, bidders react by playing a corresponding FMFE. By Theorem 4.1, we know that such an FMFE always exist. Furthermore, when Assumption 4.1 holds, the FMFE is unique. In cases for which we do not know whether the assumption holds, we will assume that advertisers play the FMFE computed by our best-response algorithm.[10]

[10] Assuming that the equilibrium being played is the one selected by a specific algorithm is a prevalent approach in the analysis of dynamic games for which uniqueness results are extremely rare. For example, Iyer et al. (2014) use this approach in a repeated auction setting, and many of the references in Doraszelski and Pakes (2007) use it in other industrial organization games.

5.1. Reserve Price: Homogeneous Advertisers

We first consider the case in which $\boldsymbol{\Theta}$ is a singleton; i.e., all advertisers have a fixed budget b, stay in the market for a deterministic time s, and share the same matching probability α and valuation parameter γ. By Proposition 4.1, we know that in this case, a unique FMFE exists, and we can characterize it in quasi-closed form. We leverage this result to study the publisher's decisions. Throughout this section, we drop the dependence on θ. In the following, we denote by $h_v(x) = f_v(x)/\bar{F}_v(x)$ the failure rate of the advertisers' values (who have a common distribution) and by $\xi_v(x) = xh_v(x)$ the generalized failure rate of the values. We assume that values possess strictly increasing generalized failure rates (IGFRs). This assumption is common in the pricing and auction theory literature, and many distributions satisfy this condition (see, e.g., Lariviere 2006, Myerson 1981).[11]

In the absence of budgets, the auctions are not coupled, and each auction is equivalent to a one-shot second-price auction with opportunity cost $c > 0$ and symmetric bidders with private values. In such a setting, it is well known that the optimal reserve price, which we denote by r_c^*, is independent of the number of bidders and is given by the unique solution of $1/h_v(x) = x - c$ (see, e.g., Laffont and Maskin 1980). The next result establishes a counterpart for the present case with budget constraints.

Theorem 5.1 (Optimal Reserve Price). *If $\alpha\eta s G_0(r_c^*) < b$, then r_c^* is the unique optimal reserve price. If $\alpha\eta s G_0(r_c^*) \geq b$, then the unique optimal reserve price is $\bar{r} = \sup \mathcal{R}^*$, where $\mathcal{R}^* = \{r\colon \alpha\eta s G_0(r) \geq b\}$. Furthermore, in the FMFE induced by the optimal reserve price, advertisers bid truthfully.*

The optimal reserve price admits a closed-form expression that highlights how it balances various effects. The expected expenditure for a bidder over its campaign when all bidders are truthful evaluated at r_c^*, $\alpha\eta s G_0(r_c^*)$, plays a key role in the result. In fact, when the budget is large in the sense that advertisers do not deplete their budget in expectation when the reserve price is r_c^* ($\alpha\eta s G_0(r_c^*) \leq b$), it is expected that r_c^* should still be optimal in our setting. Intuitively, if the budget does not bind, the auctions decouple into independent second-price auctions. However, when $\alpha\eta s G_0(r_c^*) > b$, advertisers shade their values when the reserve price is r_c^*. In the proof, we show that in this case, the optimal reserve price must be in $\mathcal{R}^*$—that is, it must induce bidders to deplete their budgets in expectations. For all such reserve prices, the revenue rate for the publisher is given by λb, and this is the maximum revenue rate she can extract from advertisers. Hence, recalling the objective value (4) of the publisher, the optimal reserve price must be the value $r \in \mathcal{R}^*$ that minimizes the probability of selling an impression in the exchange and, therefore, the opportunity cost. Increasing the reserve price has two effects on this probability: (1) a *direct effect*, where, assuming the advertiser's strategies do not change, an increase of the reserve price decreases the probability of selling an impression in the exchange; and (2) an *indirect effect*, where a change in the reserve price also alters the strategies of the advertisers through the induced FMFE. In the proof, we show that the direct effect is dominant, implying that $\bar{r} = \sup \mathcal{R}^*$ is optimal because it minimizes the opportunity cost within $\mathcal{R}^*$.

We emphasize that the optimal reserve price with budget constraints is larger than or equal to r_c^*, the static reserve price that does not account for budgets. In fact, the optimal reserve price is $\max\{\bar{r}, r_c^*\}$ because one can show that $\bar{r} \geq r_c^*$ if and only if $\alpha\eta s G_0(r_c^*) \geq b$. Theorem 5.1 highlights that ignoring budgets can result in a suboptimal decision. In the next section, we numerically evaluate the extent of the suboptimality in markets with heterogeneous bidders. Before, we note that when advertisers are highly budget constrained, the reserve price $\bar{r}$ tends to be high, and therefore it is unlikely that two advertisers will bid higher than $\bar{r}$. In this case, the advantage of running a second-price auction becomes limited, and its performance is similar to that of a fixed posted-price mechanism.

We finish this subsection by comparing the result above with the studies pertaining to one-shot auctions with budget constraints. In the case of a common budget for all bidders, authors have typically found that budget constraints decrease the optimal reserve price relative to the setting without budget constraints (see Laffont and Robert 1996, Maskin 2000). The reason is that with budget constraints, the reserve price is less effective in extracting rents of higher valuation types; hence, when trading off higher revenues conditional on a sale taking place with an increase in the probability of a sale, the latter has more weight than in the absence of budgets. In our case, instead, the optimal reserve price with budget constraints is larger than or equal to r_c^*. The difference with the one-shot auction is that the budget constraint is imposed over a large set of auctions as opposed to having a constraint per auction, leading to a different trade-off for the publisher. Indeed, when the budget constraint binds, the reserve price does not affect expected revenues, and the publisher is already extracting all budgets from the bidders. Therefore, the only role of the reserve price becomes one of reducing the opportunity cost by decreasing the probability of a sale. As we saw, this is achieved by increasing the reserve price while still extracting the maximum amount of revenue.

[11] For instance, the uniform, exponential, triangular, truncated normal, gamma, Weibull, and log-normal distribution have IGFRs.

5.2. Reserve Price: Heterogeneous Advertisers

Although it was possible to obtain essentially a closed-form solution for the publisher's optimal reserve price in the case of homogeneous advertisers, it is not generally possible to derive such a result for the case of heterogeneous advertisers. However, one may always numerically analyze the impact of the publisher's decisions on the advertisers' equilibrium outcome under different scenarios by solving for the FMFE using the characterization in Corollary 4.1 for different auction parameters. We provide such a study in this section and start by describing an algorithm to compute the FMFE.

5.2.1. Algorithm to Compute the FMFE. For each model instance, we solve for the FMFE using the following myopic best-response algorithm over the space of dual multipliers. The algorithm starts from an arbitrary vector of multipliers $\boldsymbol{\mu}$.

Algorithm 1 (Best-Response Algorithm for an FMFE)

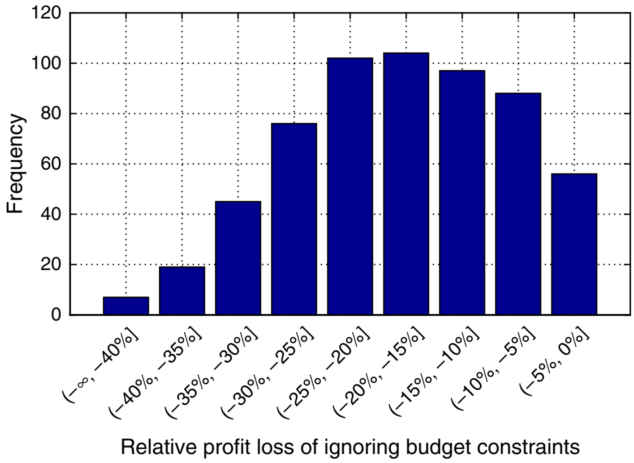
1. $\mu_\theta^0 := \mu_\theta, \ \forall\, \theta \in \Theta; \ \ i := 0$
2. **repeat**
3. $\quad \mu_\theta^{i+1} := \arg\min_{\mu' \geq 0} \Psi_\theta(\mu'; \boldsymbol{\mu}^i), \ \forall\, \theta \in \Theta$
4. $\quad \Delta := \|\boldsymbol{\mu}^{i+1} - \boldsymbol{\mu}^i\|_\infty; \ \ i := i+1$
5. **until** $\Delta < \epsilon$.

If the termination condition is satisfied with $\epsilon = 0$, we have an FMFE (see Equation (3)). Small values of ϵ allow for small errors associated with limitations of numerical precision. Although we cannot prove the convergence of the algorithm, in practice, it converged in a small number of iterations. In fact, for fixed auction parameters, solving for the FMFE takes a few seconds on a laptop computer.

5.2.2. Measuring the Impact of Budgets on the Optimal Reserve Price. The analysis with homogeneous bidders highlighted that ignoring budgets can lead to suboptimal reserve prices. In this section, we measure the extent of the suboptimality in markets with heterogeneous bidders. We believe this exercise is particularly relevant because several papers that study online advertising ignore budgets when setting optimal reserve prices in the ad exchange (see, e.g., Balseiro et al. 2014, Celis et al. 2014, Chen 2011, Ghosh et al. 2009a).

The setup for our numerical experiments is as follows. We consider randomly generated instances with a heterogeneous population of advertisers with five types. Budgets for each type are sampled from a discrete uniform distribution with support $\{1, 2, \ldots, 10\}$. Additionally, we experiment with the proportion of these types by choosing the probabilities p_θ of an arriving advertiser being of type θ uniformly from the probability simplex. Throughout the experiments, we fix the matching probability $\alpha = 0.1$ and the campaign length to $s = 10$, but we select the arrival rate λ uniformly in

Figure 1 (Color online) Histogram of the Relative Profit Loss of Ignoring Budget Constraints for Randomly Generated Instances



Notes. The relative profit loss is given by $\Pi(\boldsymbol{\mu}(r_c^*), r_c^*)/\Pi(\boldsymbol{\mu}(r^*), r^*) - 1$, where $\boldsymbol{\mu}(r)$ denotes the FMFE multipliers at reserve price r. The histogram is restricted to those instances in which the mean advertiser' truthful expenditure at r_c^* exceeds the mean budget; i.e., $\sum_\theta p_\theta G_\theta(0, r_c^*) \geq \sum_\theta p_\theta b_\theta$.

$[1, 5]$ so that the average number of matching bidders in an auction $\alpha\lambda s$ varies from one to five. Advertisers have the same distribution of values, which is drawn uniformly from the set $\text{Exp}(\gamma)$, $\mathcal{N}(\gamma, 1)$, and $\text{Unif}[0, 2\gamma]$, with γ uniformly sampled from $[1, 5]$ (the supports of valuations are truncated to $[0, 10]$). From the perspective of the publisher, we study scenarios with different opportunity costs c for the alternative channel by choosing the cost uniformly from $[1, 5]$. Additionally, we consider 10 levels for the impressions allocated to the exchange, as given by η.[12] In total, we examine 920 different scenarios.

For each model instance, we compute two reserve prices. First, we compute the optimal static reserve price r_c^* as given in §5.1, which assumes advertisers always bid truthfully and therefore ignores budget constraints. Second, the reserve price r^* that solves optimization problem (4), and therefore considers the rational response of budget-constrained advertisers via an FMFE, is computed.

From the numerical experiments, we obtain two conclusions that are robust across all model instances. First, consistent with the results from the homogeneous case, the reserve price r^* is larger than r_c^*. Second, the extent of suboptimality associated with ignoring budgets and selecting r_c^* instead of r^* can be significant with profit losses up to 40%. A histogram of the relative profit loss across the generated instances is shown in Figure 1. Overall, our results show that ignoring the

[12] In particular, we consider 10 uniformly spaced points in the interval $[0, 1.25 \max_\theta \bar{\eta}_\theta]$, where $\bar{\eta}_\theta$ is the least rate of impressions guaranteeing that a population of type θ bidders in isolation is budget constrained when the reserve is r_c^*.

rational response of budget-constrained advertisers can yield significant profit losses for the publisher.

5.2.3. Structure of the Optimal Reserve Price. In this section, we study in more detail the structure of the optimal reserve price r^* in markets with heterogeneous bidders to illustrate the trade-offs the publisher faces in these settings. For this purpose, it is useful to depict the optimal reserve price and the resulting shading multipliers as a function of the allocation of impressions to the exchange η. Figure 2 shows such dependence for a given set of parameters with two types. Notice that, when the publisher prices optimally, the high-budget type always bids truthfully. However, in contrast with the homogeneous case, this is not necessarily true for the low-budget type: for some levels of supply, low-type advertisers will shade their bids under the optimal reserve price.

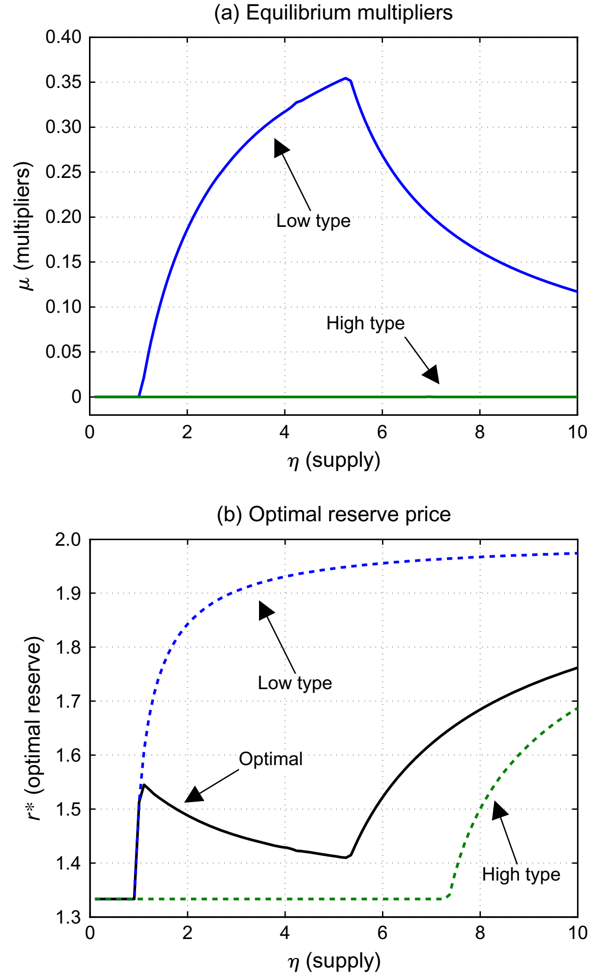
Focusing on the optimal reserve price, we observe that advertisers do not have a chance to deplete their budgets for low levels of supply. In this case, advertisers bid truthfully and r_c^* is the optimal reserve price. As the rate of impressions increases, the expenditures increase up to the point at which the low-type becomes budget constrained. From then on, the publisher needs to balance two effects. On the one hand, because the low type is now shading her bids, the publisher has an incentive to increase the reserve price so as to minimize the number of impressions won and the opportunity cost. The latter is achieved by $\bar{r}_1(\eta)$, the optimal reserve price if all advertisers shared the same budget b_1 (the top dashed line). On the other hand, the publisher has an incentive to price close to r_c^* to extract the surplus from the high-type advertisers, who are not depleting their budgets. The trade-off is such that, initially, the weight of the low-budget type bidders is higher and it is optimal for the publisher to price close to $\bar{r}_1(\eta)$, thus increasing the reserve price with the allocation of impressions. At this price, however, the expenditure of the high-budget type is well below its budget, and the publisher may be leaving money on the table. When enough impressions are allocated to the exchange, this effect becomes dominant and the publisher tries to extract this surplus by pricing closer to r_c^*; thus the sudden kink and decrease in the optimal reserve price. If the publisher keeps increasing the allocation of impressions, eventually both types become budget constrained. Similar to the homogeneous case, the publisher is now better off pricing in such a way that both types deplete their budgets, but with the high-type bidding truthfully so that the number of impressions won by the advertisers is minimized. For this reason, at some point the optimal reserve price starts increasing again.

In our numerical experiments, a similar structure and trade-off appears when there are more than two types of advertisers with different budgets in the population, with one new kink in the optimal reserve price for each additional type.

6. FMFE as a Near-Optimal Best Response

In this section, we aim to provide further support for the concept of FMFE introduced in §3.2 along two dimensions. First, we rigorously justify that playing an FMFE strategy when all other advertisers play the FMFE strategy is a near-optimal best response in large-sized markets, i.e., when both the number of advertisers and the number of auctions are appropriately large. Second, we aim to illustrate theoretically

Figure 2 **(Color online) Equilibrium Multipliers and Optimal Reserve Price as a Function of the Rate of Impressions for an Instance with** $\alpha = 0.1$, $\lambda = 1$, $s = 40$, Unif[0, 2] **Valuation Distribution,** $c = \frac{2}{3}$, $b = (1, 8)$, **and** $p = (\frac{1}{5}, \frac{4}{5})$



Notes. For illustration purposes, we only consider two types and different parameters than above. In panel (a), equilibrium multipliers serve as a function of the allocation of impressions. In panel (b), the solid line corresponds to optimal reserve price, and the dashed lines denote the optimal prices one would set for a homogeneous population with budget b_1 (low type) or b_2 (high type). The reserve price r_c^* is equal to $\frac{4}{3}$.

and numerically the main trade-offs faced by advertisers and why FMFE strategies are potentially near optimal even when the number of advertisers is small, lending further practical support to the concept.

Preliminaries. To achieve the above goals, we focus on a simplified version of the problem, the case of *synchronous campaigns*—that is, when all campaigns start at the same time and finish simultaneously. This simpler model corresponds, for example, to the case when advertisers have periodic (daily or weekly) budgets. It captures some of the key characteristics of the market and allows us to highlight the main issues at play in a relatively transparent fashion. The general case of *asynchronous campaigns* introduces a significant additional layer of complexity, and we provide an asymptotic approximation result pertaining to the latter in Balseiro et al. (2012).[13]

We next describe the synchronous model and adapt the FMFE to this setting. There is a fixed number of agents in the market, which we denote by K. All campaigns start at time 0 and finish at a common time s, and neither arrivals nor departures are allowed during the time horizon $[0, s]$. Agents are indexed by $k = 1, \dots, K$. Similar to before, the kth agent is characterized by a type vector, $\theta_k = (b_k, \alpha_k, \gamma_k) \in \mathbb{R}^3$. Types are publicly known and revealed at the beginning of the horizon. Although this assumption is not necessary for our analysis, we make it to simplify some arguments and notation.

Now, the expected expenditure function of the kth advertiser of a single auction when advertisers shade their bids according to a vector of multipliers $\boldsymbol{\mu} \in \mathbb{R}_+^K$, denoted by $G_k(\boldsymbol{\mu}; r)$, is given as in §4 but with the maximum competing bid given by $D_{-k} = \max_{i \neq k, M_i = 1}\{V_i/(1+\mu_i)\} \vee r$, where we let $M_k = 1$ indicate that the kth agent participates in the auction and we ignored the index n to simplify the notation. A similar analysis to the one performed in the case of asynchronous campaigns yields that the vector of multipliers in the FMFE can be characterized as the solution of the following NCP:

$$\mu_k \geq 0 \perp \alpha_k \eta s G_k(\boldsymbol{\mu}; r) \leq b_k, \quad \forall k = 1, \dots, K. \tag{5}$$

Moreover, similar results about the existence and uniqueness of the FMFE also apply to this setting.

[13] Because of the asynchronous nature of the market, for this result we extend the propagation of chaos argument of Graham and Méléard (1994) and Iyer et al. (2014) to accommodate the additional fluid approximation and the queuing dynamics of the number of advertisers in the market, which leads to a more restrictive scaling than our result below for synchronous campaigns. An interesting technical avenue for future research is to show whether the scaling under which we obtain our asymptotic approximation result for synchronous campaigns holds in broader settings. This generalization is likely to have other applications in mean-field models beyond the one presented in this paper.

6.1. Asymptotic Analysis for Large Markets

We consider a sequence of markets indexed by the number of advertisers K. For each market size K, bidders' types are given by $\{\theta_k^{(K)} = (b_k^{(K)}, \alpha_k^{(K)}, \gamma_k^{(K)})\}_{k=1}^K$, where we use superscript (K) to denote quantities associated to market size K. Similarly, we denote $\eta^{(K)}$ as the intensity of the arrival process of users in market K. We will prove an approximation result by considering a sequence of markets that satisfy the following set of assumptions on the primitives.

Assumption 6.1. *There exists positive bounded constants $\underline{g}$, $\bar{g}$, $\underline{z}$, and $\bar{a}$, such that for all market sizes K,*

(i) For any advertiser k, $b_k^{(K)}/(\alpha_k^{(K)}\eta^{(K)}s) \in [\underline{g}, \bar{g}]$.

(ii) For every pair of advertisers $k \neq i$, $\alpha_k^{(K)}/\alpha_i^{(K)} \leq \bar{a}$.

(iii) For any advertiser k, $G_k^{(K)}(\mathbf{0}; r) \geq \underline{z}$.

The first assumption states that the ratio of budget to number of matching auctions is uniformly bounded from above and below across advertisers, and the second one states that the ratio of matching probabilities of any two advertisers is uniformly bounded across advertisers. These assumptions guarantee that no advertiser has an excessive market influence by limiting budgets and the number of matching auctions in which they participate. The third assumption ensures that, in equilibrium, all advertisers have a positive expected expenditure per auction so that no advertiser is systematically outbid in equilibrium. Thus, these assumptions simply guarantee that, for every market along the sequence considered, there is no dominant or irrelevant advertiser. These assumptions do not impose further heterogeneity restrictions across advertisers.

We denote the kth advertiser *history* up to time t by $h_k(t)$. The history encapsulates all available information up to time t, including the advertisers' types, the realizations of her values up to that time, her bids, the budgets of all advertisers, and the result of every past auction. We define a pure strategy β for advertiser k as a mapping from histories to bids, and we denote by $\mathbb{B}^{(K)}$ the space of strategies that are nonanticipating and adaptive to the history in market K. We study the expected payoff of advertiser k when she implements a strategy $\beta^{(K)} \in \mathbb{B}^{(K)}$, and all other advertisers follow FMFE strategies $\boldsymbol{\beta}^{\mathrm{F},(K)}$ for market size K. The latter amounts to shading bids according to the multipliers that solve the NCP (5) while bidders have remaining budgets. This expected payoff is denoted by $J_k^{(K)}(\beta^{(K)}, \boldsymbol{\beta}_{-k}^{\mathrm{F},(K)})$, where the expectation is taken over the actual market process. In this notation, $J_k^{(K)}(\beta_k^{\mathrm{F},(K)}, \boldsymbol{\beta}_{-k}^{\mathrm{F},(K)})$ measures the *actual* expected payoff of the FMFE strategy for the advertiser in the exchange, which takes into account that advertisers may run out of budget before the end of the horizon. It is obvious that $J_k^{(K)}(\beta_k^{\mathrm{F},(K)}, \boldsymbol{\beta}_{-k}^{\mathrm{F},(K)}) \leq \sup_{\beta \in \mathbb{B}^{(K)}} J_k^{(K)}(\beta, \boldsymbol{\beta}_{-k}^{\mathrm{F},(K)})$. We will analyze the gap $\sup_{\beta \in \mathbb{B}^{(K)}} J_k(\beta, \boldsymbol{\beta}_{-k}^{\mathrm{F},(K)}) - J_k^{(K)}(\beta_k^{\mathrm{F},(K)}, \boldsymbol{\beta}_{-k}^{\mathrm{F},(K)})$

to bound the suboptimality of FMFE relative to unilaterally deviating to a best-response strategy. In what follows, $O(\cdot)$ stands for Landau's big O notation as K goes to infinity.

Theorem 6.1. *Suppose that Assumption* 6.1 *holds. Consider a sequence of markets indexed by K in which all bidders, except the kth bidder, follow FMFE strategies* $\boldsymbol{\beta}^{\mathrm{F},(K)}$ *in market K. Suppose that the kth advertiser unilaterally deviates and implements a nonanticipating and adaptive strategy* $\beta^{(K)} \in \mathbb{B}^{(K)}$ *in market K. The expected payoff of these deviations compared with the FMFE strategy satisfies*

$$\frac{1}{\alpha_k^{(K)}\eta^{(K)}s}\big(J_k^{(K)}(\beta^{(K)},\boldsymbol{\beta}_{-k}^{\mathrm{F},(K)}) - J_k^{(K)}(\beta_k^{\mathrm{F},(K)},\boldsymbol{\beta}_{-k}^{\mathrm{F},(K)})\big) = O\big(\alpha_k^{(K)} + (\alpha_k^{(K)}\eta^{(K)}s)^{-1/2}K^{1/2}\big).$$

The bound in Theorem 6.1 states that

$$1 - \frac{J_k^{(K)}(\beta_k^{\mathrm{F},(K)},\boldsymbol{\beta}_{-k}^{\mathrm{F},(K)})}{\sup_{\beta\in\mathbb{B}^{(K)}} J_k^{(K)}(\beta,\boldsymbol{\beta}_{-k}^{\mathrm{F},(K)})}$$

converges to zero as K grows to infinity when (1) the matching probabilities $\alpha_k^{(K)}$ converge to zero and (2) $K = o(\alpha_k^{(K)}\eta^{(K)}s)$—that is, the expected number of auctions a bidder participates in grows at a faster rate than the number of advertisers. In addition, the assumption imposes that the expected number of auctions a bidder participates in and the budget $b_k^{(K)}$ grow at the same rate. Typically, the scaling will also impose that the expected number of advertisers per auction remains constant (even though the overall number of advertisers grows large). These conditions naturally represent many ad exchange markets in which the number of auctions a bidder participates in is typically much larger than the number of competitors, the expected expenditure per auction is typically small compared with the budget, and the number of competitors per auction is small.

The key idea of the proof of Theorem 6.1 is to bound, in some appropriate way, the impact that the kth advertiser may have on the competitors and, based on that, bound the value that may be obtained by deviating from the FMFE strategy. To do so, we first exploit the fundamental observation that, independently of the kth advertiser's strategy, the competing advertisers bid exactly as prescribed by the FMFE while they have budgets remaining. Second, we exploit the fact that not all advertisers match the same impressions, and as a result, the impact of a single advertiser on any other specific advertiser (in terms of running out of budget) is limited. In particular, we establish that all advertisers will run out of budget close to the end of their campaigns no matter which strategy the deviant advertiser implements. Hence, the competitive landscape coincides with that predicted by the FMFE for most of the campaign. Based on this, we bound the performance of an arbitrary strategy by that of a strategy with the benefit of hindsight (which has complete knowledge of the future realizations of bids and values). This yields the result.

Finally, it is worthwhile to put this result in perspective with regard to typical revenue management heuristic fluid-based prescriptions. In most such settings, the bounds obtained (see, e.g., Talluri and van Ryzin 1998) are on the order of $n^{-1/2}$, where n is a proxy for the number of opportunities (akin to the number of auctions one participates in our setting). In the present context, this term is present as $(\alpha_k^{(K)}\eta^{(K)}s)^{-1/2}$, but it is multiplied by $K^{1/2}$ to control for the fact that there are K bidders that could potentially run out of budget before the length of the campaign. Moreover, the term $\alpha_k^{(K)}$ in the bound controls for the potential impact bidder k may have on any competitor, which is bounded by the expected fraction of auctions in which they compete together.

6.2. Analysis for Small Markets

Recall that the FMFE concept involves two approximations: (1) a fluid one motivated by the fact that advertisers participate in a large number of auctions during the course of their campaigns and (2) a mean-field approximation motivated by the fact that, in the presence of many advertisers, it may not be necessary to track the state of each individual competitor. The first approximation is natural in the setting of ad exchanges where advertisers participate in many repeated auctions and spend a small fraction of the budget in each one of them. In addition, whereas in some ad exchange markets the number of advertisers may be large, it is also useful to study the validity of the second approximation when this is not the case and the same set of advertisers meet repeatedly in common auctions. For this reason, we next isolate the impact of the mean-field approximation and analyze it numerically for markets with a small number of advertisers.

To do so, we propose studying the best response to other advertisers playing the FMFE in a *fluid model* in which there is a continuous flow of arriving impressions at rate η, auctions occur continuously in time, payments are infinitesimal, and budgets are depleted deterministically. The fluid model can be understood as an appropriately normalized market obtained in the limit as budgets and the number of impressions are simultaneously scaled to infinity while the number of players is fixed.

6.2.1. Fluid Model. We introduce a fluid model in which impressions arrive continuously at a rate $\eta = 1$, the time horizon has a length s, and there are K advertisers in the market running synchronous campaigns. We denote by $\mathbf{x}(t) \in \mathbb{R}_+^K$ the vector of budgets remaining of the advertisers at time t as the state vector of the market and by $\mathbf{b}$ the vector of initial budgets. At each point in time, an advertiser determines an action in the space of bidding strategies $\mathcal{B} \triangleq [\underline{V}, \overline{V}] \to \mathbb{R}$, which maps a valuation to a bid. A *control policy* β: $\mathbb{R}_+ \times \mathbb{R}_+^K \to \mathcal{B}$ maps a point in time and state vector to an action.

The dynamics are given by the following. Let the functional u_k: $\mathcal{B}^K \to \mathbb{R}$ denote the instantaneous rate of expected utility obtained by the advertiser k when competing advertisers bid according to a given strategy profile. When the profile is $\mathbf{w} \in \mathcal{B}^K$, we have that

$$u_k(\mathbf{w}) = \alpha_k \mathbb{E}\big[\mathbf{1}\{D_{-k} \le w_k(V_k)\}(V_k - D_{-k})\big],$$

with the expectation taken over the valuation random variable and the maximum competing bid, which is given by $D_{-k} = \max_{i \ne k,\, M_i = 1}\{w_i(V_i)\} \vee r$, where we let $M_i = 1$ indicate that the ith agent participates in the auction. Similarly, we let the functional g_k: $\mathcal{B}^K \to \mathbb{R}$ denote the instantaneous rate of expected expenditure incurred by the advertiser k when advertisers bid according to a given strategy profile, which is given by

$$g_k(\mathbf{w}) = \alpha_k \mathbb{E}\big[\mathbf{1}\{D_{-k} \le w_k(V_k)\} D_{-k}\big].$$

Best-Response Problem. We study the benefit of a unilateral deviation to a strategy that keeps track of the full market state when competitors implement FMFE strategies. In this setting, the FMFE strategies are given by $\beta_i^{\mathrm{F}}(t, \mathbf{x})(v) = v/(1+\mu_i)\mathbf{1}\{x_i > 0\}$, where the multipliers $\boldsymbol{\mu}$ solve the NCP given in (5). The problem faced by advertiser k of determining the optimal payoff of a unilateral deviation when competitors implement the FMFE strategies is given by

$$\begin{aligned} \max_{\beta_k(t,\mathbf{x})} \ & \int_0^s u_k\big(\beta_k(t, \mathbf{x}(t)), \beta_{-k}^{\mathrm{F}}(t, \mathbf{x}(t))\big)\,\mathrm{d}t \\ \text{s.t.} \ & \frac{\mathrm{d}\mathbf{x}(t)}{\mathrm{d}t} = -\mathbf{g}\big(\beta_k(t, \mathbf{x}(t)), \beta_{-k}^{\mathrm{F}}(t, \mathbf{x}(t))\big), \quad t \ge 0, \\ & \mathbf{x}(0) = \mathbf{b}, \quad \mathbf{x}(s) \ge 0. \end{aligned} \tag{6}$$

To simplify our arguments, for the rest of this section, we assume that the reserve price $r = 0$. Moreover, we assume the following tie-breaking rule: when the advertiser under focus and her competitors have run out of budget, the focal advertiser may still bid zero and win the remaining auctions.[14]

[14] This is without loss of generality because by not bidding in a small fraction of the campaign, the advertiser under focus can guarantee that the competitors deplete first, and by saving an infinitesimal budget, she can win all the auctions with no competition for the remaining of the campaign.

6.2.2. Best-Response Analysis. We consider the case when advertisers have equal budgets, distribution of values, and matching probabilities. We do allow, however, for advertiser k to have a different budget than its competitors. Because competitors are symmetric and the dynamics in the fluid model are deterministic, the budgets of the competitors deplete at the same rate. Thus, one can simplify the state by keeping track of the budget of only one competitor.

Some definitions are in order. Let (μ_k, μ_{-k}) denote the multipliers associated with an FMFE. Let V_k^{FMFE} denote the total utility obtained by advertiser k when implementing the FMFE strategy $\beta_k^{\mathrm{F}}(t, \mathbf{x})$.

Next, we define an alternative strategy. Let H: $\mathcal{B} \times \mathbb{R}^2 \to \mathbb{R}$ be a functional given by

$$H(w, \mathbf{p}) = u_k(w, \mathbf{w}_{-k}^{\mathrm{F}}) - p_k g_k(w, \mathbf{w}_{-k}^{\mathrm{F}}) - p_{-k} g_{-k}(w, \mathbf{w}_{-k}^{\mathrm{F}}),$$

where $w \in \mathcal{B}$ is a bidding strategy, $w_i^{\mathrm{F}}(v) = v/(1+\mu_i)$ are the FMFE bidding strategies, $g_{-k}(\cdot)$ denotes the instantaneous rate of expected expenditure incurred by one of the competitors of firm k, and $p_k, p_{-k} \in \mathbb{R}$. Consider the following problem:

$$V_k^{\mathrm{D}} \triangleq \inf_{p_k \ge 0,\, p_{-k}} \ \big\{\alpha s \mathbb{E} V + p_k b_k + p_{-k} b_{-k}\big\} \tag{7a}$$

$$\text{s.t.} \ \sup_{w \in \mathcal{B}} H(w, \mathbf{p}) \le \alpha \mathbb{E} V, \tag{7b}$$

which is a convex optimization problem since the set $\mathcal{P} = \{\mathbf{p} \in \mathbb{R}^2\colon \sup_{w \in \mathcal{B}} H(w, \mathbf{p}) \le \alpha \mathbb{E} V\}$ is convex. The latter follows because the lower-level set of a convex function is convex, and the pointwise supremum of linear functions is a convex function (see, e.g., Boyd and Vandenberghe 2009). Additionally, let V_k^{D} denote the value of (7), with the convention that it is $-\infty$ if it is unbounded; when it is bounded, let p_k^* and p_{-k}^* denote a corresponding optimal solution. Assuming it is well defined, let $\bar{w} \in \arg\max_{w \in \mathcal{B}} H(w, \mathbf{p}^*)$ be the bidding strategy that verifies the supremum in constraint (7b).

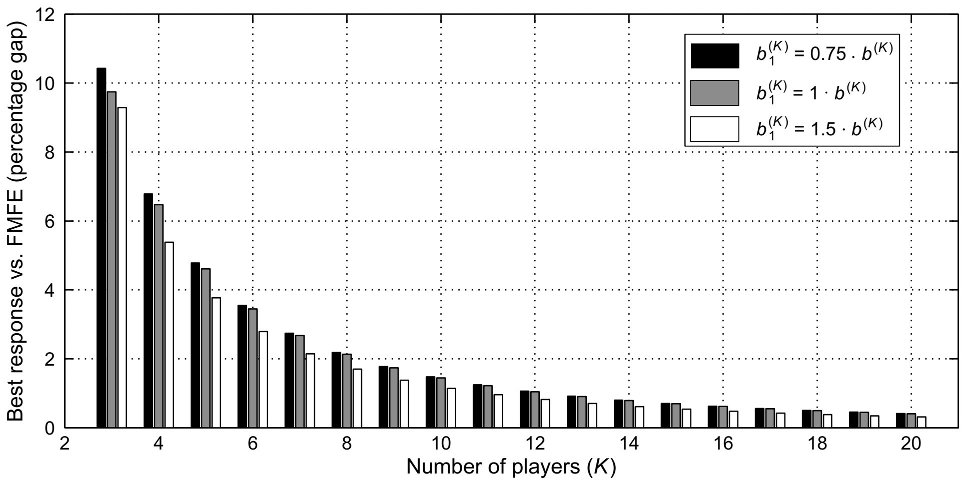
Theorem 6.2 (Best-Response Strategy). *Suppose that bidders' values possess increasing failure rates and have bounded support and that the reserve price is zero. Suppose that all competing advertisers use FMFE strategies. Then,*

(i) *If $V_k^{\mathrm{D}} \le V_k^{\mathrm{FMFE}}$, the FMFE strategy is the optimal control for advertiser k in problem* (6).

(ii) *If $V_k^{\mathrm{D}} > V_k^{\mathrm{FMFE}}$, $\bar{w}$ is well defined, and the optimal strategy for advertiser k in problem* (6) *is to bid according to $\bar{w}(\cdot)$ until competitors deplete their budgets and zero afterwards. Furthermore, this strategy yields exactly V_k^{D}.*

In other words, the result states that the value of the deviant advertiser's control is the maximum of V_k^{D} and V_k^{FMFE}. Furthermore, the result provides a crisp characterization of an optimal policy: one would only need to compute two candidate strategies, the

Figure 3 FMFE vs. Best Response



Note. Advertisers are homogeneous with arrival rate $\eta = 1$, campaign length $s = 16$, competitors' budgets $b^{(K)} = 4/K$, matching probabilities $\alpha^{(K)} = 2/K$, and uniform valuations with support $[0, 2]$.

FMFE strategy and $\bar{w}(\cdot)$, to determine a best response and the associated payoff. We show in the proof that, when $V_k^{\mathrm{D}} > V_k^{\mathrm{FMFE}}$, the competitors will deplete their budgets before the end of the horizon under $\bar{w}$, allowing advertiser k to take advantage of the time during which she operates alone in the market. This result highlights the only type of profitable deviation that one may witness compared with FMFE: use a stationary strategy to deplete competitors faster than what the FMFE strategy does. The strategy involves bidding above one's value in some auctions and carefully balances the lower expected net utility in the first part of the campaign with the benefit of facing no competition at the end of the campaign.[15]

Quite remarkably, one may establish that, in some special cases of interest, the strategy $\bar{w}$ admits a very simple structure: in the cases of uniform and exponential distributions, one may restrict attention to affine bidding functions when searching for a best response (see Corollary 3.1 in §3 of the supplementary appendix). Furthermore, one may establish that in this fluid model the losses of playing FMFE relative to a best response are at most of order $O(\alpha_k)$, a behavior we illustrate numerically next.

6.2.3. Numerical Experiments: FMFE Suboptimality Gap. Intuitively, when there are multiple players in the market, depleting the budgets of the competitors becomes more costly, and as a result, the benefit introduced from deviating from the FMFE strategies becomes negligible. To investigate this, we compare the campaign utility of an advertiser in the fluid model under the FMFE strategy with that of the best response as the number of competitors increases for many problem instances with different parameters. We present the results of a representative instance in Figure 3.[16] Budgets and matching probabilities decrease with the number of competitors so that the average number of matching advertisers per auction remains invariant, equal to 2. We plot the relative suboptimality gap as a function of the number of advertisers. For a given number of advertisers, we analyze the gap when all competitors have the same budget, but we allow the budget of the advertiser under analysis to change and be 75%, 100%, or 150% of the individual budgets of competitors. This allows us to study the gap when the deviant advertiser has varying degrees of market influence.

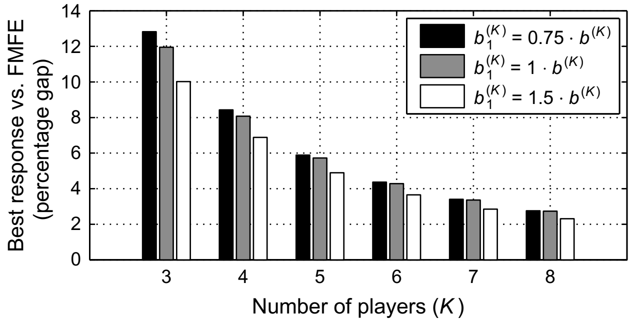
We observe that as the number of players increases, the suboptimality of playing the FMFE decreases fast. As a matter of fact, for the case of identical advertisers ($b_1^{(K)} = b_2^{(K)} = \cdots = b_k^{(K)}$), the FMFE strategy yields utility within 2.5% of that obtained by a best response as soon as there are more than six advertisers in the market. In addition, when the deviant firm has a smaller budget, the advertiser's ability to deplete the firm's competitors decreases.

In Figure 4, we analyze the same setting as earlier except that now we fix the matching probability to $\alpha = 1$. In other words, all advertisers participate in all auctions. In some settings, it is possible to imagine that a small number of advertisers would focus on the same viewer types and hence would compete more intensely. In such a setting, the suboptimality gap of the FMFE decreases fast as the number of competitors increases, dropping below 5% when there are more than five

[15] Lu et al. (2015) also identify similar strategies in which one advertiser tries to deplete the budget of its competitor in a stylized sponsored search auction duopoly model under complete information.

[16] All results can be obtained from the authors upon request.

Figure 4 **FMFE vs. Best Response**



Note. Advertisers are homogeneous with arrival rate $\eta = 1$, campaign length $s = 16$, budgets $b^{(K)} = 4/K$, matching probabilities $\alpha^{(K)} = 1$, and uniform valuations with support $[0, 2]$.

players in the market and getting around 2% when there are eight. We highlight here that the suboptimality gap we estimate in these examples is conservative in that the benchmark policy has unrealistic informational requirements; in practice, bidders would not be able to perfectly monitor competitors' budgets. Hence, their ability to strategize to deplete competitors' budgets would be even more limited.

The fluid analysis and our numerical results above suggest that the value of tracking the market state is small even in the presence of few competitors. In other words, a given bidder has a limited ability to strategize and impact the market when all other competitors play an FMFE strategy. This provides further practical support to use FMFE as a solution concept to study competition in ad exchanges.

7. Conclusions

7.1. A Framework for the Analysis of the Impact of Different Levers

In this paper, the analysis has focused on optimally setting the reserve price. However, the proposed framework based on FMFE is general and may be used to study other important auction design decisions for the publisher. In fact, it is possible to show that the framework proposed allows us to quantify the impact of increasing the allocation of impressions sent to the exchange vis-à-vis collecting the opportunity cost up front on the bidding behavior of advertisers and to optimize this allocation while accounting for budgets.

We also show how one may optimize other dimensions that may be under the control of the publisher such as the extent of user information to disclose to the advertisers. On the one hand, more information enables advertisers to improve targeting, which results in higher bids conditional on participating in an auction. On the other hand, as more information is provided, fewer advertisers match with each user, resulting in thinner markets, which could decrease the publisher's profit.[17] We show that given any mapping from user information to advertiser valuation distribution, one may apply our framework to quantify the impact of budgets on the key trade-offs at play. In particular, we demonstrate this through a stylized model for information disclosure with homogeneous bidders.

These results, available in §4 of the supplementary appendix, complement the ones in the paper and reinforce the importance of reserve price optimization. In particular, we show that proper adjustment of the reserve price is key in (1) making it profitable for the publisher to try selling all impressions in the exchange before utilizing the alternative channel and (2) compensating for the thinner markets created by greater disclosure of viewers' information.

7.2. Building on the Framework

Overall, our results provide a new approach to study ad exchange markets and the publishers' decisions. The techniques developed build on two fairly distinct streams of literature, revenue management and mean-field models, and are likely to have additional applications. The sharp results regarding the publisher's decisions could inform how these markets are designed in practice. At the same time, our framework opens up the door to study a range of other relevant issues in this space. For example, one interesting avenue for future work may be to study the impact of ad networks, which aggregate bids from different advertisers and bid on their behalf, on the resulting competitive landscape and auction design decisions. Similarly, another interesting direction to pursue is to incorporate common advertisers' values and analyze the impact of cherry-picking and adverse selection. Finally, our framework and its potential extensions can provide a possible structural model for bidding behavior in exchanges and open the door to pursue an econometric study using transactional data in exchanges.

Supplemental Material

Supplemental material to this paper is available at http://dx.doi.org/10.1287/mnsc.2014.2022.

Acknowledgments

An earlier version of this paper was circulated under the title "Auctions for Online Display Advertising Exchanges: Approximations and Design." The authors thank Awi Federgruen, Vahab Mirrokni, Hamid Nazerzadeh, Nicolas Stier-Moses, and John Turner for helpful insights as well as participants at the 2012 MSOM Society Annual Conference, 21st International Symposium on Mathematical Programming, INFORMS

[17] This trade-off is discussed in Levin and Milgrom (2010). Fu et al. (2012) study this problem in the context of a static auction without budget constraints and show that if the auctioneer implements the optimal mechanism, then additional data lead to additional revenue.

Annual Meeting 2012, INFORMS Revenue Management and Pricing Section Conference 2013, NET Institute Conference 2013, 14th ACM Conference on Electronic Commerce, and the audiences at various universities for their feedback. They also thank the area editor and three anonymous referees for valuable input that significantly improved the paper. Financial support from the NET Institute (http://www.netinst.org) and the Deming Center (http://www.gsb.columbia.edu/deming) is greatly acknowledged.

Appendix. Selected Proofs

A.1. Proof of Proposition 3.1

We prove the result in three steps. First, we derive the dual of the primal problem by introducing a Lagrange multiplier for the budget constraint. Second, we determine the optimal dual solution through first-order conditions. Third, we show that complementary slackness holds and that there is no duality gap. To simplify notation, we drop the dependence on F_d when clear from the context.

Step 1. We introduce a Lagrange multiplier $\mu \geq 0$ for the budget constraint and let

$$\mathcal{L}_\theta(w, \mu) = \alpha\eta s\,\mathbb{E}\big[\mathbf{1}\{D \leq w(V)\}(V - (1+\mu)D)\big] + \mu b$$

denote the Lagrangian for type θ. (For simplicity, we omit the subindex θ for other quantities.) The dual problem is given by

$$\begin{aligned}
&\inf_{\mu\geq 0}\sup_{w(\cdot)} \mathcal{L}_\theta(w, \mu)\\
&\quad= \inf_{\mu\geq 0}\Big\{\alpha\eta s \sup_{w(\cdot)}\big\{\mathbb{E}[\mathbf{1}\{D \leq w(V)\}(V-(1+\mu)D)]\big\} + \mu b\Big\}\\
&\quad= \inf_{\mu\geq 0}\big\{\alpha\eta s\,\mathbb{E}[\mathbf{1}\{(1+\mu)D \leq V\}(V-(1+\mu)D)] + \mu b\big\}\\
&\quad= \inf_{\mu\geq 0}\big\{\alpha\eta s\,\mathbb{E}[V-(1+\mu)D]^+ + \mu b\big\},
\end{aligned}$$

where the second equality follows from observing that the inner optimization problem is similar to the problem faced by a bidder with value $v/(1+\mu)$ seeking to maximize its expected utility in a second-price auction, in which case it is optimal to bid truthfully. Let $\Psi_\theta(\mu) = \alpha\eta s\,\mathbb{E}[V-(1+\mu)D]^+ + \mu b$. Notice that the term within the expectation is convex in μ; given that expectation preserves convexity, the dual problem is convex. As a consequence of the previous analysis one obtains for any given multiplier $\mu \geq 0$, the policy $w(v) = v/(1+\mu)$ maximizes the Lagrangian.

Step 2. To characterize the optimal multiplier, we shall analyze the first-order conditions of the dual problem. Consider the function $l(x) = \mathbb{E}[V-x]^+ = \int_x^\infty \bar{F}_v(y)\,\mathrm{d}y$. The function l is bounded by $\mathbb{E}\,V$ and continuously differentiable by assumption. Because valuations are independent and conditioning on the maximum bid, we may write the dual objective as $\Psi_\theta(\mu) = \alpha\eta s\,\mathbb{E}[l((1+\mu)D)] + \mu b$. The integrability of D and the differentiability of l, in conjunction with the dominated convergence theorem, yield that Ψ_θ is differentiable with respect to μ (and thus continuous). The derivative is given by $(d/d\mu)\Psi_\theta = b - \alpha\eta s\,\mathbb{E}[\mathbf{1}\{D \leq V/(1+\mu)\}D]$, which is equal to the expected remaining budget by the end of the campaign when the optimal bid function is employed.

Suppose $\alpha\eta s\,\mathbb{E}[\mathbf{1}\{D \leq V\}D] \leq b$—i.e., Ψ_θ admits a nonnegative derivative at $\mu = 0$. Since Ψ_θ is convex, the optimal multiplier is $\mu^* = 0$. Suppose now $\alpha\eta s\,\mathbb{E}[\mathbf{1}\{D \leq V\}D] > b$. The derivative of Ψ_θ is continuous and converges to b as $\mu \to \infty$ by another application of the dominated convergence theorem. We deduce that the equation $\alpha\eta s\,\mathbb{E}[\mathbf{1}\{D \leq V/(1+\mu)\}D] = b$ admits a solution and the optimal multiplier μ^* solves the latter.

Step 3. Combining both cases, one obtains that the optimal multiplier μ^* and the corresponding bid function $\beta_\theta^{\mathrm{F}}(v) = v/(1+\mu^*)$ satisfy $\mu^*(b - \alpha\eta s\,\mathbb{E}[\mathbf{1}\{D \leq \beta_\theta^{\mathrm{F}}(V)\}D]) = 0$, and thus the complementary slackness conditions hold. Additionally, from the first-order conditions of the dual, we get that the bid function $\beta_\theta^{\mathrm{F}}(\cdot)$ is primal feasible. We conclude by showing that the primal objective of the proposed bid function attains the dual objective. That is,

$$\begin{aligned}
&\alpha\eta s\,\mathbb{E}\big[\mathbf{1}\{D \leq \beta_\theta^{\mathrm{F}}(V)\}(V-D)\big]\\
&\quad= \mathcal{L}_\theta(\beta_\theta^{\mathrm{F}}, \mu^*) + \mu^*\big(b - \alpha\eta s\,\mathbb{E}[\mathbf{1}\{D \leq \beta_\theta^{\mathrm{F}}(V)\}D]\big)\\
&\quad= \mathcal{L}_\theta(\beta_\theta^{\mathrm{F}}, \mu^*) = \Psi_\theta(\mu^*),
\end{aligned}$$

where the second equality follows from the complementary slackness conditions and the last from the fact that $\Psi_\theta(\mu^*) = \sup_{w(\cdot)} \mathcal{L}_\theta(w, \mu^*)$, and the fact $\beta_\theta^{\mathrm{F}}$ is the optimal bid function.

A.2. Proof of Theorem 6.1

We prove the result in two steps. First, we lower bound the expected performance of the kth advertiser when all advertisers (including herself) implement the FMFE strategy in terms of the objective value of the fluid problem (1). Second, we upper bound the expected payoff of *any* strategy the kth advertiser may implement when the remaining implement the FMFE strategies via a hindsight bound.

Proposition A.1 (Lower Bound). *Suppose that Assumption* 6.1 *holds and all advertisers implement FMFE strategies* $\boldsymbol{\beta}^{\mathrm{F}}$. *The expected payoff of the kth advertiser is lower bounded by*

$$\frac{1}{\alpha_k\eta s} J_k(\beta_k^{\mathrm{F}}, \boldsymbol{\beta}_{-k}^{\mathrm{F}}) \geq \bar{J}_k^{\mathrm{F}} - O\big((\alpha_k\eta s)^{-1/2}K^{1/2}\big),$$

where $\bar{J}_k^{\mathrm{F}} \triangleq J_k^{\mathrm{F}}/(\alpha_k\eta s)$ *is the normalized objective value of the problem* (1).

The performance metric $J_k(\beta_k^{\mathrm{F}}, \boldsymbol{\beta}_{-k}^{\mathrm{F}})$ may differ from the FMFE value function, given by the objective value of the approximation problem J_k^{F}, since the former takes into account that bidders may run out of budget before the end of their campaigns. The proof is based on the fundamental observation that advertisers bid exactly as prescribed by the FMFE while they have budgets remaining. This allows one to consider an alternate system where advertisers are allowed to bid (i) when they have no budget and (ii) after the end of their campaigns. Thus, in the alternate system the expected performance *exactly* coincides with that of the approximation problem J_k^{F}. Using a coupling argument, the proof shows that the expected performance in the original and alternate systems coincide until the first time *some* advertiser runs out of budget, which in turn is shown to be close to the end of the horizon via a martingale argument.

Proposition A.2 (Upper Bound). *Suppose that Assumption* 6.1 *holds and all advertisers implement FMFE strategies* $\boldsymbol{\beta}^{\mathrm{F}}$ *and the kth advertiser implements an alternative strategy* $\beta \in \mathbb{B}$. *The expected payoff of the kth advertiser is bounded from above by*

$$\frac{1}{\alpha_k\eta s} J_k(\beta, \boldsymbol{\beta}_{-k}^{\mathrm{F}}) \leq \bar{J}_k^{\mathrm{F}} + O\big(\alpha_k + (\alpha_k\eta s)^{-1/2}K^{1/2}\big).$$

To prove the result, we first upper bound the performance of an arbitrary strategy by that of a strategy with the benefit of hindsight (which has complete knowledge of the future realizations of bids and values). This is akin to what is typically done in revenue management settings (see, e.g., Talluri and van Ryzin 1998), with the exception that here, the competitive environment (which is the counterpart of the demand environment in revenue management settings) is endogenous and determined through the FMFE consistency requirement. As a result, the optimal hindsight policy may force competitors to run out of budget so as to reduce competition. To facilitate the analysis of the expected performance of the hindsight policy, the proof considers the same alternate system in which competitors bid regardless of the budget, in which the hindsight policy can be analyzed simply via linear programming duality theory. Because the original and alternate systems coincide until some advertiser runs out of budget, we are left again with the problem of showing that advertisers run out of budget close to the end of the campaign.

The proof concludes by showing that the kth advertiser has a limited impact on the system, in the sense that competitors run out of budget—in expectation—close to the end of their campaigns no matter which strategy the advertiser implements. To this end, the proof exploits that any two advertisers compete a limited number of times during their campaigns to bound the potential impact the kth advertiser may have on her competitors. This result relies heavily on the matching probability decreasing with the scaling.

A.3. Proof of Proposition A.1

Consider an alternate system in which advertisers are allowed to bid (i) when they have no budget and (ii) after the end of their campaigns. The argument revolves around the fact that the performance of the advertiser in consideration (referred to as advertiser k) in the real and alternate coincide until the first time some advertiser runs out of budget. This follows from the fact that advertisers bid exactly as prescribed by the FMFE while they have budgets remaining.

To study the performance on the alternate system, we shall consider the sequence $\{(Z_{n,k}, U_{n,k})\}_{n\ge 1}$ of realized expenditures and utilities of advertiser k in the alternate system. In view of our mean-field assumption, this sequence is i.i.d. and independent of the impressions' interarrival times. The kth advertiser's expenditure in the nth auction is $Z_{n,k} = M_{n,k}\mathbf{1}\{D_{n,-k} \le \beta_k^{\mathrm{F}}(V_k)\}D_{n,-k}$, and her corresponding utility is $U_{n,k} = M_{n,k}\mathbf{1}\{D_{n,-k} \le \beta_k^{\mathrm{F}}(V_{n,k})\}(V_{n,k} - D_{n,-k})$. Additionally, let $b_k'(t) = b_k - \sum_{n=1}^{N(t)} Z_{n,k}$ be the evolution of the kth advertiser's budget in this alternate system, where $N(t)$ is the number of impressions arrived by time t.

The following stopping time will play a key role in the proof. Let $\tilde{N}_k$ be the first auction in which advertiser k runs out of budget—that is, $\tilde{N}_k = \inf\{n \ge 1\colon b_k'(t_n) < 0\}$. This stopping time is relative to all auctions in the market and not restricted to the auctions in which the kth advertiser participates. Similarly, let $\tilde{N}$ denote the first auction in which some advertiser runs out of budget—that is, $\tilde{N} = \min_{i=1}^K \tilde{N}_i$.

Next, we estimate from below the performance of the kth advertiser. Denoting by I_k the number of auctions in which advertiser k participates during his campaign (that is, $I_k = \sum_{n=1}^{N(s)} M_{n,k}$) and by $\tilde{I}_k$ the number of auctions in which advertiser k participates until some agent runs out of budget (that is, $\tilde{I}_k = \sum_{n=1}^{\tilde{N}} M_{n,k}$), one obtains by using a coupling argument that the performance of both systems coincides until time $\tilde{N}$, and as a result,

$$
\begin{aligned}
J_k(\beta^{\mathrm{F}}, \boldsymbol{\beta}_{-k}^{\mathrm{F}}) &\ge \mathbb{E}\left[\sum_{n=1}^{\tilde{N}\wedge N(s)} U_{n,k}\right] \\
&\ge \mathbb{E}\left[\sum_{n=1}^{N(s)} U_{n,k}\right] - \overline{V}\,\mathbb{E}\left[\sum_{n=1}^{N(s)} M_{n,k} - \sum_{n=1}^{\tilde{N}} M_{n,k}\right]^+ \\
&= \mathbb{E}\left[\sum_{n=1}^{N(s)} U_{n,k}\right] - \overline{V}\,\mathbb{E}[I_k - \tilde{I}_k]^+ \\
&\ge \mathbb{E}\left[\sum_{n=1}^{N(s)} U_{n,k}\right] - \overline{V}\,\mathbb{E}[I_k - \alpha_k\eta s]^+ - \mathbb{E}[\alpha_k\eta s - \tilde{I}_k]^+,
\end{aligned}
$$

where the first inequality follows from discarding all auctions after the time some advertiser runs out of budget, the second from the fact that $0 \le U_{n,k} \le M_{n,k}\overline{V}$, and the third from the fact that for every $a, b, c \in \mathbb{R}$ we have that $(a-c)^+ \le (a-b)^+ + (b-c)^+$. In the remainder of the proof, we address one term at a time.

Term 1. Notice that in the alternate system the number of matching impressions in the campaign is independent of the utility, and thus we have that

$$
\mathbb{E}\left[\sum_{n=1}^{N(s)} U_{n,k}\right] = \alpha_k\eta s\,\mathbb{E}[U_{1,k}] = \Psi_k(\mu_k; F_d) + \mu_k(G_k(\boldsymbol{\mu}) - \beta_k) = J_k^{\mathrm{F}},
$$

where the second equality follows from the fact that $\beta_k^{\mathrm{F}}(x) = x/(1+\mu_k)$ and $U_{n,k} = (V_{n,k} - (1+\mu_k)D_{n,k})^+ + \mu_k Z_{n,k}$, and the last follows from complementarity slackness and the optimality of the FMFE multipliers.

Term 2. Note that, for any random variable X and constant x, we have that $\mathbb{E}(X-x)^+ \le (\mathbb{E}X - x)^+ + \sqrt{\mathrm{Var}(X)/2}$ by the upper bound on the maximum of random variables given in Aven (1985). Because the agent participates in each auction with probability α_k, we have that I_k is a Poisson random variable with mean $\alpha_k\eta s$, and one obtains that

$$
\frac{1}{\alpha_k\eta s}\,\mathbb{E}[I_k - \alpha_k\eta s]^+ \le (2\alpha_k\eta s)^{-1/2} = O\big((\alpha_k\eta s)^{-1/2}\big).
$$

Term 3. Define $\tilde{I}_{k,i}$ as the number of auctions that advertiser k participates until agent ith runs out of budget—that is, $\tilde{I}_{k,i} = \sum_{n=1}^{\tilde{N}_i} M_{n,k}$. Using this notation, we obtain that the number of auctions the kth advertiser participates in until someone runs out of budget can be alternatively written as $\tilde{I}_k = \sum_{n=1}^{\min_i \tilde{N}_i} M_{n,k} = \min_i \sum_{n=1}^{\tilde{N}_i} M_{n,k} = \min_i \tilde{I}_{k,i}$. Using this identity, we obtain that

$$
\begin{aligned}
\mathbb{E}[\alpha_k\eta s - \tilde{I}_k]^+ &= \mathbb{E}\Big[\alpha_k\eta s - \min_i \tilde{I}_{k,i}\Big]^+ = \mathbb{E}\Big[\max_i\{\alpha_k\eta s - \tilde{I}_{k,i}\}^+\Big] \\
&\le \max_i\big\{\alpha_k\eta s - \mathbb{E}\tilde{I}_{k,i}\big\}^+ + \sqrt{\sum_i \mathrm{Var}[\tilde{I}_{k,i}]},
\end{aligned}
$$

where the inequality follows from the upper bound on the maximum of random variables given in Aven (1985). That is, for any sequence of random variables $\{X_i\}_{i=1}^n$, we have that $\mathbb{E}[\max_i X_i] \le \max_i \mathbb{E}X_i + \sqrt{((n-1)/n)\sum_i \mathrm{Var}(X_i)}$. Dividing by the expected number of impressions in the horizon and using the bounds on the mean and variance of the stopping times of Lemma 1.3 of the supplementary appendix, we get

$$\frac{1}{\alpha_k \eta s}\mathbb{E}[\alpha_k \eta s - \tilde{I}_k]^+ \\ \leq \max_i \left\{1 - \frac{b_i}{\alpha_i \eta s G_i(\boldsymbol{\mu})}\right\}^+ + \frac{1}{\alpha_k \eta s}\sqrt{\sum_{i=1}^{K} O(b_i)} \\ = O\big((\alpha_k \eta s)^{-1} K^{1/2} \bar{b}^{1/2}\big) = O\big((\alpha_k \eta s)^{-1/2} K^{1/2}\big),$$

where the second inequality follows from the fact that the expected expenditure in the FMFE never exceeds the budget—that is, $\alpha_i \eta s G_i(\boldsymbol{\mu}) \leq b_i$)—and by setting $\bar{b} = \max_i b_i$, and the last follows because $\alpha_k \eta s = O(\bar{b})$ from Assumption 6.1.

References

Abraham I, Athey S, Babaioff M, Grubb M (2013) Peaches, lemons, and cookies: Designing auction markets with dispersed information. *Proc. 14th ACM Conf. Electronic Commerce* (*EC '13*) (ACM, New York), 7–8.

Adlakha S, Johari R, Weintraub GY (2015) Equilibria of dynamic games with many players: Existence, approximation, and market structure. *J. Econom. Theory* 156:269–316.

Alaei S, Arcaute E, Khuller S, Ma W, Malekian A, Tomlin J (2009) Online allocation of display advertisements subject to advanced sales contracts. *Proc. Third Internat. Workshop Data Mining Audience Intelligence Advertising* (*ADKDD '09*) (ACM, New York), 69–77.

Araman V, Fridgeirsdottir K (2011) Cost-per-impression pricing and campaign delivery for online display advertising. Working paper, New York University, New York.

Araman VF, Popescu I (2010) Media revenue management with audience uncertainty: Balancing upfront and spot market sales. *Manufacturing Service Oper. Management* 12(2):190–212.

Arnon A, Mansour Y (2011) Repeated budgeted second price ad auction. Persiano G, ed. *Algorithmic Game Theory*, Lecture Notes in Computer Science, Vol. 6982 (Springer, Berlin), 7–18.

Asmussen S (2003) *Applied Probability and Queues*, 2nd ed. Stochastic Modelling and Applied Probability, Vol. 51 (Springer, New York).

Aven T (1985) Upper (lower) bounds on the mean of the maximum (minimum) of a number of random variables. *J. Appl. Probab.* 22(3):723–728.

Balseiro SR, Besbes O, Weintraub GY (2012) Repeated auctions with budgets in ad exchanges: Approximation result. Technical report, Columbia University, New York.

Balseiro SR, Feldman J, Mirrokni VS, Muthukrishnan S (2014) Yield optimization of display advertising with ad exchange. *Management Sci.* 60(12):2886–2907.

Bergemann D, Said M (2010) *Dynamic Auctions* (John Wiley & Sons, New York).

Boyd SP, Vandenberghe L (2009) *Convex Optimization* (Cambridge University Press, Cambridge, UK).

Celis LE, Lewis G, Mobius M, Nazerzadeh H (2014) Buy-it-now or take-a-chance: Price discrimination through randomized auctions. *Management Sci.* 60(12):2927–2948.

Che YK, Gale I (1998) Standard auctions with financially constrained bidders. *Rev. Econom. Stud.* 65(1):1–21.

Che YK, Gale I (2000) The optimal mechanism for selling to a budget-constrained buyer. *J. Econom. Theory* 92(2):198–233.

Chen Y-J (2011) Optimal dynamic auctions for display advertising. Working paper, University of California, Berkeley, Berkeley.

Doraszelski U, Pakes A (2007) A framework for applied dynamic analysis in IO. Armstrong M, Porter RH, eds. *Handbook of Industrial Organization*, Vol. 3 (North-Holland, Amsterdam), 1887–1966.

Dudey M (1992) Dynamic Edgeworth-Bertrand competition. *Quart. J. Econom.* 107(4):1461–1477.

Eick SG, Massey WA, Whitt W (1993) The physics of the $M_t/G/\infty$ queue. *Oper. Res.* 41(4):731–742.

Facchinei F, Pang JS (2003a) *Finite-Dimensional Variational Inequalities and Complementarity Problems*, Springer Series in Operations Research and Financial Engineering, Vol. 1 (Springer, New York).

Facchinei F, Pang JS (2003b) *Finite-Dimensional Variational Inequalities and Complementarity Problems*, Springer Series in Operations Research and Financial Engineering, Vol. 2 (Springer, New York).

Farias V, Monsch M, Perakis G (2011) Tractable notions of equilibrium in revenue management. Presentation at the MSOM Annual Conference, June 23, INFORMS, Hanover, MD.

Fu H, Jordan P, Mahdian M, Nadav U, Talgam-Cohen I, Vassilvitskii S (2012) Ad auctions with data. *Proc. 5th Internat. Conf. Algorithmic Game Theory* (*SAGT '12*) (Springer-Verlag, New York), 168–179.

Gallego G, Hu M (2014) Dynamic pricing of perishable assets under competition. *Management Sci.* 60(5):1241–1259.

Gallego G, van Ryzin G (1994) Optimal dynamic pricing of inventories with stochastic demand over finite horizons. *Management Sci.* 40(8):999–1020.

Ghosh A, McAfee P, Papineni K, Vassilvitskii S (2009a) Bidding for representative allocations for display advertising. Leonardi S, ed. *Internet and Network Economics, 5th Internat. Workshop* (*WINE 2009*), Lecture Notes in Computer Science, Vol. 5929 (Springer, Berlin), 208–219.

Ghosh A, Rubinstein BI, Vassilvitskii S, Zinkevich M (2009b) Adaptive bidding for display advertising. *Proc. 18th Internat. Conf. World Wide Web* (*WWW '09*) (ACM, New York), 251–260.

Graham C, Méléard S (1994) Chaos hypothesis for a system interacting through shared resources. *Probab. Theory Related Fields* 100(2):157–174.

Gummadi R, Key PB, Proutiere A (2011) Optimal bidding strategies in dynamic auctions with budget constraints. *49th Annual Allerton Conf. Comm., Control, Comput. (Allerton)* (IEEE, New York) 588.

Horn RA, Johnson CR (1991) *Topics in Matrix Analysis* (Cambridge University Press, Cambridge, UK).

Internet Advertising Bureau (2013) Internet advertising revenue report, 2012 full year results. Technical report, Pricewaterhouse-Coopers, New York.

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

Jofre-Bonet M, Pesendorfer M (2003) Estimation of a dynamic auction game. *Econometrica* 71(5):1443–1489.

Laffont JJ, Maskin E (1980) Optimal reservation price in the Vickrey auction. *Econom. Lett.* 6(4):309–313.

Laffont JJ, Robert J (1996) Optimal auction with financially constrained buyers. *Econom. Lett.* 52(2):181–186.

Lariviere MA (2006) A note on probability distributions with increasing generalized failure rates. *Oper. Res.* 54(3):602–604.

Levin J, Milgrom P (2010) Online advertising: Heterogeneity and conflation in market design. *Amer. Econom. Rev.* 100(2):603–607.

Lu S, Zhu Y, Dukes A (2015) Position auctions with budget-constraints: Implications for advertisers and publishers. Working paper, University of Southern California, Los Angeles.

Martínez-de-Albéniz V, Talluri K (2011) Dynamic price competition with fixed capacities. *Management Sci.* 57(6):1078–1093.

Mas-Colell A, Whinston MD, Green JR (1995) *Microeconomic Theory* (Oxford University Press, Oxford, UK).

Maskin ES (2000) Auctions, development, and privatization: Efficient auctions with liquidity-constrained buyers. *Eur. Econom. Rev.* 44(4–6):667–681.

Muthukrishnan S (2009) Ad exchanges: Research issues. Leonardi S, ed. *Internet and Network Economics, 5th Internat. Workshop* (*WINE 2009*), Lecture Notes in Computer Science, Vol. 5929 (Springer, Berlin), 1–12.

Myerson RB (1981) Optimal auction design. *Math. Oper. Res.* 6(1): 58–73.

Najafi-Asadolahi S, Fridgeirsdottir K (2014) Cost-per-click pricing for display advertising. *Manufacturing Service Oper. Management* 16(4):482–497.

Netmining (2011) Mapping the display landscape: A fundamental guide to data, targeting and the future of display. Report, Netmining, New York.

Pai MM, Vohra R (2014) Optimal auctions with financially constrained bidders. *J. Econom. Theory* 150:383–425.

Roels G, Fridgeirsdottir K (2009) Dynamic revenue management for online display advertising. *J. Revenue Pricing Management* 8(5):452–466.

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

Turner J (2012) The planning of guaranteed targeted display advertising. *Oper. Res.* 60(1):18–33.

Vranica S (2013) Automated ad buying surges online. *Wall Street Journal* (October 13), http://online.wsj.com/articles/SB10001424052702304500404579131770474789490.

Vulcano G, van Ryzin G, Maglaras C (2002) Optimal dynamic auctions for revenue management. *Management Sci.* 48(11):1388–1407.

Weintraub GY, Benkard CL, Roy BV (2008) Markov perfect industry dynamics with many firms. *Econometrica* 76(6):1375–1411.

Yang J, Vee E, Vassilvitskii S, Tomlin J, Shanmugasundaram J, Anastasakos T, Kennedy O (2012) Inventory allocation for online graphical display advertising using multi-objective optimization. *Proc. First Internat. Conf. Oper. Res. Enterprise Systems* (*ICORES '12*), 293–304.