Robust Stackelberg buyers in repeated auctions

Clément Calauzènes
Criteo AI Lab

Thomas Nedelec
Criteo AI Lab
ENS Paris Saclay

Vianney Perchet
CREST, Ensae
Criteo AI Lab

Noureddine El Karoui
Criteo AI Lab
UC, Berkeley

Abstract

We consider the practical and classical setting where the seller is using an exploration stage to learn the value distributions of the bidders before running a revenue-maximizing auction in a exploitation phase. In this two-stage process, we exhibit practical, simple and robust strategies with large utility uplifts for the bidders. We quantify precisely the seller revenue against non-discounted buyers, complementing recent studies that had focused on impatient/heavily discounted buyers. We also prove the robustness of these shading strategies to sample approximation error of the seller, to bidder's approximation error of the competition and to possible change of the mechanisms.

Introduction

Repeated auctions play an important role in modern economics as they are widely used in practice to sell e.g. electrical power or digital goods such as ad placements on big online platforms. Understanding the precise interactions between the buyers and the sellers in these auctions is key to assess the balance of power on big online platforms. In practice, most of the online auctioneers are using tools at the intersection of classical auction literature Myerson (1981) and statistical learning theory. They take advantage of the enormous amount of data they gather on the behavior of the buyers to learn optimal auctions and maximize revenue for the platforms.

Myerson showed how to design an incentive-compatible revenue-maximizing auction once the seller knows the value distributions of the buyers Myerson (1981). If the seller has perfect knowledge of these distributions, she can define the allocation and payment rules maximizing her expected revenue.

How does the seller learn these value distributions in practice to create her optimal revenue-maximizing auction ? A very rich line of work Ostrovsky and Schwarz (2011); Cole and Roughgarden (2014); Medina and Mohri (2014); Huang et al. (2018) has assumed that the seller has access to a finite sample of bidders' valuations, coming from past bids in truthful auctions. The game they consider is divided in two stages. The first round consists in several truthful auctions such as a second price auction without reserve or with random reserve where the bidders bid truthfully providing the seller with draws from their value distribution. Then, if the bidders are not strategic in the first stage, the seller can learn an approximation of the revenue-maximizing auction based on these samples and run a revenue-maximizing auction in the second stage.

On big online platforms, the same dominant bidders are actually repeatedly interacting with a seller, billions of times a day in the case of online advertising. This setting has been considered from either the seller's point of view in Amin et al. (2013, 2014); Cesa-Bianchi et al. (2013) or the bidders' standpoint Kanoria and Nazerzadeh (2014); Tang and Zeng (2018); Nedelec et al. (2018, 2019). All these works can be seen as a special instance of a Stackelberg game Dockner et al. (2000) where the bidders know the rules of the mechanism and have the choice of the bid distribution they will disclose to the seller. The main takeaway from these works is that if one of the seller or bidders is extremely dominant in term of *patience* – i.e. longer time horizon of revenue optimization – then this player can get the best revenue/utility to hope for: the payment of a *revenue-maximizing auction* if the seller is strongly dominant, the utility of a 2^{nd}*-price auction with no reserve price* if the bidders are strongly dominant and are all strategic. Yet, these payments have been exhibited in extreme asymptotic cases. In the line of work following from Amin et al. (2013); Mohri and Munoz (2015), the bidder is assumed to be finitely

Proceedings of the 23[rd]International Conference on Artificial Intelligence and Statistics (AISTATS) 2020, Palermo, Italy. PMLR: Volume 108. Copyright 2020 by the author(s).

patient while the seller is infinitely patient. In this strongly unbalanced setting, the seller is able to begin with exploration stages long enough to force the bidder to be truthful, allowing the seller to play the revenue-maximizing auction in the (longer) exploitation phase. On the contrary, if the bidders are infinitely patient and the seller has to update the mechanism in finite time, Tang and Zeng (2018); Nedelec et al. (2018) exhibited optimal strategies that bidders can enforce. The remaining crucial question is: what happens in between these extreme cases, in more realistic conditions?

Our work aims at providing answers to this question by studying the robustness of the bidders' strategic behavior to more realistic conditions:

1. Are these strategic behaviors robust to strategic sellers implementing two-stage processes (exploration/exploitation) such as the selling algorithms described in Amin et al. (2013); Golrezaei et al. (2019)? How is the value shared between seller and bidders in non-asymptotic patience conditions?
2. Are the possible strategic behaviors robust to the seller optimizing the mechanism with a finite number of samples?
3. How does the lack of knowledge of the competition (i.e. other bidders's value distributions) negatively impact the strategic bidder's utility?

1 Framework and contributions

As it is classical in auction theory Krishna (2009), we assume the valuation of a bidder $v \in \mathbb{R}$ is drawn from a specific distribution F (the distribution can be different from one bidder to the other); a bidding strategy is a mapping β from $\mathbb{R}_+$ into $\mathbb{R}_+$ that provides the actual bid $B = \beta(v)$ when the value is v. As a consequence, the distribution of bids F_B is the push-forward of F by β. We assume from now on that the support of F is $[0, b) \subseteq [0, +\infty]$. If $b = +\infty$, we additionally assume that F verifies $1 - F(x) = o(x^{-1})$ to avoid problems with the definition of the optimal reserve price. Unless otherwise noted F is assumed to be regular.

The seller has an estimate of the bidders' bid distribution, typically coming from access to past bids. Then, she uses an approximation of a revenue-maximizing auction based on her estimate of the bid distributions. We will consider that this is a lazy 2^{nd}-price auction Paes Leme et al. (2016) whose reserve price is optimized to maximize the *seller's revenue*. For the lazy 2^{nd}-price auction, the optimal reserve price happens to be the *monopoly price* r^*_β which maximizes the *monopoly revenue*

$$R_B(r) = r(1 - F_B(r))$$

extracted by the seller on the current bidder. We can formalize this as a two-step process Amin et al. (2013); Cole and Roughgarden (2014):

General two-step process with commitment: A general two-step process with commitment is a tuple of the form $P = (G, H, r, F)$ that is defined as follows:

1 – exploration The seller runs a lazy 2^{nd}-price auction. The current bidder faces a competition G. In this step, the potential randomized or deterministic reserve price is denoted by the distribution H.

2 – exploitation The seller runs a lazy 2^{nd}-price auction with reserve price r, possibly personalized. The current bidder faces the same competition G as in the first step.

To make explicit the dependency on the bidder's strategy β on the design of the auction of the second step when optimizing the reserve price based on the observation of the first period, we use the notation $P^\beta = (G, H, r^*_\beta, F)$. The tradeoff between exploration and exploitation from the seller standpoint was introduced in Amin et al. (2013) and refined in Mohri and Munoz (2015); Golrezaei et al. (2019). They introduce a parameter α, $0 \leq \alpha < 1$, to define this trade-off, assuming the ratio of length between the first and the second stage is equal to $\alpha/1-\alpha$. In Amin et al. (2013), they show that if bidders are non-discounted buyers, there must exist a good strategy for them in this mechanism, forcing the seller to suffer a regret linear in the number of auctions. We are interested in solving the Stackelberg game faced by bidders when they know the seller is using this classical mechanism to learn their value distribution. Bidders are the leaders in this framework since they know the mechanism used by the seller and can choose their strategy accordingly.

We assume that the strategic bidder commits to the same strategy β (inducing the same distribution of bids F_B) in both phases. This assumption accounts for the fact that in practice sellers regularly update their priors based on recent past bids, forcing the bidders to commit in the long-term to the bid distribution they want the seller to use as prior. Otherwise, the seller might discover new aspects of the buyer's bid distribution and change their mechanism accordingly more than once as time evolves.

In this framework, the objective of the bidder is to choose a strategy β to maximize a weighted sum of her utility in both phases:

$$U_\alpha(\beta) = \alpha U_1(\beta) + (1 - \alpha) U_2(\beta) \qquad (1)$$

where U_i is the expected utility of the bidder in stage i. The $\alpha \in [0,1]$ quantifies the length of each stage Amin et al. (2013). More precisely, U_1 is the expected utility of a lazy second price without reserve/with random reserve and U_2 is the expected utility of a lazy second price with monopoly reserves corresponding to F_B.

Tang and Zeng (2018); Nedelec et al. (2018) exhibited strategies for the case where the bidders only optimize for the utility of the second phase, i.e. $\alpha = 0$. However, in practice, sellers use an exploration stage to learn the value distribution. Taking care of this exploration phase is of great importance for the bidders since optimizing only the second stage can lead to large loss of utility during the first stage.

In Section 2, we extend the approach of Tang and Zeng (2018); Nedelec et al. (2018) to the two-stage game showing that the strategies which are optimal when the bidder is only optimizing her utility in the second stage are suboptimal when she takes into consideration the exploration stage of the seller. We find in Section 3 the optimal strategy in this framework and study the behavior of this strategy as a function of the length of the exploration stage. To study whether the strategies are robust to the presence of other strategic bidders in the game, we also prove the existence of a Nash equilibrium under certain conditions and we compute the utility of the bidders and the revenue of the seller at this Nash equilibrium when it exists.

Unlike Amin et al. (2013); Golrezaei et al. (2019), we consider the objective function of the bidders to be the expected utility of the two stages instead of the utility computed on a finite number of auctions. Indeed, as there exists a high number of repeated auctions, billions a day in the case of online advertising, optimizing directly the expected utility makes sense from a bidder's standpoint. We extend our results to finite sample sizes in Section 4.1. We finally show the strategies are robust to an estimation of the bidders' competition in Section 4.2 and to a change of mechanism in Section 4.3. It shows their robustness to most of the difficulties bidders practically face and provides concrete solutions to solve these problems.

2 Understanding the strategic reaction of the players

In order to get a good understanding of how the welfare is shared between seller and bidders when α moves from 0 to 1, we need to study the strategies of the different players. First, on the seller's side, she chooses the distribution of reserve prices H faced by the strategic bidder in the first phase. At this point, we assume the seller does not have any knowledge about the bidders (except the support of the value distribution). Hence, she mostly has two choices for H: either no reserve price or a uniform distribution Amin et al. (2014). In the second stage, the seller is assumed welfare-benevolent, i.e. if she has the choice between two reserve prices equivalent in terms of payment, she chooses the lowest one. Hence, in the second stage, the seller chooses $r^*_\beta = \inf \operatorname{argmax}_r R_B(r)$.

From a bidder's standpoint, (Tang and Zeng, 2018, Th. 6.2) exhibits the best response for the extreme case $\alpha = 0$. Unfortunately, this best response induces a complicated optimization problem for the seller: R_B is non-convex with several local optima and the global optimum is reached at a discontinuity of R_B (see Fig. 1). This is especially problematic if sellers are known to optimize reserve prices conditionally on some available context using parametric models such as Deep Neural Networks Dütting et al. (2019) whose fit is optimized via first-order optimization and hence would regularly fail to find the global optimum. This is undesirable for the bidders: in any Stackelberg game, the leader's advantage comes from being able to predict the follower's strategy.

To address this issue, Tang and Zeng (2018); Nedelec et al. (2018) proposed a *thresholding strategy* that is the best response in the restricted class of strategies that ensure R_B to be concave as long as R_X is so, when $\alpha = 0$. We show this strategy can be extended to the two-stage process with a slight modification. The requirement is to ensure R_{B_i} to be *quasi*-concave as long as R_X is so. Then, we have the following result.

Theorem 1. *Given a two-step process (G, H, r^*_β, F) with $r^*_\beta = \inf \operatorname{argmax}_r R_B(r)$ and such that F is quasi-regular[1], $\forall \alpha \in [0,1]$, $\exists 0 \le x_0 \le x_1$ such that the best quasi-regular response (maximizing $U_\alpha(\beta)$ with F_B regular) is*

$$\tilde{\beta}_{x_0,x_1}(x) = \mathbb{1}_{[x \le x_0]} x + \mathbb{1}_{[x_0 < x \le x_1]} \frac{R}{1-F(x)} + \mathbb{1}_{[x > x_1]} x$$
$$\textit{where} \qquad R = x_1(1 - F(x_1))$$

Moreover, we have $x_1 = \sup\{x : x(1-F(x)) \ge R\}$ and $x_0 \le \overline{x}_0 = \inf\{x : x(1-F(x)) \ge R\}$.

The proof is in Appendix B.2.

Remark. *The thresholding strategy from Tang and Zeng (2018); Nedelec et al. (2018) is a special case of β_{x_0,x_1} in the case when $x_0 = 0$ and we will denote it $\tilde{\beta}_{x_1}$ for simplicity. As we will see in Section 3.1, when α is small enough (but not 0 yet), $x_0 = 0$ is optimal. It thus makes sense to study both these strategies.*

This class of strategies solves the main drawback of the best response from Tang and Zeng (2018). If the

[1] We say a distribution is quasi-regular if the associated revenue curve R_{X_i} is quasi-concave

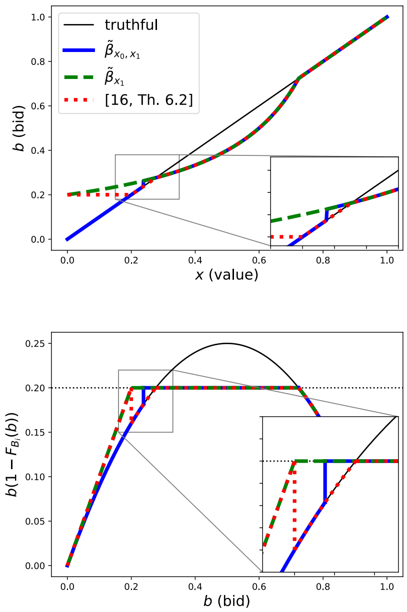


Figure 1: Top: illustration of strategies. Bottom: function optimized by the seller $b(1-F_B(b))$ depending on the strategy for values distribution F being $\mathcal{U}(0,1)$. R is set at 0.2, x_1 is set such that $x_1(1-F(x_1)) = R$ (right root) and $x_0 \simeq 0.27$ for $\tilde{\beta}_{x_0,x_1}$.

value distribution F_i is quasi-regular, then the bid distribution F_B is also quasi-regular, offering a quasi-concave maximization problem (thus “predictable" for the buyer) to the seller. Figure 1 illustrates the different strategies as well as the corresponding optimization problems and the virtual values associated to the push-forward bid distributions. Understanding the strategic answer of the bidders helps avoiding worst-case scenario reasoning when studying the revenue of the sellers against strategic bidders. We now quantify precisely how the welfare is shared between strategic bidders and seller.

3 Welfare sharing between seller and buyers

This section presents how the welfare is shared in the two-stage process with strategic bidders. First, we show how to numerically compute the best response and illustrate the variation of utility and payment with α. After remarking and showing that the *thresholded* strategy Nedelec et al. (2018) is optimal for greater values of α than 0, we focus on this strategy and show the existence of a Nash equilibrium. It proves that even in the case of multiple strategic bidders, there exists strategies that enable bidders to take advantage of the learning stage of the seller.

3.1 Welfare sharing with best quasi-regular response

We first show how to obtain numerically the best response $\tilde{\beta}_{x_0,x_1}$ by solving the maximization problem of U_α restricted to the class of strategies described in Th. 1. This allows us to show how the bidder's utility and payment vary with α.

Theorem 2. *Given a two-stage process (G, H, r^*_β, F) with $r^*_\beta = \inf \operatorname{argmax}_r R_B(r)$, the best response is of the form of $\tilde{\beta}_{x_0^*,x_1^*}$ and the utility $U_\alpha(\tilde{\beta}_{x_0,x_1})$ has the following derivatives:*

$$\frac{\partial U_\alpha(\tilde{\beta}_{x_0,x_1})}{\partial x_0} = \alpha(1-F(x_0))G(x_0)H(x_0)$$
$$- x_0 f(x_0) G_\alpha\left(\frac{x_1(1-F(x_1))}{(1-F(x_0))}\right)$$
$$\frac{\partial U_\alpha(\tilde{\beta}_{x_0,x_1})}{\partial x_1} = f(x_1)\psi(x_1)\Bigg(G_\alpha(x_1)$$
$$- \mathbb{E}_X\left(\mathbf{1}_{[x_0 \le X \le x_1]} \frac{X}{(1-F(X))} g_\alpha\left(\frac{x_1(1-F(x_1))}{(1-F(X))}\right)\right)\Bigg)$$

where $G_\alpha(x) = \alpha G(x)H(x) + (1-\alpha)G(x)$ and $\psi(x) = x - \frac{1-F(x)}{f(x)}$.

The proof is in Appendix B.3. We now use these results to compute numerically the best response in the following two-stage process: $P_1 = (G{=}U_{[0,1]}, H{=}0, r^*_{\tilde{\beta}_{x_0^*,x_1^*}}, F{=}U_{[0,1]})$ (uniformly distributed reserve in stage 1)

The results are presented in Fig. 2. First, we recover the results from the literature corresponding to the extremes values of α. For $\alpha = 1^-$, the payment and utility in the second stage are the ones of a 2^{nd}-price with monopoly price of value distribution. As α decreases, the total payment over the two stages quickly decreases (and utility increases) towards values that are even more favorable to the bidder than the ones of a 2^{nd}-price with no reserve price. This is consistent with the observations from Nedelec et al. (2018) (case $\alpha = 0$): if only one bidder is strategic, the payment of this strategic bidder can decrease even further than the one of a 2^{nd}-price with no reserve price. In Appendix, we provide the same study with random reserve and show that the presence of exploration of the reserve prices in the first period (difference between left and right) does not impact much the payment/utility of the second period. It slightly decreases the advantage of the bidder, but not significantly.

Finally, we observe two regimes for x_0^* (Fig. 2, right): for $\alpha < 0.8$ it is 0 and for $\alpha > 0.8$ it is the value that saturates the inequality constraint described by Th. 1 between x_0 and x_1. This illustrates why we cannot only

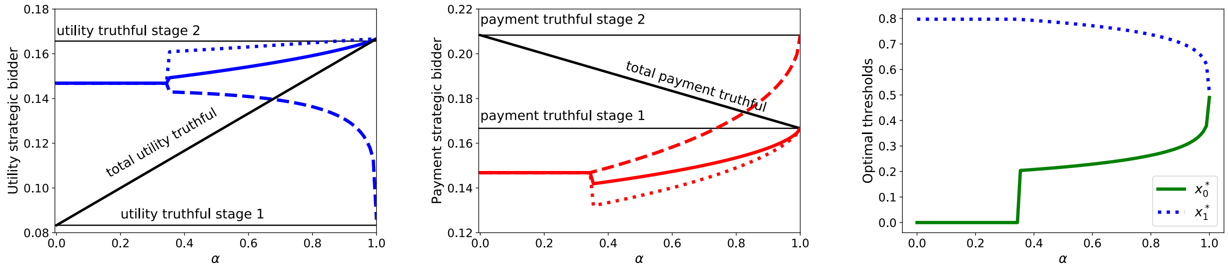


Figure 2: Welfare sharing with random reserve in stage 1: Left – Utility of strategic bidder in the two-stage game as a function of α. Middle – Payment of strategic bidder in the two-stage game as a function of α. The dotted lines corresponds to phase 1, the dashed ones to phase 2 and the solid line for the full game. The blue color is used for the bidder's utility and the red for her payment. The black color is used for the baseline of truthful bidding. Right: evolution of x_0^*, x_1^* with α.

use first-order optimality condition: the optimal value is often reached on the edge of the feasible domain. We formalize this latter observation and shows that when α is small enough, the reserve value x_0^* is exactly 0.

Lemma 1. *Given a two-step process* (G, H, r_β^*, F) *with* $r_\beta^* = \inf \operatorname{argmax}_r R_B(r)$. *If* $\exists y \in [0, b], \forall x < y, \frac{G(x)H(x)}{f(x)} < x$ *and* $\frac{G(x)H(x)}{f(x)}$ *is bounded on* $[0, x^*]$ *with* $x^* = \inf \operatorname{argmax}_x R_X(x)$, *then* $\exists \alpha_0 > 0$ *such that* $\forall \alpha < \alpha_0$,

$$(x_0^*, x_1^*) \in \operatorname*{argmax}_{x_0, x_1} U_\alpha(\tilde{\beta}_{x_0, x_1}) \quad \Rightarrow \quad x_0^* = 0$$

The proof can be found in Appendix B.3. $\frac{G(x)H(x)}{f(x)}$ being bounded on $[0, x^*]$ is satisfied if F is MHR (monotonous hazard rate). $\frac{G(x)H(x)}{f(x)} < x$ is the key assumption. Intuitively, it can be interpreted as *"if the bidder has enough mass under the competition, she can overbid at almost no cost on small values to decrease the reserve value to 0."* This theorem shows that the *thresholding* strategies proposed in Tang and Zeng (2018); Nedelec et al. (2018) are actually optimal for α small enough and not just for $\alpha = 0$. In the following subsections, we carry out a formal analysis for such strategies $\tilde{\beta}_{0,x_1} = \tilde{\beta}_{x_1}$ as we can more easily characterize the best response and prove Nash equilibria.

3.2 Phase transition for the classical thresholding strategies in the two-stage game

We suppose in this subsection that the seller commits to thresholding, i.e. $x_0 = 0$ with the notation above. The seller uses monopoly reserves in the second phase.

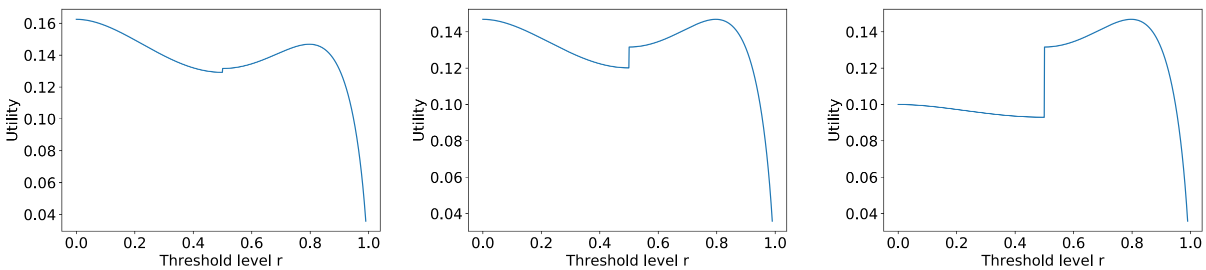
Theorem 3. *We call* $\mathcal{G}_\alpha = \alpha G_1 + (1 - \alpha) G_2$, G_1 *and* G_2 *being the distributions of the competition faced by the bidder in the two phases, possibly including reserve prices. We assume* $\psi^{-1}(0) > 0$. *If the buyer uses the thresholding strategy* $\tilde{\beta}_{x_1}$ *of Theorem 1 and commits to it, we have for their utility, if the seller is welfare benevolent : the utility has (in general) a discontinuity at* $x_1 = \psi^{-1}(0)$. *For* $x_1 < \psi^{-1}(0)$, *we have* $U(0) \geq U(x_1)$. *For* $x_1 > \psi^{-1}(0)$, *the first order condition are the same as in Nedelec et al. (2018); Tang and Zeng (2018) where the distribution of the competition is now* $\mathcal{G}_\alpha = \alpha G_1 + (1 - \alpha) G_2$. *Call* $x_1^*(\alpha)$ *the unique solution of this problem. Hence, the optimal threshold is* $\operatorname{argmax}_{x_1 \in \{0, x_1^*(\alpha)\}} U(x_1)$. *An optimal threshold at* $x_1 = 0$ *corresponds to bidding truthfully.*

Lemma 2 (Phase transition in α)**.** *Assume the setup of Theorem 3. Suppose that* $G_1 = G_2$, *i.e. the distribution of the competition is the same in the two phases and there is no reserve price in the first phase. Then there is a critical value* α_c *for which, if* $\alpha < \alpha_c$, *it is preferable for the bidder to threshold and if* $\alpha > \alpha_c$, *it is preferable for the bidder to bid truthfully.*

The proofs of the theorem and the lemma can be found in Section D and we provide more details on the utility there. A graphical illustration of these results can be found in Figure 3.

3.3 Existence of a Nash equilibrium in the two-stage game

Most of the literature on strategic buyers in repeated auctions have focused on the posted price setting Amin et al. (2013); Medina and Mohri (2014), though see Golrezaei et al. (2019) for a recent extension to second price auction. To complement this line of work, we exhibit now the existence of a Nash equilibrium between strategic bidders in the two-stage game. Our approach here is the following. All players have essentially two strategies, according to Theorem 3: they can either threshold optimally above the monopoly price or bid truthfully. We consider the first order conditions for a Nash equilibrium if all players threshold and we then



(a) $\alpha = .95$ (b) $\alpha = .76$ (c) $\alpha = .2$

Figure 3: **Utility of the strategic buyer in 2 stage game for various threshold levels r's, zero reserves**. We consider a setup where we have two bidders, both have Unif[0,1] value distribution. One is strategic. There is no reserve price in the first stage. Different α's are displayed. At $\alpha \simeq .762$, the truthful strategy is essentially equivalent to optimal thresholding. For smaller α's, the optimal thresholding is preferable (right). For higher α, truthful bidding is preferred (a). This illustrates the results of Theorem 3

.

study the best response of the remaining player.

Theorem 4. *We consider the symmetric case where all bidders have the same value distribution F and it is regular. We assume for simplicity that the distribution is supported on [0,1]. Suppose this distribution has density that is continuous at 0 and 1 with $f(1) \neq 0$ and ψ crosses 0 exactly once and hence is positive beyond that crossing point.*

In the case where $\alpha = 0$, there is a unique symmetric Nash equilibrium. The revenue of the bidders in this case is the same as in a second price auction with no reserve.

For any fixed $\alpha > 0$, there exists $\alpha_{c,thresh}$ (possibly 0 or ∞) such that if $\alpha \leq \alpha_{c,tresh}$ there exists a symmetric Nash equilibrium in the class of thresholded strategies. It is the same as in the case $\alpha = 0$ and hence the revenue of the bidders is again the same as in a second price auction with no reserve.

There exists $\alpha_{c,truthful}$ (possibly 0 or ∞) such that if $\alpha \geq \alpha_{c,truthful}$ there exists a symmetric Nash equilibrium in the class of thresholded strategies and it corresponds to bidding truthfully.

The proof for the case $\alpha = 0$ is in Appendix C. The remainder of the proof is in Appendix D. The presence of other strategic bidders does not prevent bidders from taking advantage of the learning stage of the seller.

4 Robustness of bidding strategies to the information structure of the game

We study the robustness of the bidding strategies to several variants of the two-stage game: 1) the seller is using an approximation of the bid distribution to compute the reserve price of the second stage, 2) the bidder does not have an estimation of the competition to compute the optimal strategy for the two-stage game and 3) the seller replaces the lazy second price with monopoly reserve with another type of auction. Their simplicity and robustness make them relevant in real-world interactions.

4.1 Robustness to sample approximation of the seller and ERM/optimization algorithms

In practice, the seller needs to estimate the distribution of the buyer and hence does not have a perfect knowledge of the bid distribution F_{B_i}. The buyer needs to find a robust shading method, making sure that the seller has an incentive to lower her reserve price, even if she misestimates the bid distribution.

We call ψ_F the virtual value function associated with the distribution F, i.e. $\psi_F(x) = x - \frac{1-F(x)}{f(x)}$.

Lemma 3. *Suppose that the buyer uses a strategy β under her value distribution F. Suppose the seller thinks that the value distribution of the buyer is G. Call λ_F and λ_G the hazard rate functions of the two distributions Then the seller computes the virtual value function of the buyer under G, denoted $\psi_{B,G}$, as*

$$\psi_{B,G}(\beta(x)) = \psi_{B,F}(\beta(x)) - \beta'(x)\left(\frac{1}{\lambda_G(x)} - \frac{1}{\lambda_F(x)}\right).$$

As an aside, we note that by definition we have $\frac{1}{\lambda_G(x)} - \frac{1}{\lambda_F(x)} = \psi_F(x) - \psi_G(x)$. We have the following useful corollary pertaining to the thresholded strategies described in Section 3.

Corollary 1. *If the buyer uses the strategy $\tilde{\beta}_r^{(\epsilon)}(x)$*

defined as

$$\tilde{\beta}_r^{(\epsilon)}(x) = \left(\frac{(r-\epsilon)(1-F(r))}{1-F(x)} + \epsilon \right) \boldsymbol{1}_{[x \le r]} + x \boldsymbol{1}_{[x>r]} \ ,$$

we have, for $x \neq r$, $\psi_{B,F}(\tilde{\beta}_r^{(\epsilon)}(x)) = \epsilon \boldsymbol{1}_{[x\le r]} + \psi_F(x) \boldsymbol{1}_{[x>r]}$ *. In particular, we have for* $x \neq r$

$$\left|\psi_{B,F}(\tilde{\beta}_r^{(\epsilon)}(x)) - \psi_{B,G}(\tilde{\beta}_r^{(\epsilon)}(x))\right| \le |\psi_F(x) - \psi_G(x)| \left[(r-\epsilon) \boldsymbol{1}_{[x\le r]} + \boldsymbol{1}_{[x>r]} \right] \ .$$

If for all x, $\psi_{B,F}(\tilde{\beta}_r^{(\epsilon)}) \ge \epsilon$ *and* $|\psi_F(x) - \psi_G(x)| \le \delta$, *we have* $\psi_{B,G}(\tilde{\beta}_r^{(\epsilon)}(x)) \ge \epsilon - \delta \max((r-\epsilon), 1)$ *.*

Hence, a natural way to quantify the proximity of distributions in this context is of course in terms of their virtual value functions. Furthermore, if the buyer uses a shading function such that, under her strategy and with her value distribution, the perceived virtual value is positive, as long as the seller computes the virtual value using a nearby distribution, she will also perceive a positive virtual value and hence have no incentive to put a reserve price above the lowest bid. In particular, if δ comes from an approximation error that the buyer can predict or measure, she can also adjust her ϵ so as to make sure that the seller perceives a positive virtual value for all x. We pursue this specific question in more details in the next subsection.

The previous results already give some results about the impact of empirically estimating the value distribution F by the empirical cumulative distribution function $\hat{F}_n$ on setting the reserve price. However because our approximations are formulated in terms of hazard rate, applying those results would yield quite poor approximation results in the context of setting the monopoly price through ERM. This is due to the fact that estimating a density pointwise in supremum norm is a somewhat difficult problem in general, associated with poor rates of convergence. We refer the interested reader to Tsybakov (2009) for more details on this question.

So we now focus on the specific problem of empirical minimization and take advantage of its characteristics to obtain better results than would have been possible by applying the results of the previous section naively.

Theorem 5. *Suppose the buyer has a continuous and increasing value distribution* F*, supported on* $[0, b]$, $b \le \infty$*, with the property that if* $r \ge y \ge x$, $F(y) - F(x) \ge \gamma_F(y-x)$*, where* $\gamma_F > 0$*. Suppose that* $\sup_{t\ge r} t(1-F(t)) = r(1-F(r))$*. Suppose the buyer uses the strategy* $\tilde{\beta}_r^{(\epsilon)}$ *defined as*

$$\tilde{\beta}_r^{(\epsilon)}(x) = \left(\frac{(r-\epsilon)(1-F(r))}{1-F(x)} + \epsilon \right) \boldsymbol{1}_{[x \le r]} + x \boldsymbol{1}_{[x>r]} \ ,$$

Assume she samples n *values* $\{x_i\}_{i=1}^n$ *i.i.d according to the distribution* F *and bids accordingly in second price auctions. Call* $x_{(n)} = \max_{1\le i\le n} x_i$*. In this case the (population) reserve value* x^* *is equal to 0. Assume that the seller uses empirical risk minimization to determine the monopoly price in a (lazy) second price auction, using these* n *samples. Call* $\hat{x}_n^*$ *the reserve value determined by the seller using ERM. We have, if* $C_n(\delta) = n^{-1/2}\sqrt{\log(2/\delta)/2}$ *and* $\epsilon > x_{(n)} C_n(\delta)/F(r)$ *with probability at least* $1-\delta_1$*,*

$$\hat{x}_n^* < \frac{2rC_n(\delta)}{\epsilon\gamma_F} \text{ with probability at least } 1 - (\delta + \delta_1) \ .$$

In particular, if ϵ *is replaced by a sequence* ϵ_n *such that* $n^{1/2}\epsilon_n min(1, 1/x_{(n)}) \to \infty$ *in probability,* $\hat{x}_n^*$ *goes to 0 in probability like* $n^{-1/2}\max(1, x_{(n)})/\epsilon_n$*.*

Informally speaking, our theorem says that using the strategy $\tilde{\beta}_r^{(\epsilon_n)}$ with ϵ_n slightly larger than $n^{-1/2}$ will yield a reserve value arbitrarily close to 0. Hence the population results we derived in earlier sections apply to the sample version of the problem. We give examples and discuss our assumptions in Appendix E.2 where we prove the theorem.

We note that the flexibility afforded by ϵ is two-fold: when $\epsilon > 0$, the extra seller revenue is a strictly decreasing function of the reserve price; hence even if for some reason reserve price movements are required to be small, the seller will have an incentive to make such move. The other reason is more related to estimation issues: if the reserve price is determined by empirical risk minimization, and hence affected by even small sampling noise, having ϵ big enough will guarantee that the mean extra gain of the seller will be above this sampling noise. Of course, the average cost for the bidder can be interpreted to just be ϵ at each value under the current reserve price and hence may not be a too hefty price to bear.

4.2 Robustness to the knowledge of the competition distribution

Another common situation is that bidders do not know in advance the competition G they are facing. They need to estimate it from past interactions with the seller and other bidders. First, we show that there exists some strategies that do not need a precise estimation of the competition. Then, we look at some worst-case scenario where the goal is to find the strategy with the highest utility in the worst case of G. We are optimizing the bidding strategy in the class of thresholded strategies.

Theorem 6 (Thresholding at the monopoly price)**.** *Consider the one-strategic setting in a lazy second price auction with* F_{X_i} *the value distribution of the* *strategic bidder i with a seller computing the reserve prices to maximize her revenue. Consider β_{tr} the truthful strategy. Consider the case of $\alpha = 1$. Then there exists β which does not depend on G and :* ***1)*** *$U_i(\beta) \geq U_i(\beta_{tr})$, U_i being the utility of the strategic bidder.* ***2)*** *$R_i(\beta) \geq R_i(\beta_{tr})$, R_i being the payment of bidder i to the seller. Then, $\tilde{\beta}^{(\epsilon)}_{\psi^{-1}(0)}$ fulfill these conditions for $\epsilon \geq 0$ small enough.*

For $\epsilon = 0$, we call this strategy the thresholded strategy at the monopoly price. In the two-stage game, there is a critical value α_c for which, if $\alpha < \alpha_c$, it is preferable for the bidder to threshold at the monopoly price and if $\alpha > \alpha_c$, it is preferable for the bidder to bid truthfully.

The formal proof in the general case is in Appendix F. The nice and crucial property of thresholding at the monopoly price is that it does not depend on a specific knowledge of the competition.

Numerical applications in the case $\alpha = 0$: In the case of two bidders with uniform value distribution, the strategic bidder utility increases from 0.083 to 0.132, (a 57% increase). In the case of two bidders with exponential distribution,with parameters $\mu = 0.25$ and $\sigma = 1$, the utility of the strategic bidders goes from 0.791 to 1.025 (a 29.5% increase). This theorem shows that even if a fine-tuned knowledge of the competition helps bidders increase their utility, there exists a strategy where bidders do not need to know the competition and can still significantly increase their utility. Due to lack of space, we provide more results on the impact of knowledge of the competition in Appendix F.

4.3 Robustness to a change of mechanism

We finally show that the thresholded strategies are robust to certain changes of mechanism. This makes sense in the framework of Stackelberg games with no commitment where the player performs a certain optimization given the objective function of the other player whereas this other player can decide to change her optimization problem.

Myerson auction. In the one-strategic setting, in the symmetric case where all bidders have initially the same value distribution, we show that the utility gain for the strategic bidder of thresholding at the monopoly price is higher in the Myerson auction than in the lazy second price auction.

Lemma 4. *Consider the case where the distribution of the competition is fixed. Assume all bidders have the same value distribution F_X, and that F_X is regular. Assume that bidder i is strategic and that the $K-1$ other bidders bid truthfully. Let us denote by β_{tru} the truthful strategy and β_{thr} the thresholded strategy at the monopoly price. The utility of bidder i in the Myerson auction U_i^{Myer} and in the lazy second price auction U_i^{Lazy} satisfy*

$$U_i^{Myer}(\beta_{thr}) - U_i^{Myer}(\beta_{tru}) \geq U_i^{Lazy}(\beta_{thr}) - U_i^{Lazy}(\beta_{tru}).$$

The proof is given in Appendix G.1. Numerics with $K = 2$ bidders with $\mathcal{U}[0,1]$ value distribution: The utility is 1/12 in the truthful case in both lazy second price and Myerson since these auctions are identical in the symmetric case. The utility of the thresholded strategy at monopoly price is 7/48 in the Myerson auction, i.e. 75% more than the utility with the truthful strategy. We note that the gain is larger than for a second price auction with monopoly reserve where the extra utility was 57%.

Eager second price auction with monopoly price. Monopoly reserves are not optimal reserve prices for this version of the second price auction in general but in practice, the optimal ones are NP-hard to compute Paes Leme et al. (2016). We recall that the eager second price auction is a standard second price auction between bidders who clear their reserves.

Lemma 5. *Consider the same setting and notations as Lemma 4. The utility of bidder i in the eager second price auction with monopoly reserves, U_i^{Eager}, and in the lazy second price auction U_i^{Lazy} with the same reserves, satisfy :*

$$U_i^{Eager}(\beta_{thr}) - U_i^{Eager}(\beta_{tru}) \geq U_i^{Lazy}(\beta_{thr}) - U_i^{Lazy}(\beta_{tru}).$$

The proof is given in Appendix G.2. These two lemmas show that thresholded strategies can increase the utility of strategic bidders even if the seller runs a different auction in the second stage than the lazy second price auction with monopoly price. We plan to design games where the seller can change the mechanism and bidders can update their bidding strategy at any point.

5 Conclusion

We propose novel optimal strategies for the canonical problem of auction design with unknown value distributions that are learned in an exploration phase and exploited thereafter. The Stackelberg game we consider exhibit complex solutions but we provide simple strategies that can be made robust to various setups reflecting how much information bidders have about the game they are participating in. This allows buyers to quantify the price of revealing information about their values in repeated auctions. It also opens new avenues for research on the seller side by providing new realistic strategies that may adopted by the strategic bidder.

References

Michele Aghassi and Dimitris Bertsimas. Robust game theory. *Mathematical Programming*, 2006.

Kareem Amin, Afshin Rostamizadeh, and Umar Syed. Learning prices for repeated auctions with strategic buyers. *Proceedings of NIPS*, 2013.

Kareem Amin, Afshin Rostamizadeh, and Umar Syed. Repeated contextual auctions with strategic buyers. *Proceedings of NIPS*, 2014.

Itai Ashlagi, Constantinos Daskalakis, and Nima Haghpanah. Sequential mechanisms with ex-post participation guarantees. *Proceedings of EC*, 2016.

N. Cesa-Bianchi, C. Gentile, and Y. Mansour. Regret minimization for reserve prices in second-price auctions. *Proceedings of SODA*, 2013.

Richard Cole and Tim Roughgarden. The sample complexity of revenue maximization. *Proceedings of Theory of computing*, 2014.

Peerapong Dhangwatnotai, Tim Roughgarden, and Qiqi Yan. Revenue maximization with a single sample. *Games and Economic Behavior*, 91, 2015.

Engelbert J. Dockner, Steffen Jorgensen, Ngo Van Long, and Gerhard Sorger. *Differential Games in Economics and Management Science*. Cambridge University Press, 2000.

Paul Dütting, Zhe Feng, Harikrishna Narasimhan, and David C Parkes. Optimal auctions through deep learning. *Proceedings of ICML*, 2019.

Alessandro Epasto, Mohammad Mahdian, Vahab Mirrokni, and Song Zuo. Incentive-aware learning for large markets. *Proceedings of WWW*, 2018.

Negin Golrezaei, Adel Javanmard, and Vahab Mirrokni. Dynamic incentive-aware learning: Robust pricing in contextual auctions. *Proceedings of NeurIPS*, 2019.

Zhiyi Huang, Yishay Mansour, and Tim Roughgarden. Making the most of your samples. *SIAM Journal on Computing*, 2018.

Yash Kanoria and Hamid Nazerzadeh. Dynamic reserve prices for repeated auctions: Learning from bids. *Proceedings of WINE*, 2014.

V. Krishna. *Auction Theory*. 2009.

P. Massart. The tight constant in the dvoretzky-kiefer-wolfowitz inequality. *The Annals of Probability*, 18, 1990.

Andres M Medina and Mehryar Mohri. Learning theory and algorithms for revenue optimization in second price auctions with reserve. *Proceedings of ICML*, 2014.

Mehryar Mohri and Andres Munoz. Revenue optimization against strategic buyers. *Proceedings of NIPS*, 2015.

Jamie H Morgenstern and Tim Roughgarden. On the pseudo-dimension of nearly optimal auctions. *Proceedings of NIPS*, 2015.

R. B. Myerson. Optimal auction design. *Mathematics of Operation Research*, 6(1), 1981.

Thomas Nedelec, Marc Abeille, Clément Calauzènes, Noureddine El Karoui, Benjamin Heymann, and Vianney Perchet. Thresholding the virtual value: a simple method to increase welfare and lower reserve prices in online auction systems. *arXiv preprint arXiv:1808.06979*, 2018.

Thomas Nedelec, Noureddine El Karoui, and Vianney Perchet. Learning to bid in revenue-maximizing auctions. *Proceedings of ICML*, 2019.

M. Ostrovsky and M. Schwarz. Reserve prices in internet advertising auctions: A field experiment. *Proceedings of EC*, 2011.

Renato Paes Leme, Martin Pal, and Sergei Vassilvitskii. A field guide to personalized reserve prices. *Proceedings of WWW*, 2016.

Pingzhong Tang and Yulong Zeng. The price of prior dependence in auctions. *Proceedings of EC*, 2018.

Alexandre B. Tsybakov. *Introduction to nonparametric estimation*. 2009.