Article

A Near-Optimal Bidding Strategy for Real-Time Display Advertising Auctions

Srinivas Tunuguntla and Paul R. Hoban

Abstract

This article introduces a near-optimal bidding algorithm for use in real-time display advertising auctions. These auctions constitute a dominant distribution channel for internet display advertising and a potential funding model for addressable media. The proposed efficient, implementable learning algorithm is proven to rapidly converge to the optimal strategy while achieving zero regret and constituting a competitive equilibrium. This is the first algorithmic solution to the online knapsack problem to offer such theoretical guarantees without assuming a priori knowledge of object values or costs. Furthermore, it meets advertiser requirements by accommodating any valuation metric while satisfying budget constraints. Across a series of 100 simulated and 10 real-world campaigns, the algorithm delivers 98% of the value achievable with perfect foresight and outperforms the best available alternative by 11%. Finally, we show how the algorithm can be augmented to simultaneously estimate impression values and learn the bidding policy. Across a series of simulations, we show that the total regret delivered under this dual objective is less than that from any competing algorithm required only to learn the bidding policy.

Keywords

bidding strategies, internet display advertising, online advertising, online knapsack problems, stochastic optimization

Online supplement: https://doi.org/10.1177/0022243720968547

In 2018, U.S. advertisers spent $48 billion on internet display advertising, a 20% increase year-over-year. Display ad spending now represents 45% of all digital ad spending (Silverman 2019), and analysts expect double-digit growth through at least 2021 (eMarketer 2017). Much of this growth stems from advertiser preferences for more precise audience targeting, which is viewed as a key driver of return on investment (*Forbes* 2015).

The demand for precise targeting has made real-time bidding (RTB) a dominant distribution channel for display ad impressions. In 2020, U.S. advertisers are expected to spend $26 billion purchasing online display advertising through real-time auctions. With a compound annual growth rate of 24%, this will represent nearly half of all display ad spending (Hoelzel and Ballve 2015). Under the RTB approach, individual impressions are sold in real time, frequently via second-price (Vickrey) auctions. Each time a user requests a web page from a publisher's server, auctions are held to determine which advertisements will be served alongside the web page's content. The immediacy and individual nature of RTB targeting allows advertisers to control the timing, location, and recipient of each exposure.

Because of the speed and volume of RTB transactions, advertisers must manage their campaigns via automated targeting and bidding strategies. There are, on average, 1.6 million RTB auctions per second, and each concludes within milliseconds (Shen et al. 2015). During this brief window, targeting algorithms are used to forecast the expected value of each impression opportunity. Advertisers commonly quantify this value as the probability that the recipient undertakes a desired outcome of interest (e.g., click, website visit, purchase), which is known as a conversion (Gordon et al. 2019). The advertiser's bidding algorithm then converts the forecasted impression value to a monetary offer. Advertisers leverage a wide variety of bidding algorithms, including bidding a constant amount per opportunity, pacing campaign spend evenly over time, and dynamically adjusting bids on the basis of forecasted impression value (Google 2018b; Heise, Abou Nabout, and Skiera 2016; Zhang, Yuan, and Wang 2014). Because targeting algorithms and bidding strategies vary across advertisers, two sets of impressions that are valued equivalently by one advertiser may sell for dramatically different prices. Thus, a bidding strategy that balances each

Srinivas Tunuguntla is a doctoral candidate in Marketing, Wisconsin School of Business, University of Wisconsin-Madison, USA (email: stunuguntla@wisc.edu). Paul R. Hoban is Assistant Professor of Marketing, Wisconsin School of Business, University of Wisconsin-Madison, USA (email: paul.hoban@wisc.edu).

impression's value with its cost can significantly improve campaign impact.

Despite following Vickrey protocol, bidding one's valuation is not generally a dominant bidding strategy in RTB auctions. Under Vickrey protocol, each bidder submits a single, sealed bid, and the winner pays an amount equal to the second highest bid. When participating in a series of independent Vickrey auctions or a single such auction, bidding an amount equal to one's true valuation is a well-known, weakly dominant strategy (Vickrey 1961). However, challenges in accurately attributing future incremental profits to individual impressions (i.e., the "attribution problem") have led to the widespread use of intermediate valuation metrics and budgets. These budgets induce interdependence across auctions, as each bid must balance the value and cost of the current impression opportunity with the uncertain values and costs of future opportunities. As a result, truthful bidding is no longer optimal.

In this article, we develop a near-optimal bidding algorithm congruent with the features and constraints of the RTB ecosystem. We term this strategy "near-optimal" because it performs within 2% of the theoretic upper bound possible only with perfect foresight and outperforms the best available alternative by 11%. Furthermore, the algorithm leverages only the limited data available to bidders; the campaign budget, the forecasted impression value, and the sequence of impression costs incurred by the focal advertiser. It is capable of processing the enormous volume and velocity of auctions while meeting the rapid response times required by the advertising exchanges. It ensures that the advertiser's budget constraint is satisfied, and it remains agnostic to competitors' strategies and the distribution of auction attributes. Moreover, it constitutes a unique competitive equilibrium, ensuring that no advertiser has incentive to unilaterally deviate. Finally, it can be combined with Thompson sampling to simultaneously estimate impression values with respect to their incremental impact and optimize bidding behavior. Notably, it can do both while still outperforming the best available alternative algorithm, even when the alternative is provided accurate valuation information and asked only to optimize bidding behavior. In addition, Thompson sampling induces exogenous variation in the advertiser's bidding behavior, mitigating selection biases common in display advertising measurement.

To achieve this, we frame the advertiser's challenge as one of budget-constrained value maximization and draw parallels to the common online knapsack problem. The advertiser wants to purchase the set of impressions delivering greatest value conditional on its budget. In the offline version of the knapsack problem, commonly applied to constrained resource allocation problems in marketing (Anderson, Lodish, and Weitz 1987; Mantrala, Sinha, and Zoltners 1992), firms are assumed to know, with certainty and at the outset, the value and cost of each marketing instrument in which they might invest. This is decidedly unrealistic in the RTB setting, where advertisers must bid for each impression while the values and costs of future impressions remain uncertain. This uncertainty perfectly describes the online knapsack problem, and the extant literature

has generally relied on algorithms assuming the distribution of future values and/or costs are at least partially known a priori.[1] In reality, these distributions are generally unknown and decidedly difficult to estimate (Cai et al. 2017; Zhang, Yuan, and Wang 2014). Without assuming knowledge of future impression values or costs, we prove near-optimal guarantees on the algorithm's regret, a first among algorithmic solutions to the general class of online knapsack problems.

Our algorithm has immediate and significant implications for how $26 billion of display advertising impressions are purchased. Assuming an advertiser employs the best available alternative, it should expect an 11% increase in campaign effectiveness from adopting the proposed approach. Alternatively, it could obtain equivalent campaign value by spending just 89% of its current budget. This is a lower bound on the potential benefits, as many advertisers employ strategies that perform significantly worse than the best available alternative. For example, Google's tools allow advertisers to bid a constant amount (Google 2018b), which our approach outperforms by 16% even when the constant amount is optimized with perfect foresight.

We expect this impact to grow as more marketers participate directly in these exchanges and more advertising formats adopt the RTB model. In a 2017 Association of National Advertisers survey, 35% of respondents indicated that they are expanding their in-house programmatic media buying capabilities, partially in an effort to reduce intermediary costs (Wolfe 2017). Without an efficient bidding strategy, these firms risk eroding much, if not all, of the potential cost savings. Our strategy also has implications for developing technologies such as addressable TV and radio that increasingly look to RTB as a funding model. DISH Network began testing an RTB exchange for addressable TV in 2015, reaching 8 million households across 210 designated market areas (Liyakasa 2015). In 2017, 15%–17% of advertisers purchased addressable television impressions, and an additional 20%–30% planned to experiment with the media in 2018 (Joe 2018). As the underlying auction formats mirror those in the RTB display advertising space, our bidding strategy and associated findings will be directly applicable.

The rest of the article is organized as follows. We review the RTB landscape, present the relevant literature on display advertising and bidding strategies, and examine concepts of competitive equilibrium in this ecosystem. Next, we formalize the advertiser's problem and present the proposed algorithm. We then compare the proposed algorithm with an optimal bidding strategy requiring perfect foresight. We prove that the proposed algorithm converges rapidly to this optimal strategy, achieves zero regret, and constitutes a competitive equilibrium. Then, we show how the proposed bidding algorithm can be combined with Thompson sampling to simultaneously estimate

---

[1] For a general discussion, see Kleywegt and Papastavrou (1998) and Dean, Goemans, and Vondrák (2008); for examples in online display advertising, see Chakrabarty, Zhou, and Lukose (2007) and Amar and Renegar (2018).

impression values and learn the near-optimal bidding policy. We show that the total regret delivered under this dual objective is less than that from any competing algorithm required only to learn the bidding policy and seeded with a priori optimal parameters. This underscores the importance of employing an efficient bidding strategy, even when impression values are uncertain.

Background and Related Literature

The RTB Ecosystem

Within the RTB ecosystem, there are three fundamental roles: publishers, ad exchanges, and advertisers. Publishers manage web pages containing information in which users are primarily interested, and sell advertising nested within and around this content to generate revenue. Advertisers purchase this inventory to promote their brand, product, or service. They are ultimately responsible for designing ad creatives, establishing targeting rules, and specifying campaign objectives. Ad exchanges connect advertisers and publishers, creating a two-sided market in which publishers benefit from the presence of more buyers and advertisers are able to manage campaigns across multiple publisher websites.

When a user visits a publisher's web page, an impression opportunity is created. To fill this opportunity, the publisher sends a bid request to the ad exchange, who then queries advertisers for bids. The bid request is typically accompanied by contextual information about the web page on which the ad would be served (e.g., domain name, URL, topic, keywords) and behavioral and demographic information about the user involved (e.g., cookie id, IP address, geographic location, interests inferred from browsing histories). Each advertiser then estimates the value of the impression opportunity, combining the information accompanying the bid request with any available private information on the user's browsing and purchase history. In practice, this value is generally estimated with respect to the probability of a conversion outcome of interest (e.g., click, website visit, transaction) (Gordon et al. 2019). The advertiser then calculates a bid amount on the basis of the available inputs, and submits this to the advertising exchange. The ad exchange receives bids from all interested advertisers and determines the winner through an auction. The winning advertiser's ad is then displayed to the user on the publisher's web page. To deliver a seamless user experience, ad exchanges typically require advertisers to submit bids within 100 milliseconds. The challenges created by these short response windows are compounded by the volume of bid requests, 1.6 million per second on average (Shen et al. 2015). As a result, valuation and bidding decisions rely on automated algorithms.

Until recently, these auctions near-universally followed second-price rules. However, first-price auctions have risen in popularity, as publishers have begun sending the same impression opportunity to multiple exchanges simultaneously and determining the winner via a subsequent auction between the exchanges (Zawadziński 2018). Because multiple perverse outcomes can occur in the presence of intermediate second-price auctions (e.g., the highest bidder is not guaranteed to win the auction), many exchanges have moved to first-price rules. However, second-price auctions remain the standard when exchanges are the final arbiter of impression placement (see, e.g., OpenX 2019), and many in the industry advocate for publishers to adopt second-price rules for their auctions (PubMatic 2017). In this article, we focus on the scenario in which the final auction follows second-price rules and is preceded only by first-price auctions, if any. However, the dynamic nature of the RTB ecosystem highlights the need for research on effective bidding strategies for a variety of auction protocols.

Marketing Allocation and the Knapsack Problem

In presenting a near-optimal bidding algorithm for use in RTB auctions, we follow a long tradition of marketing research into the design and deployment of automated programs to help marketers allocate scarce resources (Leeflang and Wittink 2000; Little 1970). For example, Rust (1986) details a number of algorithms for allocating advertising spend across media, conditional on all costs and expected outcomes being known at the campaign outset. With such a priori knowledge, budget constrained resource allocation can be framed as an offline knapsack problem (Anderson, Lodish, and Weitz 1987; Mantrala, Sinha, and Zoltners 1992), for which there are known algorithmic solutions with appealing theoretical guarantees on regret (Cormen et al. 2009).

With the automation of media buying on the internet, interest in such systems has experienced a resurgence in marketing. For example, Skiera and Abou Nabout (2013) developed PRO-SAD, an optimized bidding algorithm designed to maximize keyword profit contribution for paid search media. Furthermore, some authors have tried to relax the strong assumptions on a priori knowledge. Schwartz, Bradlow, and Fader (2017) present a multi-armed bandit approach to optimizing display advertising allocation across websites, when the campaign budget and impression costs are known a priori, but impression values are uncertain. Building on the work of Danaher (1991), Paulson, Luo, and James (2018) propose a method to maximize campaign reach by optimally allocating spend across a set of predetermined publishers via RTB auctions. Notably, maximizing reach assumes that all impression values are known a priori,[2] though impression costs may vary.

In contrast, advertisers purchasing display advertising generally bid for each impression opportunity while uncertain of both the costs and values associated with future auctions. As a result, the advertiser must solve an *online* knapsack problem (Chakrabarty, Zhou, and Lukose 2007), learning the optimal policy to balance the value and cost of each opportunity with the uncertain values and costs of future realizations. Prior

---

[2] All individuals are assumed equally valuable until served the first impression, at which point the value of subsequent impressions drops to zero.

research on online knapsack problems has addressed this uncertainty by first assuming that items are drawn from a known distribution (Dean, Goemans, and Vondrák 2008; Kleywegt and Papastavrou 1998; Lueker 1998). Subsequent work relaxed this assumption, requiring only that the cost of all (Agrawal, Wang, and Ye 2014) or some (Amar and Renegar 2018) items be known a priori. Among the weakest assumptions are those made by Chakrabarty, Zhou, and Lukose (2007), who require that the bounds on the value-to-cost ratio be known for all items at the outset. However, the authors acknowledge that their algorithm's performance can vary dramatically with these bounds, about which an advertiser is likely to be uncertain.

Overall, we contribute to this literature by presenting a general solution to the online knapsack problem that does not require the advertiser to know or even forecast future impression values and costs. Despite these relaxed assumptions, we prove near-optimal guarantees on the resulting average and total regret, a first among algorithmic solutions to the general class of online knapsack problems. For advertisers, this simplifies the bidding process and ensures that the algorithm can consistently deliver campaign value closely approximating that available with perfect foresight (i.e., via the offline knapsack), while meeting the realities of the RTB market.

RTB Strategies

Despite following Vickrey protocol, constraints of the RTB ecosystem preclude bidding one's valuation as a dominant strategy (Choi et al. 2020). Across a series of Vickrey auctions, such truthful bidding is a well-known, weakly dominant strategy when bidders can accurately value each item and the auctions are independent (Vickrey 1961). Within RTB, this requires that individual impression values are accurately forecast in terms of the net present value of all incremental future cash flows, and the advertiser does not face a budget constraint. The former ensures that advertisers appropriately value impression opportunities, while the latter is required for the auctions to be independent. Neither of these criteria are generally met.

Calculating the net present value of all incremental future cash flows would require solving the attribution problem at the impression level. This includes accurately measuring display advertising's impact on e-commerce and offline profits, allowing for cross-media synergies, and capturing potential carryover (Danaher and Van Heerde 2018; Dinner, Van Heerde, and Neslin 2014). Because RTB impressions are auctioned individually, such forecasts would need to be made accurately for each impression opportunity. In addition to the practical challenges of linking individual advertising exposures to all future profits, such measurement raises significant privacy concerns by linking online and offline data at the individual level. Faced with these constraints, researchers and practitioners typically rely on intermediate valuation metrics such as page views, clicks, and online sales.

Campaign budgets are also a ubiquitous, and often required, component of the RTB ecosystem. Forecasted impression

values are uncertain, and this is amplified by the use of intermediate outcome metrics. Furthermore, advertisers frequently face binding capital constraints, limiting the amount they can spend on a campaign. Thus, budgets serve as a hedge against uncertain impression values and provide assurance that expenses do not exceed an advertiser's ability to pay. As a result, platforms generally require advertisers to establish budgets before a campaign (Google 2018a), inducing interdependence between auctions.[3]

Faced with these practical constraints, researchers have worked to develop feasible bidding strategies. These strategies generally take as input the forecasted value from the advertiser's targeting algorithm and output a monetary bid amount. Early work focused on pacing algorithms (e.g., Lee, Jalali, and Dasdan 2013), selectively bidding in a subset of auctions to ensure that advertisers do not prematurely exhaust their budget and miss potentially valuable future impression opportunities. However, in doing so, they trade total campaign value for smooth budget delivery.

Others have focused on methodologies that explicitly balance impression values and costs. Zhang, Yuan, and Wang (2014) show that impression opportunities with lower conversion probabilities tended to be undervalued in the marketplace. Capturing this trade-off between bid amount and winning probability, they provide two bidding algorithms. Using a series of simulations and a field experiment, they show that their algorithms outperform several common strategies including bidding a constant amount, bidding below some threshold amount (commonly defined by expected cost per action), and linear form bidding (i.e., bidding proportional to the expected cost per action). However, their approach requires the distribution of highest competing bids to be directly estimated. This is problematic, because each opportunity is defined by a large number of attributes, only some of which are observable by each advertiser (i.e., private information is prevalent). Despite this, we use Zhang, Yuan, and Wang's approach as a benchmark, overcoming challenges associated with estimating the distribution of highest competing bids by fixing the underlying parameters for their approach at the a posteriori optimal values. In contrast, our strategy must simultaneously bid and learn, providing a conservative test of the proposed approach.

Other researchers have directly addressed the strategic nature of the underlying auction. Balseiro, Besbes, and Weintraub (2015) introduce the fluid mean-field equilibrium (FMFE) concept. Relative to perfect Bayesian equilibrium strategies, FMFE strategies drastically reduce the information requirements for agents and dramatically increase computational feasibility. In the RTB environment, the authors show that FMFE strategies approximate best response strategies, even in thin markets with few advertisers. Balseiro and Gur (2019) develop a game-theoretic approach to bidding when all advertisers measure impression value in terms of expected

---

[3] With a budget in place, the outcome of each auction influences the resources available for subsequent bids.

profits. They cast the problem as a sequential game of incomplete information, in which advertisers try to guarantee themselves a portion of the campaign profit achievable with perfect foresight. Their strategy is based on a dynamic learning algorithm and requires only the incurred costs and a fixed learning rate parameter as inputs. This serves as an additional benchmark to the proposed strategy, and we again set the underlying parameter at its a posteriori optimal value.

To date, this literature has generally taken impression values as given, assuming that the advertiser's targeting algorithm is both optimal and accurate. In practice, there is a mismatch between advertiser objectives and the targeting algorithms deployed. While maximizing campaign value requires that impressions be rated with respect to their incremental impact, commonly employed targeting algorithms score impressions based on the associated browser's conversion probability (Choi et al. 2020). Waisman et al. (2019) tried to resolve this conflict, showing that when an advertiser is able to measure impression values in monetary units and faces no budget constraint, the problems of learning an optimal bidding policy and estimating the incremental effects of advertising exposure converge. This result stems from the advertiser's incentive to bid its valuation in the absence of a budget. Under specific assumptions regarding the distribution of impression values and costs, the authors provide an algorithm capable of simultaneously learning an optimal bidding policy and estimating advertising response.

We contribute to the literature on RTB strategies by developing an algorithm that directly maximizes campaign value while satisfying the constraints faced by advertisers. In contrast to pacing strategies, our algorithm directly maximizes total campaign value, constraining spending only to the extent necessary to balance the value and cost of each impression. In contrast to proposed game-theoretic alternatives, the proposed strategy can accommodate any valuation metric, including common intermediate measures and incremental future profits, while still constituting a unique competitive equilibrium. Furthermore, we show that the distribution of competing bids need not be known or even directly estimated, greatly simplifying the advertiser's task. Under these general conditions, we show how the algorithm can be used to learn the optimal targeting and bidding policies simultaneously, greatly expanding its applicability.

Problem Setup and Proposed Algorithm

Defining the Advertiser's Problem

At the start of each campaign, we assume that the focal advertiser allocates a budget, B, to be spent purchasing individual impressions through N real-time auctions.[4] In each auction, $n = 1, 2, \ldots, N$, the advertiser bids b_n and obtains a nonnegative value v_n conditional on serving the impression (i.e.,

winning the auction). Thus, the total value derived from the campaign is

$$V = \sum_{n=1}^{N} x_n v_n, \qquad (1)$$

where x_n is an indicator for winning the auction. The focal advertiser wins an auction when its bid is greater than the highest competing bid. Letting w_n represent the maximum competing bid, $x_n = 1(b_n > w_n)$, where $1(\cdot)$ is the indicator function.

For now, we assume that the advertiser accurately forecasts impression value. In Web Appendix B, we examine how campaign value and the relative performance of the proposed bidding strategy are impacted by inaccurate forecasts. In the "Estimating Valuation While Learning the Bidding Policy" section, we show how the proposed strategy can be combined with Thompson sampling to also estimate impression values.

Because these are Vickrey auctions, the focal advertiser pays w_n if and only if $x_n = 1$. Thus, the total cost of the campaign, C, can be written as

$$C = \sum_{n=1}^{N} x_n w_n.$$

To define the advertiser's objective, we make the following assumptions. First, we assume that the value v_n and the highest competing bid w_n at each auction are drawn from a joint distribution. That is, impression opportunities arrive randomly. Second, we assume that realized impressions do not affect the value or cost of future impression opportunities. Under these assumptions, the advertiser's objective is to maximize the total expected value purchased over all auctions, subject to the budget constraint,

$$\max \sum_{n=1}^{N} \mathbb{E}[x_n v_n]$$

$$\text{s.t.} \sum_{n=1}^{N} x_n w_n \leq B. \qquad (2)$$

Note that the highest competing bid, w_n, enters the objective function only through its impact on whether the focal advertiser wins the auction (i.e., x_n). We approach the advertiser's problem from the perspective of maximizing campaign value subject to the budget constraint, while remaining agnostic as to how value is measured. To do this, we must allow for potentially incomparable units of measure between impression value and cost. That is, we must allow impression values to be measured nonmonetarily to accommodate the common industry practice of valuing opportunities in terms of intermediate, potentially nonmonetary metrics.

However, an argument can be made that advertisers should value opportunities in terms of incremental profits. While this can be infeasible or even inappropriate in some instances (e.g., public universities targeting potential applicants), valuing

---

[4] In Web Appendix A, we show that variation in N is relatively small and predictable, indicating that it is straightforward for advertisers to convert desired campaign duration into N.

opportunities in this manner is likely desirable in most cases. In such instances, w_n would also enter the objective function directly (e.g., $\sum_{n=1}^{N} \mathbb{E}[x_n(v_n - w_n)]$), and the advertiser would try to maximize expected campaign profits, exhausting the budget only if there exists a sufficient quantity of profitable impressions (i.e., impressions for which $v_n - w_n > 0$). In Web Appendix C, we provide just such a specification and show that the solution to this problem is equivalent to that of our more generalizable structure with a simple modification.

Motivating the Algorithm

To motivate the bidding algorithm and provide intuition, we briefly consider a version of the problem with three additional constraints. First, we require that the budget constraint in Equation 2 need only be satisfied in expectation. Second, we assume that the advertiser commits to a per-period strategy and does not update it with new information. Third, we assume that the joint distribution describing impression values and competing bids is static. We use these additional constraints only to motivate our zero-regret bidding strategy. We relax the first two assumptions when we introduce our bidding strategy, and we relax the third when we discuss potential dynamics in Web Appendix D.

Given these assumptions, the advertiser's objective function can then be rewritten as

$$\max \; N\mathbb{E}[xv]$$
$$\text{s.t.} \; N\mathbb{E}[xw] \leq B. \tag{3}$$

Because auctions are stochastically equivalent a priori, they can be treated interchangeably. Thus, maximizing the expected value from each auction also maximizes total campaign value.

Consider the Lagrangian of the optimization problem in Equation 3,

$$\mathcal{L}(\lambda) = N\mathbb{E}[xv] + \lambda(B - N\mathbb{E}[xw])$$
$$= \lambda B + N\mathbb{E}[x(v - \lambda w)]. \tag{4}$$

Recall that x is an indicator for winning the auction (i.e., $x = 1(b > w)$). Given any λ, the Lagrangian is maximized if $x = 1$ when $v - \lambda w > 0$ and $x = 0$ otherwise. Intuitively, the advertiser only wants to win an auction for which the value, scaled by λ, exceeds the cost (i.e., $v/\lambda > w$). Because these are second-price auctions, this can be guaranteed by bidding $b = v/\lambda$.

To specify the complete bidding strategy, it remains to optimize for the choice of λ. Differentiating Equation 4 with respect to λ gives $\delta\mathcal{L}/\delta\lambda = B - N\mathbb{E}[xw]$. Setting this equal to zero provides the first order condition,

$$\frac{B}{N} = \mathbb{E}[xw]. \tag{5}$$

Thus, the optimal value of λ, which we refer to as λ^*, is such that average cost per auction is equal to the available budget per auction. We leverage this critical insight in "The Algorithm" subsection to develop a learning algorithm that converges to λ^* using only the campaign budget and revealed sequence of costs. The advertiser's optimal bidding strategy is then

$$b_n = \frac{v_n}{\lambda^*}. \tag{6}$$

This bidding strategy has several intuitive features. First, it ensures that the advertiser wins the subset of auctions for which the value per expenditure, v_n/w_n, is greatest. Conditional on the budget, λ^* defines a threshold for this ratio, and the advertiser will purchase all impressions for which the value per expenditure exceeds λ^*. Furthermore, this threshold is a decreasing function of the campaign budget, indicating that when the budget increases, λ^* decreases, the focal advertiser's bids increase, and the advertiser wins a larger subset of the auctions. If the budget decreases, the opposite occurs.

Second, λ^* represents the expected change in total campaign value when the budget is increased by \$1. This stems directly from the fact that λ is the Lagrange multiplier in Equation 4 and λ^* is its optimal value. For example, if impression value were measured as the incremental probability of conversion, λ^* is the number of additional conversions that could be obtained for a \$1 increase in budget. In this way, the value of λ^* from previous campaigns can inform budget decisions for future campaigns.

Third, the optimal bid depends only on the focal advertiser's valuation and λ^*. Thus, conditional on these measures, the optimal bid is independent of both the underlying attributes and competitor bids, alleviating the need to forecast future realizations of these quantities. Such forecasts are particularly challenging because of the dimensionality of the problem (Cai et al. 2017; Zhang, Du, and Wang 2016). We formally discuss this independence in Web Appendix E. This also implies the bidding strategy is agnostic with respect to the advertiser's unit of measure. λ^* converts any nonmonetary unit of measure for v_n to a dollar value for bidding while still generalizing to cases in which v_n is measured monetarily.

The Algorithm

In this subsection, we present a learning algorithm that converges to λ^*. The algorithm leverages only the campaign budget, the impression value output by the advertiser's targeting algorithm, and the incurred cost of each auction. The incurred cost is the second-highest bid when an advertiser wins an auction and zero otherwise. Thus, the algorithm requires only the information that is readily available to all advertisers bidding in real-time exchanges.

From Equation 5, λ^* is such that the average cost per auction is equal to the available budget per auction. Thus, if λ^* were known at the start of the campaign and the advertiser bid according to Equation 6, the strong law of large numbers guarantees

$$\lim_{n \to \infty} \left(\frac{B}{N} - \frac{1}{n} \sum_{i=1}^{n} c_i \right) = 0, \tag{7}$$

where $c_i = w_i 1(b_i > w_i)$ is the cost incurred at auction i.

If an advertiser were to select $\lambda \neq \lambda^*$, Equation 7 would not hold, but the resulting value would reflect the difference between λ and λ^*. An advertiser using $\lambda < \lambda^*$ would submit larger bids, win more auctions, and incur greater costs. Thus, the average cost per auction would exceed the available budget per auction, and the left hand side of Equation 7 would be negative. Conversely, the left-hand side of Equation 7 is positive when $\lambda > \lambda^*$. Furthermore, larger differences between λ and λ^* produce more extreme differences between the budget and cost per auction. This suggests that an iterative stochastic approximation algorithm (Pasupathy and Kim 2011) can be used to learn λ^* by comparing the available budget per auction to the revealed sequence of auction costs, motivating the proposed algorithm. The advertiser begins with an arbitrary λ. It updates this value after each auction based on the difference between the history of incurred costs, c_i, $i \in \{1, \ldots, n\}$, and the average available budget per auction, $\rho = B/N$. Let λ_n represent the iterate before auction n. At auction n, the advertiser bids $b_n = v_n / \lambda_n$ and observes cost $c_n = w_n 1(b_n > w_n)$. The iterate λ_{n+1} is then computed as

$$\lambda_{n+1} - \frac{1}{n} \sum_{i=1}^{n} \lambda_i = -\frac{1}{\mu} \left(\rho - \frac{1}{n} \sum_{i=1}^{n} c_i \right). \tag{8}$$

Intuitively, λ_{n+1} is set such that the difference between it and the average λ to that point is proportional to the difference between the average budget per auction and the average incurred cost per auction. This proportion is defined by $\mu > 0$, a hyperparameter influencing the rate of convergence that is set by the advertiser at the start of the campaign. In the "Empirical Performance" section, we show that the algorithm's performance is generally robust to the choice of μ. The negative sign on the right-hand side reflects the fact that expected cost is a decreasing function of λ. In the "Empirical Performance" section, we also present a stylized example that provides the operational details of the algorithm and demonstrates the convergence of the bidding strategy.

This algorithm relaxes the first two assumptions from the "Motivating the Algorithm" subsection. First, the algorithm stops at the end of N auctions, or earlier if the remaining budget is less than the maximum potential competing bid, $\overline{w}$. This guarantees the budget constraint is satisfied ex post instead of only in expectation (i.e., campaign spend cannot exceed the budget). Second, the advertiser explicitly updates its strategy at each auction through the aforementioned stopping rule and changes to λ_n.

Analysis

In this section, we show that the proposed algorithm converges rapidly to the optimal bidding strategy, delivers a nearly

optimal amount of forecasted value, and constitutes a unique competitive equilibrium. We present formal proofs for the underlying lemmas and theorems in Web Appendix F. We begin by establishing that the aforementioned update rule results in a sequence $\{\lambda_n\}$ that converges to λ^* at a rate of $O(1/\sqrt{n})$, the same as achieved by the central limit theorem. Next, we establish a theoretical upper bound on the total value achievable by any bidding strategy, introducing an oracle with perfect foresight about the values and competing bids for all auctions at the campaign's start. We define regret to be the difference between the expected value obtained by the oracle and that delivered by the proposed strategy. We then show that the algorithm achieves zero regret, indicating that the expected regret per auction approaches zero as the number of auctions grows large. We further show that the total regret achieved by the proposed algorithm is near optimal, nearly matching the lower bound on regret derived by Marchetti-Spaccamela and Vercellis (1995). Finally, we establish that the proposed bidding strategy constitutes a unique competitive equilibrium by framing the interactions between competing advertisers as a dynamic game.

Proving Convergence

Lemma 1. Using the update rule in Equation 8, with probability at least $1 - (1/\sqrt{N})$,

$$\forall n \in \{1, \ldots, N\}, \ |\lambda_n - \lambda^*| = O\left(\sqrt{\frac{\log(N)}{n}} \right). \tag{9}$$

Lemma 1 states that the sequence $\{\lambda_n\}$ converges to λ^* at a rate of $O(1/\sqrt{n})$ with high probability.[5] That is, the proposed algorithm converges to the optimal strategy in Equation 6.

Oracle and Regret

To quantify the performance of any bidding strategy, we compare the expected total value obtained through the strategy to that achievable with perfect foresight. Consider an oracle that has access to all the values and competing bids at the start of the campaign. The oracle's objective reduces to the well-known zero-one knapsack problem. In the canonical example of the problem, a traveler is packing for a trip. He would like to bring all of his possessions but is limited by the weight he can carry. Thus, he packs the combination of possessions that will deliver the most value, subject to the amount of weight he can carry and the indivisibility of each item. For the oracle, each item is an auction, the weight constraint is the campaign budget, and

---

[5] The convergence rate of λ_n cannot be derived directly from the central limit theorem, because λ_n depends on all prior values λ_i, $i \in 1, 2, \ldots, n-1$ and the incurred costs c_i, $i \in 1, 2, \ldots, n-1$. Because the distribution of c_n depends on λ_n, it depends implicitly on the previous observed costs. As a result, the observed costs are neither independently nor identically distributed across auctions, violating the assumptions of the central limit theorem.

the weight and value of each item are given by the cost and valuation, respectively.

Clearly, the optimal value obtained by the oracle is an upper bound on the value obtained by any bidding strategy lacking foresight into the future values and costs. The relaxed knapsack problem, in which items are assumed divisible, provides an upper bound to the zero-one knapsack solution, and has the benefit of being solved in polynomial time. With this relaxation, the optimal assortment can be selected by sorting the items in descending order according to the ratio of value and cost, and then selecting the largest prefix of items that does not violate the budget constraint. In other words, there exists a critical threshold, Ω, such that all items in the optimal assortment satisfy $v_n/w_n > \Omega$ (Cormen et al. 2009). To satisfy this condition, the oracle can bid b_n, such that $b_n = v_n/\Omega$. Because the solution allows for the inclusion of partial (vs. whole) items, the optimal value of the relaxed knapsack problem may overstate the optimal value of the zero-one knapsack problem (Cormen et al. 2009). Thus, it also constitutes an optimistic upper bound on the optimal value of any online bidding strategy in which items are indivisible. The algorithm proposed in "The Algorithm" subsection approximates this solution by learning Ω.

We define the regret resulting from any bidding strategy to be the difference between the expected value obtained by the oracle's solution to the relaxed knapsack problem and the expected total value obtained by the bidding strategy. A zero-regret bidding strategy is a strategy whose average regret per auction tends to zero with probability 1 when the number of auctions tends to infinity. Intuitively, zero-regret strategies are guaranteed to converge to the oracle's strategy, given a sufficient number of auctions.

Proving Appropriate Pacing

To prove that the proposed algorithm is a zero-regret strategy, we first establish that it is unlikely to prematurely exhaust the campaign budget and the risk of doing so decreases as the number of auctions grows large. Let T be the auction after which the algorithm stops. This could be N or an earlier auction if the remaining budget drops below the maximum competing bid, $\overline{w}$. The following theorem derives a probabilistic bound on the stopping auction T.

Theorem 1. There exist constants C and N_0 such that for $N > N_0$,

$$\mathbb{P}\left[T \geq N - C\sqrt{N \log(N)} \right] \geq 1 - \frac{2}{\sqrt{N}}.$$

Theorem 1 states that the algorithm exhausts the capacity with at most $O\left[\sqrt{N \log(N)}\right]$ auctions remaining, with high probability. Because $\{\lambda_n\}$ converges rapidly to λ^*, and λ^* is such that the expected cost per auction equals the average budget per auction, the total cost after auction N is close to $N\rho$, or

equivalently B. As a result, with high probability, the algorithm does not prematurely exhaust the budget.

Proving Regret Bounds

Theorem 2. The expected total regret is $O\left[\sqrt{N \log(N)}\right]$.

Theorem 2 states that the algorithm achieves zero regret. Intuitively, the total regret is small, because the algorithm converges rapidly to the optimal strategy and does not prematurely exhaust the budget. Because the regret $O\left[\sqrt{N \log(N)}\right]$ is sublinear in N, the average regret per auction tends to zero with probability 1 when number of auctions, N, tends to infinity.

The total regret achieved by the algorithm is also near-optimal. Marchetti-Spaccamela and Vercellis (1995) show that the expected regret even when λ^* were known at the start of the campaign is $O(\sqrt{N})$. The proposed algorithm achieves this lower bound with only an additional logarithmic factor. This drives the impressive empirical performance shown in the "Empirical Performance" section.

Establishing a Unique Competitive Equilibrium

When characterizing auction behavior, it is important to carefully consider the competitive landscape, including potential dynamic interactions between bidders. We frame these interactions as advertisers competing in a dynamic game, where the bidding strategy employed by one advertiser affects the bidding landscape for others. However, the number of potential competitors and limited competitive information generally precludes analysis relative to classic game-theoretic concepts of equilibrium (e.g., Nash equilibrium).

For advertisers, formulating direct competitive responses, a common requirement in classic game-theoretic analyses, is decidedly unrealistic and practically infeasible. In the RTB environment, bidders generally know little about the competition they face. They are typically unaware of the number of competing advertisers, much less their identities, valuations, budgets, objectives, or bidding strategies. During a campaign, advertisers generally bid on millions of impressions and compete with thousands of advertisers. Budgets vary greatly by advertiser and campaign, and not all advertisers actively bid in all auctions. Campaign objectives vary along myriad dimensions, and are not always directly comparable. Furthermore, the exchanges generally share only the second-highest bid, and only with the winning bidder. As a result, solving for classic equilibrium concepts in such a game is generally intractable, both analytically and computationally, and is unlikely to reflect the data generating process.

Instead, we leverage the concept of an FMFE to show that our strategy closely approximates the behavior of rational bidders in these markets. Introduced by Balseiro, Besbes, and Weintraub (2015), the FMFE approximation relies on two assumptions. First, the mean field approximation assumes that the number of advertisers is sufficiently large, such that no one advertiser can unilaterally influence the bidding landscape. As a result, advertisers can rely on a stationary distribution

characterizing the maximum competing bids. Second, the fluid approximation assumes that the number of bidding opportunities for an advertiser is large and the expected cost per opportunity is small relative to the advertiser's budget. Both these criteria fit well in the RTB context. Given these criteria, Balseiro, Besbes, and Weintraub (2015) show that FMFE strategies closely approximate the behavior of rational bidders, even in small markets with few advertisers, and establish the criteria for the existence and uniqueness of the equilibrium.

Theorem 3. FMFE always exists and is unique.

Theorem 3 establishes the competitive equilibrium. In the formal proof, we show that there exists a unique profile of scaling factors λ^* such that, the bidding strategy is optimal simultaneously for each advertiser. Thus, the strategy profile resulting from these scaling factors constitute a unique equilibrium.

Empirical Performance

In this section, we explore the empirical performance of the proposed bidding algorithm. First, we present a stylized example that provides the operational details of the algorithm and demonstrates the convergence of the bidding strategy. Next, we examine the algorithm's finite sample performance across a series of 100 realistic simulations and 10 real-world campaigns. We evaluate this performance based on the difference in campaign value possible with perfect foresight (i.e., the oracle in the "Oracle and Regret" subsection) and that delivered by the proposed strategy. We also compare this performance with that of several alternative strategies identified from the extant literature and industry practice. Across the simulations and real-world campaigns, we consistently find that the proposed algorithm closely approximates the theoretic upper bound and significantly outperforms available alternatives.

Stylized Example

Consider an advertiser managing a retargeting campaign, who evaluates each impression opportunity based on two attributes: the age of the user and time elapsed since the user visited the advertiser's website. The advertiser's valuation, measured as the probability of conversion, is obtained through a logit model of the two attributes:

$$v = \frac{1}{1 + e^{-\beta_0 + \beta_1 A + \beta_2 T}}, \tag{10}$$

where A is the age of the user and T is the number of hours since the user visited the advertiser's website. For the sake of this example, we set the coefficients to be $\beta_0 = -1, \beta_1 = .1, \beta_2 = -.3$. Note that we use only two attributes for simplicity of exposition. In practice, advertisers use a large number of attributes corresponding to contextual, demographic, and behavioral information.

Next, we illustrate how the proposed algorithm can be applied to compute bids for the first few auctions of a campaign.

We set the average budget available per auction ρ equal to .5. For this example, we set the parameter μ in Equation 8 to be equal to 1. For real-world campaigns, as we discuss subsequently, we recommend setting μ to be equal to .001 or .01.

We arbitrarily initialize $\lambda_1 = 1.0$ at the start of the campaign. At auction 1, the advertiser observes that the user is 19 years old, and 1.73 hours have elapsed since their last visit to the advertiser's website (i.e., $A_1 = 19$ and $T_1 = 1.73$). The advertiser computes the value of the impression opportunity using Equation 10.

$$v_1 = \frac{1}{1 + e^{-1 + .1 \times 19 - .3 \times 1.73}} = .59.$$

The advertiser's bid is then

$$b_1 = \frac{v_1}{\lambda_1} = \frac{.59}{1.0} = .59.$$

Let the highest competing bid be $w_1 = 2.78$. Because the focal advertiser's bid is less than the competing bid, it loses the auction, and the incurred cost is $c_1 = w_1 1(b_1 > w_1) = 0$. The advertiser then updates λ using Equation 8 as follows:

$$\lambda_2 = \lambda_1 - \frac{1}{\mu}(\rho - c_1) = 1.0 - \frac{1}{1.0}(.5 - .0) = .5.$$

At auction 2, the advertiser observes $A_2 = 33$ and $T_2 = 11.12$. Thus, the value is

$$v_2 = \frac{1}{1 + e^{-1 + .1 \times 33 - .3 \times 11.12}} = .26.$$

The advertiser's bid is $b_2 = v_2/\lambda_2 = .26/.5 = .52$. Let the highest competing bid be $w_2 = 1.13$. The advertiser again loses the auction, and the incurred cost is $c_2 = 0$. To compute λ_3 using Equation 8, we must first compute the average value of lambda per auction so far $(1/2)(1.0 + .5) = .75$, and the average incurred cost per auction $(1/2)(.0 + .0) = .0$. Using these to compute λ_3:

$$\lambda_3 = \frac{1}{2}(\lambda_1 + \lambda_2) - \frac{1}{\mu}\left[\rho - \frac{1}{2}(c_1 + c_2)\right] = .75 - \frac{1}{1.0}(.5 - .0) = .25.$$

At auction 3, $A_3 = 27$ and $T_3 = 1.32$. Therefore, $v_3 = .79$ and $b_3 = v_3/\lambda_3 = 3.16$. Assuming the highest competing bid $w_3 = 1.52$, the advertiser wins the auction. Because these are second-price auctions, the incurred cost is equal to the highest competing bid $c_3 = w_3 = 1.52$. The average value of lambda per auction so far is $(1/3)(1.0 + .5 + .25) = .58$, and the average incurred cost per auction is $(1/3)(.0 + .0 + 1.52) = .51$. λ_4 is then

$$\lambda_4 = \frac{1}{3}(\lambda_1 + \lambda_2 + \lambda_3) - \frac{1}{\mu}\left[\rho - \frac{1}{3}(c_1 + c_2 + c_3)\right]$$

$$= .58 - \frac{1}{1.0}(.5 - .51) = .59.$$

Table 1 repeats this computation for the first ten auctions, while Figure 1 demonstrates the convergence of the sequence $\{\lambda_1, \lambda_2, \ldots\}$ to the optimal scaling factor Ω, as computed by an

Table 1. Stylized Example.

Auction	Lambda λ_n	Age A_n (Years)	Time Since Website Visit T_n (Hours)	Value v_n	Max Competing Bid w_n	Ratio v_n/w_n	Bid $b_n = v_n/\lambda_n$	Incurred Cos $c_n = w_n I(b_n > w_n)$	Average Lambda $\frac{1}{n}\Sigma_{i=1}^n \lambda_i$	Average Incurred Cost $\frac{1}{n}\Sigma_{i=1}^n c_i$
1	1.0	19	1.73	.59	2.78	.21	.59	0	1.0	.0
2	.5	33	11.12	.26	1.13	.23	.52	0	.75	.0
3	.25	27	1.32	.79	1.52	.52	3.15	1.52	.58	.51
4	.59	29	7.17	.44	1.06	.41	.74	0	.59	.38
5	.47	34	9.98	.36	1.82	.2	.76	0	.56	.3
6	.37	19	.53	.68	.2	3.39	1.85	.2	.53	.29
7	.32	27	3.28	.67	1.83	.37	2.13	1.83	.5	.51
8	.51	25	6.74	.37	1.26	.3	.74	0	.5	.44
9	.44	25	4.96	.5	1.82	.28	1.13	0	.49	.4
10	.39	25	14.71	.05	.03	1.82	.13	.03	.48	.36

oracle with perfect foresight into the values and computing bids for all auctions.

To calculate this value, the oracle first sorts the auctions in descending order according to their value-to-cost ratio (see Table 1, Column 7). Here, the sorted order of auctions is $\{6, 10, 3, 4, 7, 8, 9, 2, 1, 5\}$. The oracle then selects the largest prefix of impressions that does not violate the budget constraint. With an average budget per auction of \$.50 and ten auctions, the total budget is \$5. The largest prefix of impressions that can be purchased for \$5 come from auctions $\{6, 10, 3, 4, 7\}$ with cost $.2 + .03 + 1.52 + 1.06 + 1.83 = 4.64$. The oracle declines to purchase the impression in auction 8 (i.e., that with the next highest value-to-cost ratio), because this would violate the budget constraint. All the auctions in the purchased prefix satisfy $v_n/w_n > \Omega$, and the minimum value-to-cost ratio among the set is $v_7/w_7 = .37$. Thus, $\Omega = .36$.

Simulations

We begin testing the finite sample performance of our algorithm using a series of 100 realistic simulations. For each impression opportunity, the focal advertiser's valuation and the maximum competing bid are drawn from a joint distribution $f(v, w)$. The marginal distribution of the focal advertiser's valuation, $f_v(v)$, is assumed to follow a truncated normal distribution with mean .5 and standard deviation .1 (i.e., $v \sim \mathcal{N}(.5, .01)$ s.t. $0 \le v$). The conditional distribution of the maximum competing bid given valuation, $f_{w|v}(w|v)$ is assumed to follow the gamma distribution $\Gamma(2.75, v)$. By setting the scale parameter equal to the focal advertiser's valuation, this explicitly incorporates correlation in impression values across advertisers. We have tested similar designs without this correlation, with no meaningful difference in the algorithm's performance.

Notably, the previous distributions are used only to simulate the underlying sequence of auctions. Throughout this article, we assume that the distribution of $f(v, w)$ is unknown to the advertiser. Thus, none of the bidding algorithms directly
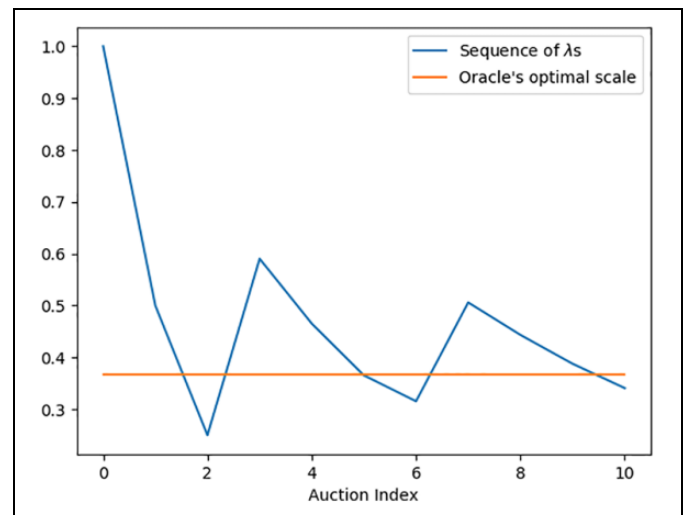

Figure 1. Stylized example: convergence of λ.
Notes: This plot shows the convergence of λ in the stylized example, and the optimal scale computed by the oracle.

leverage these functional forms. In many ways, this is similar to simulating data from known distributions to empirically evaluate the properties of nonparametric estimation methods.

The underlying distributions were selected in collaboration with a demand-side platform (DSP). This DSP purchases display advertising impressions through the real-time exchanges on behalf of hundreds of advertisers, and provided a random sample of impressions costs for 10.7 million impressions it served. This distribution is plotted in gray in Figure 2. The distribution of $f_w(w)$ realized in our simulations is plotted as a dashed line. That the latter overstates the former is intentional. As is common, the collaborating DSP only observes pricing information for the set of impressions it won. Because the DSP is more likely to win auctions when competing bids are low and lose when they are high, the observed distribution of impression costs understates the distribution of highest competing bids. $f_w(w)$ corrects for this understatement. We also
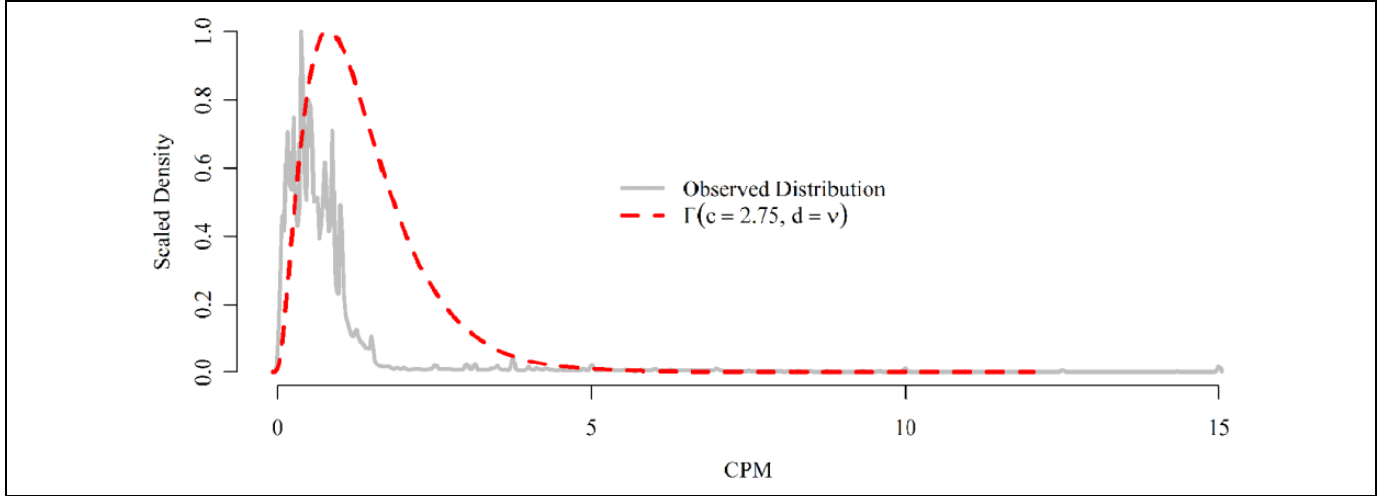
Figure 2. Observed CPM distribution over 10.7 million impressions.
Notes: This plot contains the price paid for 10.7 million randomly selected impressions at a collaborating DSP. The dashed line represents the realized distribution of highest competing bids, simulated from $\Gamma(2.75, v)$.

tested $f_{w|v}(w|v) \sim \Gamma(2.75, .5v)$, which closely approximates the distribution of costs provided by the DSP, and found no meaningful differences in the algorithm's performance.

Each of the 100 simulations contains 10 million auctions, with (v, w) sampled randomly from $f(v, w)$. μ, the learning rate parameter in Equation 8, was fixed at .001. This is in line with the stochastic gradient descent literature, which advises selecting a small value for the learning rate parameter. The algorithm's performance is fairly robust to this parameter μ. An order of magnitude change, $\mu = .01$ or $\mu = .0001$, resulted only in small change in performance ($<2\%$). In the results presented here, $\mu = .001$ across all simulations, providing further evidence of the algorithm's robustness. We recommend advertisers to set μ to be equal to .001 or .01. The budget was set to \$200, in line with the average campaign budget per 10 million auctions at the collaborating DSP. The bid amount for each opportunity is given by $b_n = v_n/\lambda_n$.

Numerical Efficiency

To test the speed at which the sequence $\{\lambda_n\}$ converges to λ^*, we start our algorithm from four distinct starting points, $\lambda_0 \in \{.1, 1.0, 10.0, 100.0\}$. For each of the 100 simulations and four starting points, the resulting sequences for λ_n can be seen in Figure 3. We measure convergence based on the first observation for which $\lambda_n \in \lambda^* \pm 5\%$ and there exists no k such that $\lambda_{n+k} \notin \lambda^* \pm 5\%$. λ_n converges to λ^* within 1 million auctions for all starting points $\lambda_0 \in \{.1, 1.0, 10.0, 100.0\}$. With an average of 1.6 million auctions per second in the real time exchanges (Shen et al. 2015), this indicates that an advertiser bidding in all auctions can expect convergence in less than one second, while convergence would take just over ten minutes when bidding in .1% of auctions.

Comparison to Alternative Algorithms and the Theoretic Upper Bound

To understand the performance of our algorithm, we compare the total campaign value delivered by our approach to several alternatives. The details of each of the benchmark algorithms are provided in Web Appendix G. First, we focus on the ORTB1 algorithm introduced by Zhang, Yuan, and Wang (2014). This specification is designed for situations in which reserve prices are low or nonexistent, which accurately describes our simulations. Through a series of simulations and a field experiment, Zhang, Yuan, and Wang show that ORTB1 outperforms a number of alternative, common bidding strategies (e.g., bidding a constant amount, bidding below some threshold amount, and bidding in proportion to the expected cost per action). By showing that our algorithm outperforms ORTB1, we provide evidence that it also outperforms these alternatives.

To provide a strenuous test of our approach, we set the c and λ parameters for ORTB1 at their a posteriori optimal values for each simulation while restricting our own algorithm to simultaneously bidding and learning λ^*. To set the parameters for ORTB1, we perform a brute-force search to find the optimal parameter values for each simulation given the data. We then use these parameters to guide bidding behavior in the same data. In doing so, we are using the data once to estimate the necessary parameters and again to guide bidding behavior. This provides ORTB1 with perfect foresight, establishing an upper bound on its performance.

Second, we compare our performance with the algorithm proposed by Balseiro and Gur (2019; BG hereinafter). We again set the underlying parameter at its a posteriori optimal value to provide a conservative comparison to the proposed strategy. As described in the "RTB Strategies" subsection,
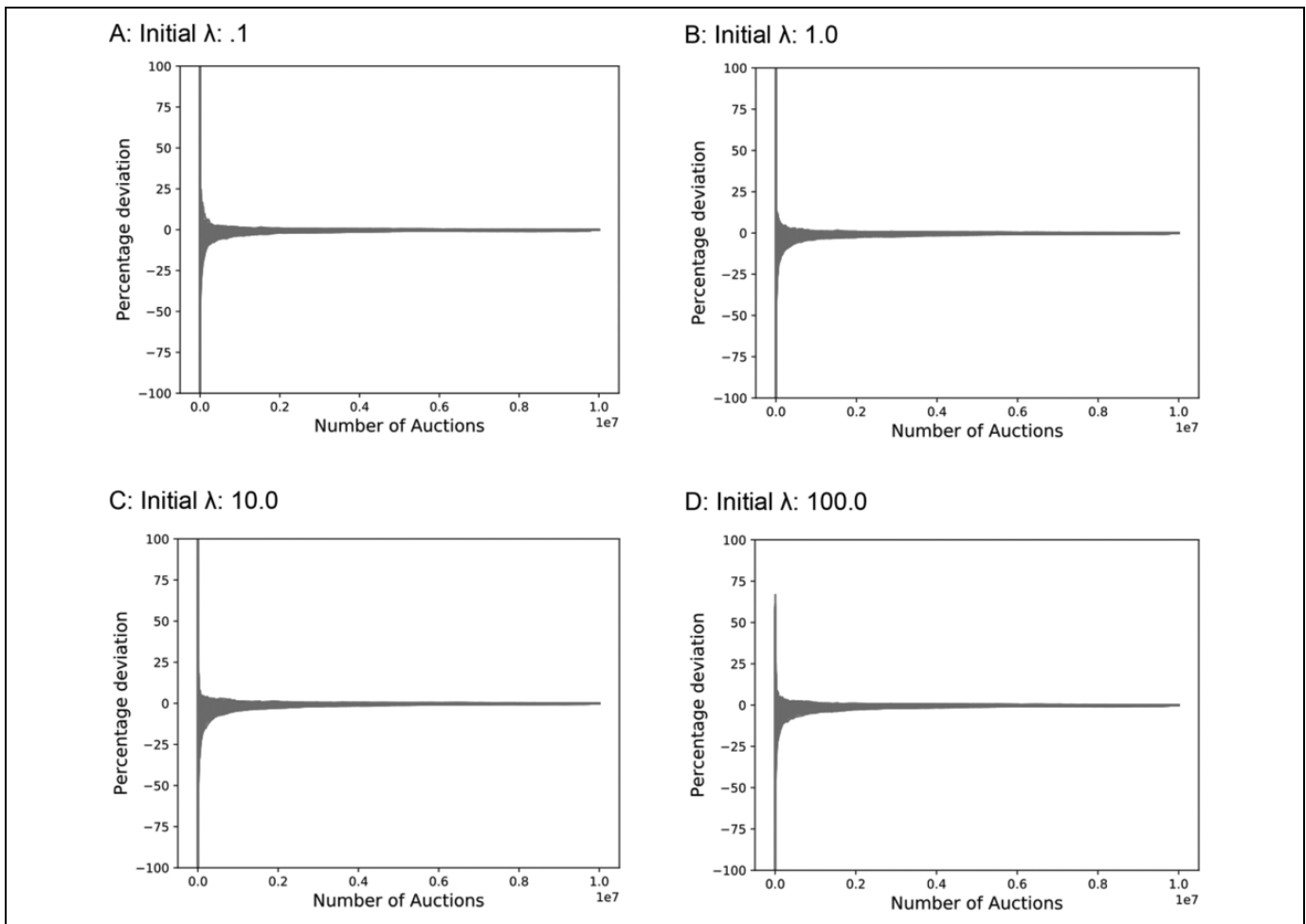
Figure 3. Convergence to λ^* with different starting points.
Notes: These plots show the percentage deviation of λ_n from λ^* across 100 simulations based on a series of starting points, $\lambda_0 \in \{.1, 1.0, 10.0, 100.0\}$. The consistent, rapid convergence of λ_n to λ^* is evident.

BG applies only when a monetary value can be assigned to each impression opportunity. Thus, we assume that the value assigned to each simulated opportunity is in monetary units. Because BG's goal is to maximize profit, the campaign budget may not be exhausted when there exist too few profitable opportunities. Thus, we measure the total value achieved as the sum of the value obtained from purchased impressions and the budget remaining at the end of the campaign. While we present a single comparison to BG for the sake of parsimony, we have compared the approaches under a wide variety of assumptions on the prevalence of profitable opportunities. We consistently find the proposed algorithm delivers significantly greater value.

Finally, we compare our performance with an algorithm that bids a fixed amount for each opportunity (Fixed-Bid). While such a strategy has been shown to be inferior to available dynamic strategies (Zhang, Yuan, and Wang 2014), it is still commonly enabled by platforms such as Google's ad network (Google 2018b). We provide this comparison to help researchers and advertisers understand the opportunity cost of

employing such an approach. To provide the most conservative possible comparison and a lower bound on this opportunity cost, we bid the optimal fixed bid amount given perfect foresight in each simulation.

Each of these comparisons is made relative to the theoretic upper bound, provided by the oracle with perfect foresight. As described in the "Oracle and Regret" subsection, the oracle has access to all values and competing bids at the start of the campaign, making the value obtained by the oracle an upper bound on the value obtained by any algorithm that has access only to the history of values and observed costs. That is, there exists no online bidding algorithm that can outperform the oracle for a single campaign, much less in expectation. By using the oracle in our comparisons, we convey both the degree to which our approach outperforms available alternatives and the extent to which any potential future algorithm could outperform the proposed strategy.

Figure 4 plots the percentage loss from ORTB1, BG, Fixed-Bid, and our algorithm as compared to the relaxed knapsack with perfect foresight. These plots are based on 100
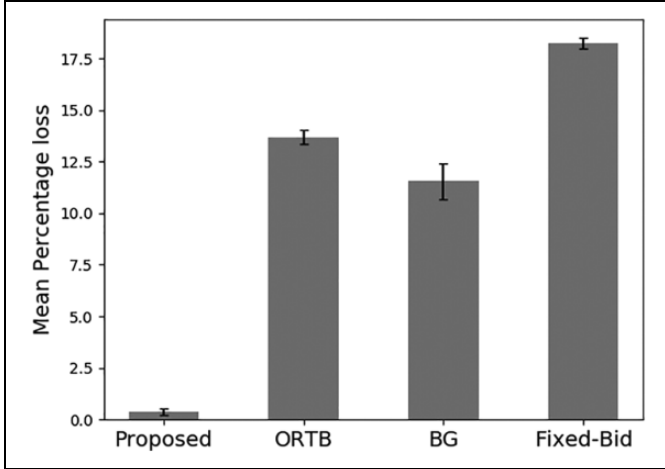
Figure 4. Mean percentage loss: static simulations.

Notes: This plot shows the mean percentage loss from the upper bound for the proposed algorithm, ORTB1, BG, and Fixed-Bid, along with the standard deviations.

simulations, each containing 10 million auctions, with (v, w) sampled randomly as described previously. In each, we initialized our algorithm at $\lambda_0 = 1.0$. Across the simulations, our algorithm delivers 99.63% of the value achieved by oracle and never delivers less than 99.12%. While BG performs admirably, it captures only 88.45% of the value delivered by the relaxed knapsack problem. Notably, it underperforms our approach in all 100 simulations. When provided with the optimal c and λ parameters for each simulation a priori, the performance of ORTB1 is similar to that of BG, capturing 86.3% of the value delivered by the oracle. Fixed-Bid underperforms the rest of the algorithms, delivering only 81.76% of the potential value.

Empirical Performance: DSP Data

We further compare the empirical performance of our algorithm using a large panel data set of real impression values, bids, and outcomes provided by a DSP. The data detail advertising valuation, bidding, and exposures for 990,360 randomly selected users between September 29, 2014, and January 2, 2015. When the data were collected, the collaborating DSP purchased RTB impressions on behalf of 401 advertisers. Each time the DSP receives a bid request, it calculates the value of the impression opportunity for each of the campaigns it manages. This value, referred to as the "Targeting Score," reflects the DSP's belief that a browser will convert for a given advertiser (e.g., visit the advertiser's website or make an online purchase) and is specific to the advertiser and impression opportunity. It is estimated using data provided by the exchange (e.g., IP address and website on which the ad will be served), purchased from third-party providers (e.g., inferred interests), and proprietary to the advertiser (e.g., purchases and browsing on its website). The DSP then decides whether and on behalf of which campaign to submit a bid.

Following bid submission, the DSP records data describing the opportunity for later analysis and reporting. This includes an identifier for the campaign for which the bid was submitted and the calculated value, v_n, that motivated that bid. To assist in evaluating the effectiveness of the bidding strategy, the DSP also records a second campaign identifier and associated targeting score, randomly selected from the set of campaigns it manages. This process is detailed in Figure 5, and the resulting data are stylized in Table 2. As a result, each time a bid is submitted we observe two valuations: one for the campaign for which the DSP submits a bid and a second for a randomly selected campaign.

Because impression costs are observed only when the DSP wins an auction, we restrict our attention to those opportunities. This does not restrict us to the set of impressions won by any given advertiser, as the DSP works for multiple advertisers simultaneously. Instead, for each advertiser, we observe impression values and costs for all auctions won by that advertiser as well as auctions won by other DSP clients for which the focal advertiser's "Targeting Score" was saved. Building on Table 2, we present the resulting data for advertiser A in Table 3.

These real-world data on counterfactual impression costs do not alter the sequence of data revealed to the online algorithms, but they do enrich our comparisons. The data allow us to compare the proposed bidding algorithm with that deployed by the focal DSP. They also establish the oracle's strategy, and therefore regret. Finally, they are used to optimize parameter values for the competing algorithms a priori, providing a conservative test of the proposed approach. In contrast, the impression cost used by the online algorithms is still equal to the observed cost when the algorithm's recommended bid exceeds that value and $.00 otherwise.

For each campaign, we thus compiled a chronologically ordered panel containing the targeting score and impression cost for a subset of opportunities. We focus on the ten campaigns with the most opportunities, the details of which are available in Table 4. Here, the second column contains the total number of opportunities for which we observe the targeting score, the third column reflects the number of impressions served by the DSP, and the fourth column contains the cost of purchasing those impressions. Note that for each campaign, the number of opportunities is greater than the number of impressions served, as discussed previously.

For each campaign, the data are used to evaluate the performance of the proposed algorithm and the alternatives described in the previous subsection. For the alternative approaches, we again fix all underlying parameters at their a posteriori optimal values while requiring our algorithm to optimize λ as it bids. We set the budget to the total observed cost incurred by the DSP for that campaign, enabling us to compare the total value achieved by the DSP to that obtained through the proposed algorithm and alternatives. This budget is binding because we observe impression values and costs even when an impression was not served.

Figure 6 plots the value of λ^* learned by the proposed algorithm for each of the campaigns. Recall that λ^* is the optimal
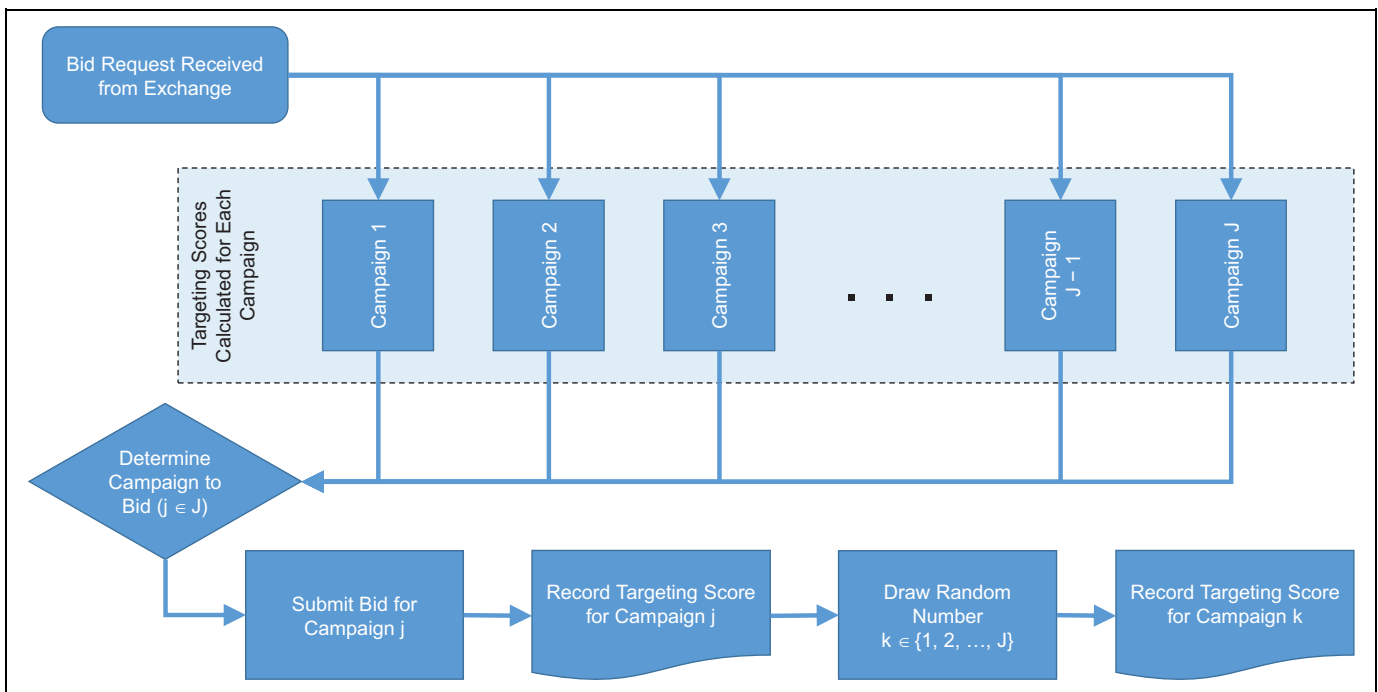
Figure 5. Process by which targeting scores are documented.
Notes: This plot shows the method by which targeting scores are recorded in our data.

Table 2. Stylized DSP Data.

	Bid Campaign			Random Campaign	
Auction	Advertiser	Value	Cost	Advertiser	Value
1	A	.31	$.50	B	.35
2	B	.21	$.75	A	.25
3	C	.37	$.55	B	.32
4	C	.47	$1.25	A	.60
5	A	.54	$.60	C	.32

Table 3. Stylized Campaign Data: Advertiser A.

Auction	Value	Cost	Won
1	.31	$.50	1
2	.25	$.75	0
4	.60	$1.25	0
5	.54	$.60	1

Table 4. DSP Campaign Summary Statistics.

Campaign	Opportunities	Impressions	Cost ($)
1	3,532,510	291,426	554.56
2	1,446,023	999,742	1,324.50
3	1,167,017	732,747	737.56
4	998,791	810,027	1,161.65
5	998,504	40,663	40.15
6	900,926	805,819	556.05
7	892,763	286,204	411.84
8	823,487	414,350	420.16
9	801,778	663,758	822.38
10	717,500	590,846	722.77

Lagrange multiplier in Equation 4 and thus represents the minimum number of conversions expected by the advertiser for every dollar expenditure. For these campaigns, the value of λ^* varies greatly between .22 and 14.33, with an average of 6.2.

As before, the efficacy of each algorithm is measured as the percentage regret relative to the theoretic upper bound. Figure 7 plots the regret produced by the DSP's bidding strategy, ORTB1, BG, Fixed-Bid, and our algorithm for each of the ten campaigns. Across campaigns, the proposed algorithm produces regret of just 1.73% on average. That is, the proposed algorithm delivers on average 98.27% of the value achieved by the oracle. In contrast, ORTB1, BG, and Fixed-Bid deliver 80.46%, 84.03%, and 81.86% of this value on average. The DSP's current algorithm performed the worst, delivering on average only 55.84% of the oracle's value. Thus, across ten real-world campaigns, our algorithm performed within 2% of the theoretic maximum and outperformed the best available alternative by more than 14%. It is worth noting that the results indicate that the DSP's loss in campaign value attributable to inefficient bidding is between 12.4% and 84% and varies greatly across campaigns.

Estimating Valuation While Learning the Bidding Policy

To this point, we have assumed that the advertiser accurately forecasts the impression value for each auction. While the

proposed algorithm consistently outperforms available alternatives when such forecasts are inaccurate (for further details, see Web Appendix B), total campaign value does decline. In this section, we show how the proposed bidding algorithm can be combined with Thompson sampling to simultaneously estimate impression values and learn the optimal bidding policy. For expositional simplicity, we focus on the static algorithm presented in the "Problem Setup and Proposed Algorithm" section. However, because the distribution of impression values may change as the advertiser updates its valuation function, the dynamic algorithm presented in Web Appendix D may offer improved performance. The underlying mechanism of estimating impression values does not depend on whether the bidding algorithm presented in the "Problem Setup" section or Web

Appendix D is used, and the results presented here are robust to either.

Multi-Arm Bandit Formulation and Algorithm

The advertiser's problem can be framed as a contextual multi-arm bandit. Each auction n is accompanied by a d-dimensional attribute vector z_n, describing the context and user characteristics. These attributes can represent features of the publisher's web page (e.g., content type), the impression (e.g., allowable size), and the user (e.g., past purchase behavior). We assume that z_n is i.i.d. across auctions. Conditional on z_n, the advertiser selects an arm to play (i.e., the bid amount). The advertiser's impression is served when its bid is greater than all others (i.e., $b_n > w_n$). Following the auction, the advertiser receives a reward Y_n, equal to one if the user converts and zero otherwise. Conversion is defined by the advertiser, generally measured as either a page view or a purchase. For other applications of the contextual multi-arm bandit to online display advertising, see Schwartz, Bradlow, and Fader (2017) and Waisman et al. (2019).

This setting deviates from the traditional contextual bandit problem in several ways. First, the advertiser's actions (i.e., bids) are continuous, not discrete and finite. Second, at each auction, the advertiser incurs a cost, equal to the highest competing bid if an impression is served and zero otherwise. Because of the budget constraint, this cost induces interdependence across auctions. Third, rewards are correlated across arms because multiple bids can produce the same reward. All bids exceeding w_n produce the same reward, as do all bids below w_n. Finally, rewards are correlated across contexts, as similar attribute vectors may have similar reward distributions. The first two challenges are addressed by the proposed bidding algorithm.
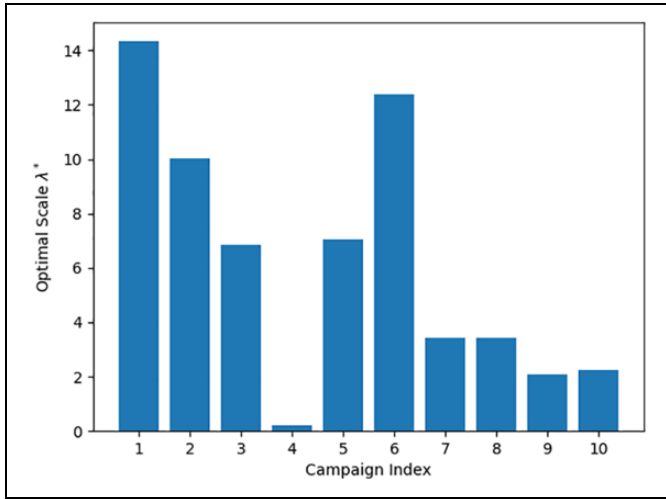

Figure 6. DSP data: λ^*.
Notes: This plot shows the learned values of λ^* for each of the DSP campaigns.
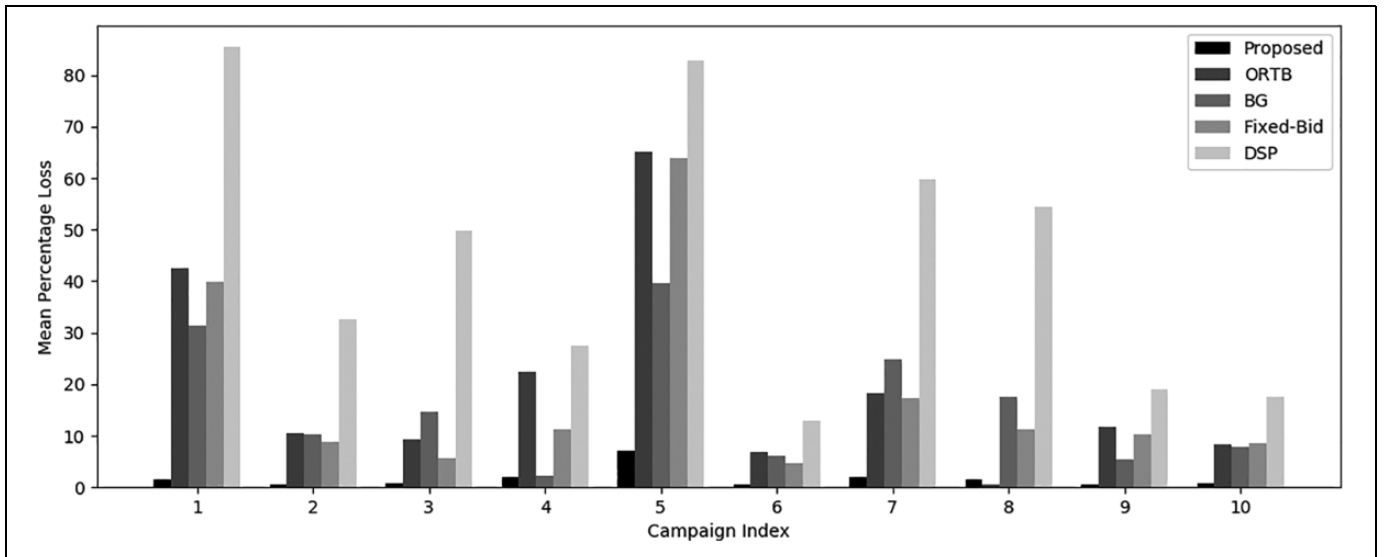

Figure 7. Mean percentage loss: empirical data.
Notes: This plot shows the mean percentage loss for ten campaigns from the upper bound for the proposed algorithm, ORTB1, Fixed-Bid, BG, and the DSP's bidding strategy.

We address the second two challenges, correlated rewards across arms and contexts, using Thompson sampling, similar to Schwartz, Bradlow, and Fader (2017). Thompson sampling, a Bayesian heuristic to choose actions in a multi-arm bandit setting, parameterizes the reward distribution and learns the associated parameters through repeated interactions. At each auction n, the advertiser uses past observations to form a belief distribution over the parameters of the reward distribution. A random sample is drawn from this distribution, and the advertiser takes the action maximizing the expected reward, conditional on the sample drawn and observed vector z_n. Following each auction, the advertiser updates its belief distribution based on the reward.

We model the probability of a conversion associated with auction n as a function of the attribute vector, z_n, and whether an impression is served, $1(b_n > w_n)$. For expositional simplicity and following Schwartz, Bradlow, and Fader (2017) and Waisman et al. (2019), we specify this relationship via a generalized linear model with additively separable terms, though note that any binary classification model will suffice. Letting h^{-1} represent the link function, the probability of a conversion associated with auction n is

$$p_n = h^{-1}\left[\beta_z^T z_n + \beta_{Imp} 1(b_n > w_n)\right], \qquad (11)$$

where β_z specifies the relative weight of each element of the attribute vector z_n and β_{Imp} represents the change in the linear predictor resulting from an impression. The advertiser's valuation of each opportunity, v_n, is then equal to the change in the probability of conversion resulting from an impression,

$$v_n = h^{-1}\left(\beta_z^T z_n + \beta_{Imp}\right) - h^{-1}\left(\beta_z^T z_n\right). \qquad (12)$$

As this measure is based on an impression's incremental impact on the outcome of interest, it overcomes many of the issues associated with valuing impressions based on the user's absolute conversion probability, a common practice as discussed in the "RTB Strategies" subsection. The uncertainty in valuation arises from the uncertainty around the parameters $\beta = [\beta_z, \beta_{Imp}]$.

We now propose an algorithm using Thompson sampling to estimate these parameters. At the start of the campaign, we assume a normal prior for the parameters, such that $\beta \sim \mathcal{N}(\bar{\beta}, \Sigma)$. This prior distribution serves as the belief distribution at the first auction. In our simulations, we set $\bar{\beta}$ equal to a vector of zeros and Σ equal to the identity matrix. In practice, advertisers might select informed priors based on estimates from prior campaigns. Following each auction, the advertiser will update its belief distribution, and we denote the belief distribution at auction n by $\mathbb{P}_n(\beta)$.

Each auction begins with the advertiser receiving the attribute vector z_n. It then draws a random sample from the belief distribution $\mathbb{P}_n(\beta)$, calculates the value of serving an impression v_n using Equation 12, and computes the bid amount b_n using Equation 6. Following the auction, the advertiser

observes whether an impression is served $1(b_n > w_n)$ and whether a conversion occurred Y_n. The advertiser then updates the belief distribution using Bayes' rule

$$\mathbb{P}_{n+1}(\beta) = \mathbb{P}_n(\beta) p_n^{Y_n}(1 - p_n)^{(1 - Y_n)}, \qquad (13)$$

where p_n is computed using Equation 11.

We now summarize the algorithm for completeness. The algorithm takes as input the total number of auctions N, campaign budget B, learning rate μ, initial scale for bidding λ_1, and prior distribution $\mathbb{P}_1(\beta)$. Recalling that $\rho = B/N$, for each auction $n = 1, 2, \ldots, N$ the advertiser

1. Receives attribute vector z_n,
2. Draws a random sample β from $\mathbb{P}_n(\beta)$,
3. Computes $v_n = h^{-1}\left(\beta_z^T z_n + \beta_{Imp}\right) - h^{-1}\left(\beta_z^T z_n\right)$,
4. Submits bid $b_n = v_n/\lambda_n$,
5. Observes whether an impression is served $1(b_n > w_n)$ and conversion Y_n,
6. Computes the probability of conversion $p_n = h^{-1}\left[\beta_z^T z_n + \beta_{Imp} 1 (b_n > w_n)\right]$,
7. Updates its belief distribution $\mathbb{P}_{n+1}(\beta) = \mathbb{P}_n(\beta) p_n^{Y_n}(1 - p_n)^{(1 - Y_n)}$, and
8. Sets $c_n = w_n 1(b_n > w_n)$ and updates λ:

$$\lambda_{n+1} - \frac{1}{n}\sum_{i=1}^{n}\lambda_i = -\frac{1}{\mu}\left(\rho - \frac{1}{n}\sum_{i=1}^{n}c_i\right)$$

The multi-arm bandit formalization of the advertiser's problem and its solution presented previously provide several advantages. First, Thompson sampling has been shown to yield near-optimal asymptotic regret bounds (Agrawal and Goyal 2013) and state-of-the-art performance with finite samples (Chapelle and Li 2011; Kaufmann, Korda, and Munos 2012). This is reflected in the empirical performance of the proposed algorithm, as we demonstrate in the next subsection. Second, the proposed algorithm can be extended to the case where the advertiser chooses between multiple creatives in addition to the bid for each auction. Here, the probability of conversion can be modeled for each creative using creative-specific coefficients β. We refer the reader to Schwartz, Bradlow, and Fader (2017), where such an approach is applied to derive the optimal allocation of impressions across a set of publishers. Third, advertisers can increase the algorithm's computational efficiency by updating the belief distribution (i.e., Step 7) only after a batch of auctions. Previously, we have assumed a batch size of one for simplicity. Finally, Thompson sampling induces exogenous variation in the advertiser's bidding behavior mitigating estimation biases associated with targeting and competitive auction dynamics. For a discussion of how such exogenous variation in bidding can produce unbiased estimates of display advertising response, see Lewis and Wong (2018). Thus, the proposed approach also benefits practitioners and empirical researchers interested in display advertising response.

Table 5. iPinYou Auction Attributes.

Bid ID	User tags	Long-term interest/news
Time stamp	Demographic/gender/male	Long-term interest/education
Log type	Demographic/gender/female	Long-term interest/automobile
iPinYou ID	In-market/3c product	Long-term interest/real estate
User agent	In-market/appliances	Long-term interest/information technology
IP	In-market/clothing, shoes, bags	Long-term interest/electronic games
Region	In-market/Beauty, Personal Care	Long-term interest/fashion
City	In-market/household, home improvement	Long-term interest/entertainment
Ad exchange	In-market/infant, mom products	Long-term interest/luxury
Domain	In-market/sports item	Long-term interest/home and lifestyle
URL	In-market/outdoor	Long-term interest/health
URL ID	In-market/health care products	Long-term interest/food
Ad slot ID	In-market/luxury	Long-term interest/divine
Ad slot width	In-market/real estate	Long-term interest/motherhood, parenting
Ad slot height	In-market/automobile	Long-term interest/sports
Ad slot visibility	In-market/finance	Long-term interest/travel, outdoors
Ad slot format	In-market/travel	Long-term interest/social
Ad slot reserve price	In-market/education	Long-term interest/art, photography, design
Creative ID	In-market/service	Long-term interest/online literature
Bidding price	In-market/electronic game	Long-term interest/3c
Paying price	In-market/book	Long-term interest/culture
Key page URL	In-market/medicine	Long-term interest/sex
Advertiser ID	In-market/food, drink	

Regret Analysis

To explore the relative impact of uncertainty over impression values and bidding policies, we use data from iPinYou, a Chinese DSP (Liao et al. 2014). The data contain information on RTB auctions for 15 display advertising campaigns managed by iPinYou from March 6–17, 2013, and October 19–27, 2013. Each auction is associated with contextual information, such as the publisher's website, a time stamp, and a measure of ad slot visibility. iPinYou augmented these contextual attributes with user specific demographics (e.g., gender, location) as well as marketing information (e.g., customer segment assignments). Table 5 provides the complete list of attributes. Finally, the data include the bids submitted by iPinYou, the costs they incurred, and user feedback, measured by clicks on the display ad impressions.

We begin by constructing a binary classification model to estimate the probability a user clicks on an impression, given the contextual and user attributes. Following Amar and Renegar (2018), we opt for a binary logistic regression model; remove unique or nearly unique features such as iPinYou ID and URL; and dummy-code categorical variables such as device type, gender, and consumer segment. In our simulations, the resulting estimates serve as the ground truth for $\beta = [\beta_z, \beta_{\text{Imp}}]$, though we assume these are unknown to the advertiser.

Similar to the previous section, we then simulate 100 campaigns, each containing 10 million auctions. For each simulated auction, the contextual and user attributes are drawn randomly from their empirical distribution in the iPinYou data set. At the beginning of each campaign, we set $\bar{\beta}$ equal to a vector of zeros and Σ equal to the identity matrix, such that each coefficient

has a standard normal prior that is independent of the other coefficients. We assume that the advertiser models user conversion as a Bernoulli random variable, meaning h^{-1} is the logit link function. In each simulation, the advertiser follows the algorithm outlined in the previous subsection.

We again compare our algorithm's performance to that of ORTB1, BG, and Fixed-Bid based on total regret. As in the "Empirical Performance" section, regret is defined as the number of conversions produced by an algorithm relative to the theoretic upper bound. We again seed competing algorithms with the optimal bidding parameters for each simulation, providing the most conservative possible test of our proposal. However, we do require that each algorithm learn impression values in real time using Steps 1–7 of the process outlined previously.

To compare the relative impacts of uncertainty over impression values and the bidding policy, we provide two upper bounds. The first, which we refer to as the TrueOracle, is produced by an algorithm that has perfect foresight into both the true valuations and the highest competing bid for each auction. This oracle is identical to that in the "Empirical Performance" section. The second, which we call the Bidding-Oracle, has perfect foresight into the highest competing bids w_n but is constrained to forecast impression values v_n using the same belief distribution $\mathbb{P}_n(\beta)$ as the algorithm to which it is being compared. Note that because this belief distribution is determined by the set of impressions purchased and each algorithm purchases a unique set of impressions in each simulation, there is a unique BiddingOracle for each algorithm in each simulation. In contrast, the TrueOracle has perfect foresight into all values, resulting in one TrueOracle for each simulation.

By comparing an algorithm's performance with each oracle, we can identify the portions of regret attributable to uncertainty
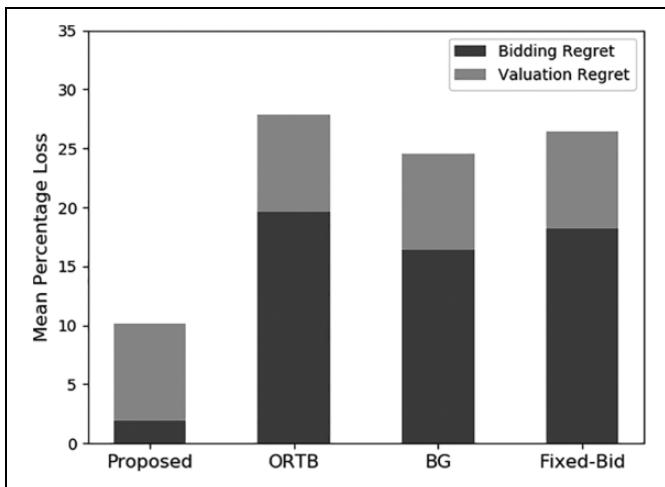
Figure 8. Mean percentage loss: uncertain valuations and bidding policy.

Notes: This plot shows the regret attributable to uncertainty over the bidding policy, uncertainty over valuations, and the total regret for each of the algorithms.

over the bidding policy and uncertainty over impression values. Because the BiddingOracle and the algorithm to which it is being compared share a belief distribution $\mathbb{P}_n(\beta)$ at each auction, they assign identical values to each opportunity. As a result, any difference in performance, and thus regret, stems entirely from the difference in bidding behavior. At the same time, the difference in performance between the TrueOracle and each BiddingOracle stems from uncertainty over valuation, as they differ only in the accuracy with which they forecast v_n. This also implies that the TrueOracle's performance is an upper bound on that of the BiddingOracle.

Figure 8 presents the total regret produced by each algorithm over the series of 100 simulated campaigns, each containing 10 million auctions. The average regret attributable to uncertainty over the bidding policy (i.e., the regret relative to the BiddingOracle) is less than 2% for the proposed algorithm, but 19.62%, 16.37%, and 18.24% for ORTB1, BG, and Fixed-Bid, respectively. The regret due to uncertain valuations (i.e., the regret of each BiddingOracle relative to the TrueOracle) is approximately 8.2% for each of the bidding algorithms. This similarity across algorithms is largely a result of using the same Thompson sampling technique to estimate $\mathbb{P}_n(\beta)$ and the large number of auctions per simulation.

Finally, it is worth noting that the total regret of the proposed algorithm, 10.1%, is less than the minimum regret attributable to the bidding policy of any other algorithm, 16.37% for BG. This indicates that the proposed algorithm consistently outperforms available alternatives, even when it is asked to simultaneously estimate impression values and learn the optimal bidding policy while the alternatives are provided with perfect foresight into impression values and seeded with optimal bidding parameters. This underscores the importance of employing an efficient bidding strategy, even when impression values are uncertain.

Conclusion

In this article, we introduced a near-optimal bidding algorithm for firms purchasing advertising through real-time auctions. This is now the dominant distribution channel for internet display advertising and a growing funding model for addressable television and radio. Because the impressions are sold as a user loads a web page, these exchanges allow advertisers to target each impression by individual, publisher, and time. With this flexibility, U.S. advertisers are expected to purchase $26 billion of display advertising through RTB exchanges in 2020, representing nearly half of all display ad spending (Hoelzel and Ballve 2015).

For the advertisers, the speed and volume of these auctions mandate the use of automated algorithms. By developing a near-optimal bidding algorithm, our work follows a long history of marketing research into the effective design and use of automated programs to help marketers allocate scarce resources (e.g., Leeflang and Wittink 2000; Little 1970; Rust 1986). It is also in line with more recent marketing research focused on the development of programmatic solutions to optimize digital advertising, (e.g., Paulson, Luo, and James 2018; Schwartz, Bradlow, and Fader 2017; Skiera and Abou Nabout 2013; Waisman et al. 2019). Finally, we contribute to the growing literature seeking solutions to the online knapsack problem. We are able to prove near-optimal average and total regret without assuming that the advertiser has any knowledge of the distributions describing future impression values and costs. To the best of our knowledge, this is unique among algorithmic solutions to the online knapsack.

In addition to being easily implemented, the proposed algorithm satisfies the real-world constraints of the RTB ecosystem. It is computationally efficient and capable of processing the volume and velocity of auctions while meeting the rapid response times required by the advertising exchanges. It is guaranteed to converge to the optimal strategy and achieves zero-regret based only on the sequence of historical auction costs incurred by the focal advertiser. Importantly, this algorithm allows advertisers to learn the optimal strategy in the course of a campaign without assuming or directly estimating the distribution describing impression values and costs. This mitigates a formidable challenge identified in the existing literature on bidding strategies. It also addresses competitive concerns, as following the proposed strategy constitutes a competitive equilibrium, allowing advertisers to be agnostic with respect to competitive strategies. Finally, we show how the bidding algorithm can be combined with Thompson sampling to simultaneously estimate impression values and learn the optimal bidding policy.

Across a series of simulations and real-world tests, the proposed algorithm significantly outperformed existing alternatives. Across 100 simulations containing 10 million auctions each, we found that our approach consistently delivered 11% more value than available alternatives and 99% of the value possible with perfect foresight. Using data from ten real-world campaigns, the proposed algorithm outperformed the next best

approach by 14% and delivered 98% of the value possible with perfect foresight. Thus, any future solution to the problem can, at best, outperform the proposed algorithm by 2%. When asked to simultaneously estimate both impression values and the optimal bidding policy, the proposed algorithm continued to outperform available alternatives, even when these alternatives were provided accurate impression values and seeded with a priori optimal bidding parameters. This provides strong evidence of the algorithm's potential impact, even when advertisers are uncertain with respect to impression values.

We expect this work to have a significant and increasing impact on how advertisers purchase media through the real-time auctions. The proposed approach outperforms the best available alternative by 11%, even when the parameters underlying competing algorithms were optimized a priori. Reframing this in terms of campaign spend, a cost-conscious advertiser that unilaterally deviates from this strategy to ours can expect to deliver equivalent campaign value while spending just 89% of the current budget. In many instances, the differences may be even more dramatic. Many large, sophisticated DSPs regularly recommend and enable advertiser bidding strategies that deviate even more significantly from the optimal (e.g., fixed-bid; Google 2018b). Partially in response to perceived inefficiencies and intermediary costs, advertisers have begun shifting programmatic media buying capabilities in-house (Wolfe 2017). Without a solid bidding strategy, these firms risk eroding much, if not all, of the potential cost savings. For advertisers, this article introduces a near-optimal bidding strategy, highlights the opportunity cost associated with alternative strategies, and presents a method to simultaneously estimate impression values based on their incremental impact. Looking forward, these implications extend beyond internet display advertising, as RTB is increasingly considered a potential funding model for addressable television and radio.

This work is not without limitations. We provide a near-optimal bidding strategy in the form of an implementable learning algorithm for use in sequential, stochastic Vickrey auctions. This is a common auction protocol in RTB for internet display advertising. However, ad exchanges and other intermediaries continually experiment with auction protocols. Soft floors and first-price auctions are two permutations currently receiving attention. Soft floors set a price threshold, unknown by the advertiser, below which the auction follows first-price protocols. While Zeithammer (2016) shows that this is revenue neutral at best, it is still used occasionally. Because of "header bidding," many exchanges now operate as intermediaries rather than the final arbiter of impression placement. As a result, some exchanges have adopted first-price auction protocols, while some publishers, which now operate the terminal auctions, have adopted second-price auctions. So long as the terminal auction follows second-price rules and preceding auctions, if any, follow first-price rules, the proposed algorithm will perform as described. However, these industry dynamics highlight the need for more research on optimal bidding strategies, especially when the auctions are sequential and the protocols are mixed.

ORCID iDs

Srinivas Tunuguntla ⓘ https://orcid.org/0000-0001-9580-3970
Paul R. Hoban ⓘ https://orcid.org/0000-0002-9868-6674

References

Agrawal, Shipra and Navin Goyal (2013), "Further Optimal Regret Bounds for Thompson Sampling," *Artificial Intelligence and Statistics*, 31, 99–107.

Agrawal, Shipra, Zizhuo Wang, and Yinyu Ye (2014), "A Dynamic Near-Optimal Algorithm for Online Linear Programming," *Operations Research*, 62 (4), 876–90.

Amar, Jonathan and Nicholas Renegar (2018), "The Second-Price Knapsack Problem: Near-Optimal Real Time Bidding in Internet Advertisement," https://arxiv.org/abs/1810.10661.

Anderson, Erin, Leonard M. Lodish, and Barton A. Weitz (1987), "Resource Allocation Behavior in Conventional Channels," *Journal of Marketing Research*, 24 (1), 85–97.

Balseiro, Santiago R., Omar Besbes, and Gabriel Y. Weintraub (2015), "Repeated Auctions with Budgets in ad Exchanges: Approximations and Design," *Management Science*, 61 (4), 864–84.

Balseiro, Santiago R. and Yonatan Gur (2019), "Learning in Repeated Auctions with Budgets: Regret Minimization and Equilibrium," *Management Science*, 65 (9), 3952–68.

Cai, Han, Kan Ren, Weinan Zhang, Kleanthis Malialis, Jun Wang, Yong Yu, et al. (2017), "Real-Time Bidding by Reinforcement Learning in Display Advertising," in *Proceedings of the Tenth ACM International Conference on Web Search and Data Mining*. Cambridge, UK: Association for Computing Machinery, 661–70.

Chakrabarty, Deeparnab, Yunhong Zhou, and Rajan Lukose (2007), "Budget Constrained Bidding in Keyword Auctions and Online Knapsack Problems," in *WWW2007, Workshop on Sponsored Search Auctions*, Vol. 10, https://archive.thewebconf.org/www2008/papers/pdf/p1243-zhou.pdf.

Chapelle, Olivier and Lihong Li (2011), "An Empirical Evaluation of Thompson Sampling," *Advances in Neural Information Processing Systems*, 24, 2249–57.

Choi, Hana, Carl F. Mela, Santiago R. Balseiro, and Adam Leary (2020), "Online Display Advertising Markets: A Literature Review and Future Directions," *Information Systems Research*, 31 (2), 556–75.

Cormen, Thomas H., Charles E. Leiserson, Ronald L. Rivest, and Clifford Stein (2009), *Introduction to Algorithms*, 3rd ed. Boston: MIT Press.

Danaher, Peter J. (1991), "Optimizing Response Functions of Media Exposure Distributions," *Journal of the Operational Research Society*, 42 (7), 537–42.

Danaher, Peter J. and Harald. J. van Heerde (2018), "Delusion in Attribution: Caveats in Using Attribution for Multimedia Budget Allocation," *Journal of Marketing Research*, 55 (5), 667–85.

Dean, Brian C., Michel X. Goemans, and Jan Vondrák (2008), "Approximating the Stochastic Knapsack Problem: The Benefit of Adaptivity," *Mathematics of Operations Research*, 33 (4), 945–64.

Dinner, Isaac M., Harald J. Van Heerde, and Scott A. Neslin (2014), "Driving Online and Offline Sales: The Cross-Channel Effects of Traditional, Online Display, and Paid Search Advertising," *Journal of Marketing Research*, 50 (5), 527–45.

eMarketer (2017), "US Digital Display Advertising Will Continue to Climb in 2018," (December 26), https://www.emarketer.com/content/us-advertisers-will-spend-nearly-48-billion-on-digital-display-ads-in-2018-emarketer-estimates.

Forbes (2015), "As Brands Turn to Digital Advertising to Reach the Right Audience, Focus on Validation Is Increasing," (May 5), https://www.forbes.com/sites/forbespr/2015/05/05/as-brands-turn-to-digital-advertising-to-reach-the-right-audience-focus-on-validation-is-increasing/?sh=d2f3b62272c7.

Google (2018a), "About Campaign Budgets," (accessed November 5, 2018), https://support.google.com/google-ads/answer/6385083.

Google (2018b), "Understanding Bidding Basics," (accessed May 24, 2018), https://support.google.com/adwords/answer/2459326?.

Gordon, Brett R., Florian Zettelmeyer, Neha Bhargava, and Dan. Chapsky (2019), "A Comparison of Approaches to Advertising Measurement: Evidence from Big Field Experiments at Facebook," *Marketing Science*, 38 (2), 193–225.

Heise, Marc, Nadia Abou Nabout, and Bernd Skiera (2016), "An Analysis of the Profitability of Even Pacing in Real-Time Bidding," working paper, Goethe University.

Hoelzel, Mark and Marcello Ballve (2015), "Programmatic Advertising: Mobile, Video, and Real-Time Bidding Drive Growth in Programmatic," *Business Insider* (March 26), https://www.businessinsider.com/programmatic-advertising-report-mobile-video-and-real-time-bidding-drive-growth-in-programmatic-2015-3.

Joe, Ryan (2018), "When It Comes to Addressable TV, AT&T Has the Scale and Verizon Has the Speed," *Ad Exchanger* (May 21), https://adexchanger.com/tv-2/when-it-comes-to-addressable-tv-att-has-the-scale-and-verizon-has-the-speed/.

Kaufmann, Emilie, Nathaniel Korda, and Rémy Munos (2012), "Thompson Sampling: An Asymptotically Optimal Finite-Time Analysis," in *Algorithmic Learning Theory*, Nader H. Bshouty, Gilles Stoltz, Nicolas Vayatis and Thomas Zeugmann, eds. Berlin: Springer, 199–213.

Kleywegt, Anton J. and Jason D. Papastavrou (1998), "The Dynamic and Stochastic Knapsack Problem," *Operations Research*, 46 (1), 17–35.

Lee, Kuang-Chih, Ali Jalali, and Ali Dasdan (2013), "Real Time Bid Optimization with Smooth Budget Delivery in Online Advertising," in *Proceedings of the Seventh International Workshop on Data Mining for Online Advertising*. Chicago: Association for Computing Machinery, 1–9.

Leeflang, Peter S. and Dick R. Wittink (2000), "Building Models for Marketing Decisions: Past, Present and Future," *International Journal of Research in Marketing*, 17 (2/3), 105–26.

Lewis, Randall A and Jeffrey Wong (2018), "Incrementality Bidding & Attribution," *SSRN* (March 12), https://papers.ssrn.com/sol3/papers.cfm?abstract_id=3129350.

Liao, Hairen, Lingxiao Peng, Zhenchuan Liu, and Xuehua Shen (2014), "iPinYou Global RTB Bidding Algorithm Competition Dataset," *Proceedings of the Eighth International Workshop on Data Mining for Online Advertising*. New York: Association for Computing Machinery, 1–6.

Little, John D. (1970), "Models and Managers: The Concept of a Decision Calculus," *Management Science*, 16 (8), B466–85.

Liyakasa, Kelly (2015), "Dish Opens a Programmatic Exchange, Enables RTB," *Ad Exchanger* (October 26), https://adexchanger.com/digital-tv/dish-opens-a-programmatic-exchange-enables-rtb/.

Lueker, George S. (1998), "Average-Case Analysis of Off-Line and On-Line Knapsack Problems," *Journal of Algorithms*, 29 (2), 277–305.

Mantrala, Murali K., Prabhakant Sinha, and Andris A. Zoltners (1992), "Impact of Resource Allocation Rules on Marketing Investment-level Decisions and Profitability," *Journal of Marketing Research*, 29 (2), 162–75.

Marchetti-Spaccamela, Alberto and Carlo Vercellis (1995), "Stochastic On-Line Knapsack Problems," *Mathematical Programming*, 68 (1–3), 73–104.

OpenX (2019), "Identifying First-Price Auctions," (accessed November 1, 2019), https://docs.openx.com/Content/demandpartners/first-price-auctions.html#understanding-first-price.

Pasupathy, Raghu and Sujin Kim (2011), "The Stochastic Root-Finding Problem: Overview, Solutions, and Open Questions," *ACM Transactions on Modeling and Computer Simulation*, 21 (3), 19.

Paulson, Courtney, Lan Luo, and Gareth M. James (2018), "Efficient Large-Scale Internet Media Selection Optimization for Online Display Advertising," *Journal of Marketing Research*, 55 (4), 489–506.

PubMatic (2017), "Understanding Auction Dynamics," (accessed November 2, 2019), https://pubmatic.com/wp-content/uploads/2017/08/PubMatic-UnderstandingAuctionDynamics.pdf.

Rust, Roland T. (1986), *Advertising Media Models: A Practical Guide*. New York: The Free Press.

Schwartz, Eric M., Eric T. Bradlow, and Peter S. Fader (2017), "Customer Acquisition Via Display Advertising Using Multi-Armed Bandit Experiments," *Marketing Science*, 36 (4), 500–522.

Shen, Jianqiang, Burkay Orten, Sahin C. Geyik, Daniel Liu, Shahriar Shariat, and Fang Bian (2015) "From 0.5 Million to 2.5 Million: Efficiently Scaling Up Real-Time Bidding," in *2015 IEEE International Conference on Data Mining (ICDM)*. Piscataway, NJ: Institute of Electrical and Electronics Engineers, 973–78.

Silverman, D. (2019), "2018 IAB Internet Ad Revenue Full Year Report," technical report, PwC Advisory Services (May), https://

www.iab.com/wp-content/uploads/2019/05/Full-Year-2018-IAB-Internet-Advertising-Revenue-Report.pdf.

Skiera, Bernd and Nadia Abou Nabout (2013), "Practice Prize Paper—PROSAD: A Bidding Decision Support System for Profit Optimizing Search Engine Advertising," *Marketing Science*, 32 (2), 213–20.

Vickrey, William (1961), "Counterspeculation, Auctions, and Competitive Sealed Tenders," *Journal of Finance*, 16 (1), 8–37.

Waisman, Caio, Harikesh S. Nair, Carlos Carrion, and Nan Xu (2019), "Online Inference for Advertising Auctions," (August 22), https://arxiv.org/abs/1908.08600.

Wolfe, John (2017), "Marketers Expanding In-House Capabilities for Programmatic Buying: ANA Study," press release, ANA (December 18), http://www.ana.net/content/show/id/47123.

Zawadziński, Maciej (2018), "Waterfalling, Header Bidding and New Auction Dynamics," Clearcode (accessed October 30, 2019), https://clearcode.cc/blog/sequential-auctions-header-bidding-first-price-second-price-auctions/.

Zeithammer, Robert (2016), "The Futility of Soft Floor Auctions," working paper.

Zhang, Weinan, Shuai Yuan, and Jun Wang (2014), "Optimal Real-Time Bidding for Display Advertising," in *Proceedings of the 20th ACM SIGKDD International Conference on Knowledge Discovery and Data Mining*. New York: Association for Computing Machinery, 1077–86.

Zhang, Weinan, Tianming Du, and Jun Wang (2016), "Deep Learning over Multi-Field Categorical Data," in *European Conference on Information Retrieval*. Padua, Italy: Springer, 45–57.