Robust Auction Design with Support Information

Jerry Anunrojwong* Santiago R. Balseiro† Omar Besbes‡

March 13, 2023

Abstract

A seller wants to sell an indivisible item to n buyers. The buyer valuations are drawn i.i.d. from a distribution, but the seller does not know this distribution; the seller only knows the support $[a, b]$. To be robust against the lack of knowledge of the environment and buyers' behavior, the seller optimizes over dominant strategy incentive compatible (DSIC) mechanisms, and measures the *worst-case* performance relative to an oracle with complete knowledge of buyers' valuations. Our analysis encompasses both the *regret* and the *approximation ratio* objectives.

For these objectives, we derive an optimal mechanism in quasi-closed form, and the associated performance, as a function of the support and the number of buyers n. Our analysis reveals three regimes of support information and a new class of robust mechanisms. i.) With "low" support information, the optimal mechanism is a second-price auction (SPA) with a random reserve, a focal class in the earlier literature. ii.) With "high" support information, we show that second-price auctions are strictly suboptimal, and we establish that an optimal mechanism belongs to a novel class of mechanisms we introduce, which we call *pooling auctions* (POOL); whenever the highest value is above a threshold, the mechanism still allocates to the highest bidder, but otherwise the mechanism allocates to a *uniformly random buyer*, i.e., pools low types. iii.) With "moderate" support information, we establish that a randomization between SPA (with a random reserve price) and POOL (with a random threshold) is optimal.

We also characterize optimal mechanisms within nested central subclasses of mechanisms: standard mechanisms that only allocate to the maximum value bidder, SPA with random reserve, and SPA with no reserve. We show strict separations in terms of performance across classes, implying that deviating from standard mechanisms is necessary for robustness. Lastly, we show that the same results hold under other distribution classes that capture "positive dependence", namely: i.i.d., mixture of i.i.d., and exchangeable and affiliated distributions, as well as i.i.d. regular distributions.

Keywords: robust mechanism design, minimax regret, maximin ratio, support information, prior-independent, standard mechanisms, second-price auctions, pooling.

*Columbia University, Graduate School of Business. Email: `janunrojwong25@gsb.columbia.edu`
†Columbia University, Graduate School of Business. Email: `srb2155@columbia.edu`
‡Columbia University, Graduate School of Business. Email: `ob2105@columbia.edu`

1 Introduction

The question of how to optimally sell an item underlies much of modern marketplaces, from online advertising and e-commerce to art auctions. Selling mechanisms are widely used in practice, and in turn they are studied in economics, computer science, and operations research under *optimal mechanism design*, starting from the pioneering work of [Myerson, 1981]. The literature often assumes that the seller knows the environment perfectly, but (i) this knowledge is often either not available or reliable, and (ii) the optimal mechanism prescribed by the theory is often too complicated or fine-tuned to the details of the environment, to be used in practice. There is therefore a need to develop mechanisms that depend less on market details, and this need is often referred to as the "Wilson doctrine" [Wilson, 1987].

The emerging literature on *robust mechanism design*, in turn, aims to design mechanisms that perform "well" in the worst case against "any" environment. This line of work often leads to interesting insights but, taken literally, they can lead to mechanisms that are too conservative. In practice, while we do not have complete knowledge about the environment, we often do have *partial* knowledge, and how to incorporate additional side information into the robust framework is essential to bring the robust theory closer to practice. In this paper, we make progress in this direction by analyzing the role of *support information* of bidder valuations, as captured by *lower bounds* and *upper bounds* on bidder valuations. This is a natural form of partial knowledge in an uncertain environment because, while we might not know the *shape* of the underlying distributions (e.g., a particular parametric form), we often do have some information on the *range* of the bidders' values, from historical data and domain knowledge. For example, in online advertising auctions, "the typical cost per click for a search ads in 2022 is in the range of \$2 to \$4" [McCormick, 2022].

More formally, consider a seller who wants to sell an item to n bidders. The bidders' valuations are unknown to the seller and are assumed to be drawn from a joint distribution $\boldsymbol{F}$. The seller does not know $\boldsymbol{F}$, and knows only the lower bound a and the upper bound b on the support of $\boldsymbol{F}$, and that the valuations belong to a given class $\mathcal{F}$. The class $\mathcal{F}$ captures qualitative assumptions on dependence between agent valuations; for convenience, we will mostly assume that the valuations are i.i.d., but later we will also show that our results extend to the case when valuations are "positively dependent."[1] Similarly, the bidders also do not know $\boldsymbol{F}$. Therefore, we focus on mechanisms that are

[1] More precisely, when valuations are either (i) mixtures of i.i.d., or (ii) exchangeable and affiliated. We will give

dominant strategy incentive compatible (DSIC). Under such a mechanism, every bidder optimally reports her true value regardless of other bidders' valuations and strategies.

We will quantify the performance of mechanisms by the *gap* between the benchmark oracle revenue and the mechanism revenue. The benchmark revenue is the ideal expected revenue the seller could have collected with knowledge of the buyers' valuations, while the mechanism revenue is the expected revenue garnered by the actual mechanism. Our framework will be general and apply to two classical notions of gaps considered in the literature: (i) the *regret* (absolute gap) is the difference between these two revenues, and (ii) the *approximation ratio* (relative gap) is the ratio of these two revenues. The seller selects a mechanism that performs well (minimizes regret or maximizes approximation ratio) in the *worst case* against all admissible distributions.

The support $[a, b]$ of the admissible distribution class captures the amount of uncertainty of the decision maker. We will parameterize this uncertainty through a/b, which we call the *relative support information*, and which is a unitless quantity ranging from 0 to 1. When $a/b \sim 0$ (either because $a \sim 0$ or $b \gg a$), we have minimal relative support information while when $a/b \sim 1$ we have maximal support information as the endpoints are close.

For the special case of pricing, i.e., when there is one bidder, the problem is well understood. It was analyzed concurrently in Bergemann and Schlag [2008] for the regret objective and in Eren and Maglaras [2010] for the approximation ratio objective. When moving to auctions with an arbitrary number of bidders n, in the present context, only the case $a = 0$, which corresponds to minimal relative support information, has been studied. Anunrojwong et al. [2022] shows that a second-price auction with appropriately randomized reserve price is an optimal minimax regret mechanism. It is also easy to see that with minimal relative support information, no mechanism can guarantee positive worst-case approximation ratio. The understanding of the interplay of support information and robust auctions is very limited outside of these special cases. How does support information affect the structure of optimal robust auctions and achievable performance?

This is the departure point of our work. We study optimal mechanism and performance across the relative support information spectrum and establish richness in the structure of the resulting robust mechanisms with three distinct information regimes corresponding to three mechanism types. In particular, our work subsumes and unifies the three studies mentioned above, characterizing an

the definitions of these in Appendix G.2.

optimal mechanism and the associated performance for an arbitrary number of bidders n and any support information $[a, b]$, for both the regret and ratio objectives. See Table 1 for a high level summary of known results and the results we develop in this paper.

Problem Type	Information level	Objective: Regret	Objective: Ratio
pricing ($n = 1$)	all a/b	Bergemann and Schlag [2008]	Eren and Maglaras [2010]
auctions ($n \geq 1$)	$a/b = 0$	Anunrojwong et al. [2022]	0
auctions ($n \geq 1$)	all a/b	—**This work**—	

Table 1: Comparison with the closest previous studies along the dimension of the number of buyers (pricing ($n = 1$) vs. auctions (arbitrary $n \geq 1$) and the level of relative support information a/b.

1.1 Summary of Main Contributions

We develop a unified framework for regret and approximation ratio through a single quantity, the minimax λ-regret, where λ-regret is the difference between λ times the benchmark revenue and the mechanism revenue, and $\lambda \in (0, 1]$ is a constant. The reduction itself is through an epigraph formulation and is fairly standard. Our main contribution, however, is the full characterization of a minimax optimal mechanism and its associated performance for λ-regret for *any* value of $\lambda \in (0, 1]$, *any* number of buyers n and *any* support $[a, b]$. Since we are primarily interested in the effect of the support $[a, b]$, we initially assume that the valuations are n i.i.d. distributions given the canonical nature of this setting.[2] Our family of optimality results across this spectrum brings to the foreground a very rich structure of optimal mechanisms, and establishes how relative support information critically impacts the structure of optimal mechanisms.

Novel mechanism class. A natural candidate for an optimal mechanism is a second-price auction with appropriate random reserve. This was shown to be optimal with zero relative support information, i.e, for $[0, b]$ support. Suppose for a moment that relative support information is high (i.e., $a \sim b$) and suppose we are restricted to the class of SPAs. Now suppose the seller considers setting any nontrivial reserve; this is *risky* because when the highest buyer's value is below the reserve the seller then does not allocate and gets zero revenue. At the same time, the benefits of a

[2]We show later that we can extend our method and results to other distribution classes that capture "positive dependence," and similar equivalences and reductions hold.

reserve price are limited since the highest and lowest values are close. On the other hand, the seller can guarantee a revenue of a with no reserve, which is close to the maximal revenue achievable of b. Hence, it should be intuitive that when relative support information is high, a SPA with no reserve is optimal among the class of SPAs. (A formal result is presented in Section 4.2.) A natural question is then whether there exists mechanisms that can outperform a SPA with no reserve from a robust perspective, and what their structures are.

We define a new mechanism class, with the aim of softening the trade-offs associated with reserve pricing in second-price auctions. These mechanisms, that we dub "pooling auctions," have an associated threshold. When the highest bid is above the threshold, the mechanism allocates to the highest bidder, as in a SPA when the highest bid is above the reserve price. However, when the highest bid is below the threshold, rather than not allocating as a SPA would do, the seller allocates uniformly at random to any of the bidders. In particular, the lowest bidder may get the item. In other words, this auction pools the low types. Slightly more formally, we can define the following parametrized mechanism classes:

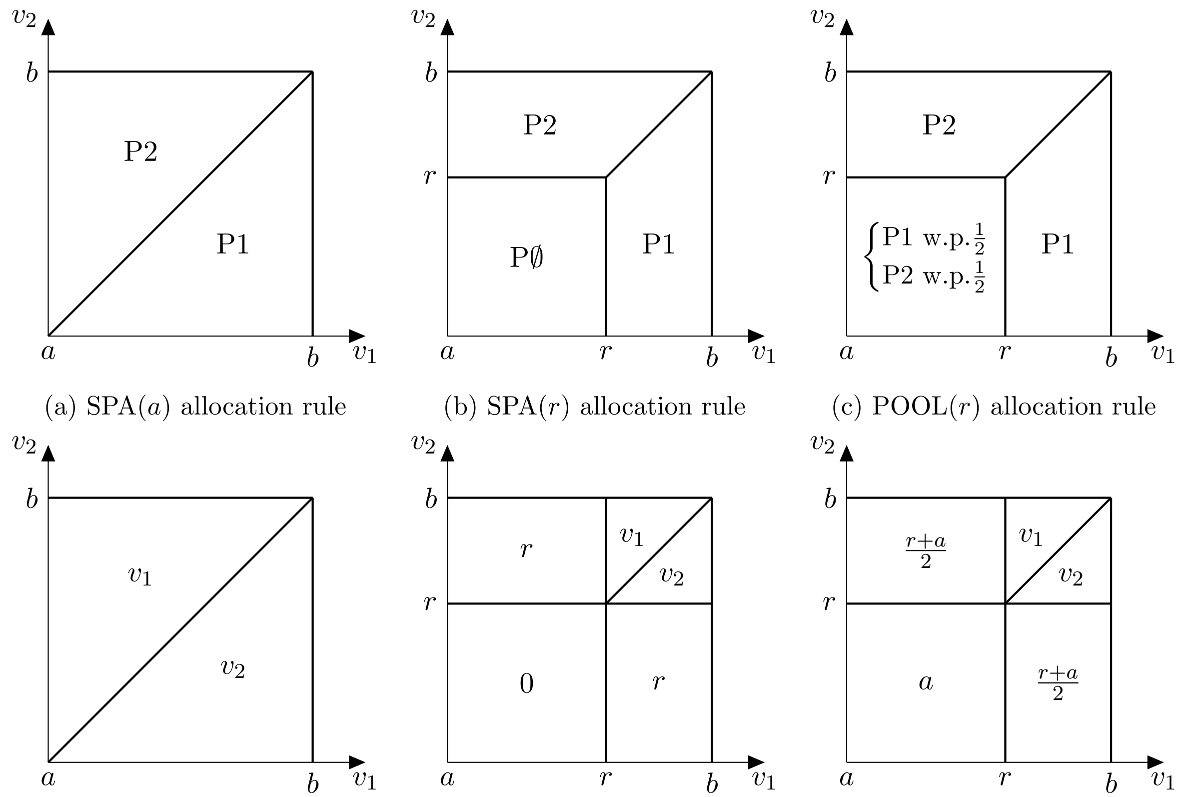
- SPA(r), second-price auction with reserve r, always allocates to the highest-value agent if the highest value is above r, otherwise the mechanism does not allocate.

- POOL(τ), the "pooling auction" with threshold τ, always allocates to the highest-value agent if the highest value is above τ, otherwise the mechanism allocates to each one of the n agents uniformly at random with probability $1/n$.

In a pooling auction, we can still use the threshold to differentiate between bidders with different values and potentially extract more revenue, without risking the zero payoff that comes from not allocating the item. Of course, this has implications for payments. We illustrate this interplay in Figure 1, where we depict, for the case of two agents, the allocation rule $x(\boldsymbol{v})$ and revenue $p_1(\boldsymbol{v})+p_2(\boldsymbol{v})$ at each valuation vector $\boldsymbol{v} = (v_1, v_2)$ for three mechanisms: SPA (no reserve), SPA(r), and POOL(r).[3]

We can see intuitively that pooling low types indeed softens the tradeoff associated with reserve pricing. By increasing the allocation for the low types, we increase the payment accrued from lower-value bidders but, at the same time, we decrease the payment accrued from higher-value bidders to

[3] The payment rule, and thus pointwise revenue, is determined via Myerson's envelope formula.



(a) SPA(a) allocation rule (b) SPA(r) allocation rule (c) POOL(r) allocation rule

(d) SPA(a) revenue (e) SPA(r) revenue (f) POOL(r) revenue

Figure 1: Allocation rules and revenue of SPA without reserve, SPA(r), and POOL(r). In the allocation rule, P1 stands for allocating to player 1, P2 for allocating to player 2, and P$\emptyset$ for not allocating.

guarantee incentive compatibility (so higher-value bidders do not pretend to be lower-value ones). When the relative support information is high (i.e., $a \sim b$), the lower-value and higher-value bidders are not too different, and this softer tradeoff has the potential to lead to more robust mechanisms.

Characterization of an optimal mechanism. Our main theorem fully characterizes a minimax optimal mechanism for any relative support information level, for the λ-regret (and hence for the minimax regret and maximin ratio). We present an abridged version of our main result here. The full version is available in Theorem 2.

Theorem 1 (Main Theorem, Succinct)**.** *Fix n and $\lambda \in (0, 1]$. Then, there exists constants $k_l < k_h$, depending only on n and λ, such that the problem admits an optimal minimax λ-regret mechanism m^*, depending on a/b, as follows.*

- *(Low Relative Support Information) For $a/b \leq k_l$, there is a probability distribution of reserves Φ such that $m^* = \mathrm{SPA}(r)$ with $r \sim \Phi$.*

- *(High Relative Support Information) For $a/b \geq k_h$, there is a probability distribution of thresholds Ψ such that $m^* = \mathrm{POOL}(\tau)$ with $\tau \sim \Psi$.*

- *(Moderate Relative Support Information) For $k_l \leq a/b \leq k_h$, then there is $v^* \in [a, b]$, such that m^* is a randomization over* SPA*s and* POOL*s in the following sense. There is a probability distribution $\mathcal{D}$ on $[a, b]$ such that when we draw a sample $r \sim \mathcal{D}$, if $r \leq v^*$, the mechanism is* $\mathrm{SPA}(r)$*, otherwise the mechanism is* $\mathrm{POOL}(r)$*.*

Our main theorem shows there always exists an optimal mechanism that is a convex combination of different mechanisms in the families $\mathrm{SPA}(r)$ and $\mathrm{POOL}(\tau)$. Therefore, an optimal mechanism can be implemented in terms of a random instance of one of these "base" mechanisms. Furthermore, three fundamental relative support information regimes emerge. In the high (resp. low) support information regime, POOL (resp. SPA) is optimal. The moderate information regime is an interpolation between the two extreme regimes.

The resulting optimal mechanism inherits qualitative features from the basis mechanisms SPA and POOL, and so their properties depend critically on the relative support information. We note that SPA is a standard mechanism, meaning that it never allocates to non-highest bidders, but POOL is not. Therefore, the optimal mechanism we have identified is standard if and only if $a/b \leq k_l$. Secondly, POOL always allocates, meaning it allocates with probability one, whereas SPA does not (because it does not allocate below the reserve). Therefore, the optimal mechanism always allocates if and only if $a/b \geq k_h$.

While the result above applies for any λ, we note that for the maximin ratio problem, the value of λ is endogenous, and it is not clear a priori in which information regime one falls. Quite interestingly, we can prove that the optimal maximin ratio mechanism is never in the SPA regime and thus some amount of pooling is always necessary in this case (see Section 3.1 and Proposition 2).

Methodology and closed-form characterization. We characterize the optimal mechanism and worst-case distribution in closed form via a saddle-point argument. In particular, if we assume that a saddle point exists and the optimal mechanism has the form outlined in the previous paragraph, we derive necessary conditions for Nature's worst-case distribution (Proposition 5) as well as the distributions of random reserve r and threshold τ (Proposition 7), under a few fairly mild technical conditions. We then prove that the resulting mechanism is optimal without any additional assumptions. Our methodology provides a unified treatment across all support information levels, and objectives (regret and approximation ratio) in one framework. We also characterize Nature's worst-case distribution as part of our analysis, which takes the following form: for $a/b \leq k_l$, the worst-case distribution is an isorevenue distribution (i.e., zero virtual value), whereas for $a/b > k_l$, the worst-case distribution has a constant positive virtual value in the interior of the support.[4]

Quantifying the value of scale information and competition. Using the machinery we develop, we can exactly compute the minimax regret and maximin ratio for any support $[a, b]$ and number of buyers n. Approximation ratio values are shown in Table 2.

a/b	10^{-4}	0.01	0.05	0.10	0.20	0.25	0.30	0.50	0.75	0.99
$n = 1$	0.0979	0.1784	0.2503	0.3028	0.3832	0.4191	0.4537	0.5906	0.7766	0.9900
$n = 2$	0.1086	0.2158	0.3228	0.4038	0.5197	0.5660	0.6077	0.7463	0.8841	0.9957
$n = 3$	0.1148	0.2406	0.3673	0.4529	0.5668	0.6110	0.6504	0.7779	0.9001	0.9963
$n = 4$	0.1194	0.2582	0.3884	0.4743	0.5869	0.6302	0.6684	0.7909	0.9066	0.9966
$n = 8$	0.1310	0.2836	0.4175	0.5035	0.6139	0.6556	0.6922	0.8080	0.9150	0.9969

Table 2: Maximin ratio as a function of relative support information a/b for various number of buyers n.

This table provides quantitative evidence that even a small amount of knowledge can lead to nontrivial guarantees on revenue. For example, Table 2 shows that, even when we only know that values can vary over a full order of magnitude ($a/b = 0.10$), we can guarantee 40.38% of the first-best with only 2 buyers. When the knowledge of the scale is more precise, say, the search ads cost per click in the introduction (\$2 to \$4), we get 74.63% with 2 buyers. With more agents, the guarantees improve (around 5% and 3% more, respectively, for an additional buyer). Note also that the guarantee is with respect to the first-best *full information* benchmark, a very strong (ideal) benchmark where we know the exact valuations of all agents.

Quantifying the power of mechanism features. We have identified an optimal mechanism that is a convex combination of base mechanisms in the SPA and POOL classes. The distinguishing feature of this mechanism is that it is *non-standard*, i.e., it allocates to non-highest bidders. We

[4] For a distribution with CDF F and density f, the virtual value at v is defined by $v - \frac{1-F(v)}{f(v)}$.

show that this feature is *necessary for optimality* by characterizing the minimax optimal mechanism and performance within the class of all standard mechanisms and showing that the optimal mechanism strictly improves over optimal standard mechanisms. More broadly, we quantify the value of different features in the mechanism class by computing the worst-case λ-regret (and thus, regret and ratio) for different nested mechanism subclasses of all DSIC mechanisms: all DSIC mechanisms ($\mathcal{M}_{\text{all}}$), all standard mechanisms ($\mathcal{M}_{\text{std}}$), SPA with random reserve ($\mathcal{M}_{\text{SPA-rand}}$), SPA with deterministic reserve ($\mathcal{M}_{\text{SPA-det}}$), and SPA with no reserve ($\mathcal{M}_{\text{SPA-a}}$). These results are also of independent interest, as they characterize the worst-case performance of commonly used mechanisms. We present in Table 3 the maximin ratio across mechanism classes and levels of relative support information, and observe that the power of mechanism features can be significant.

a/b	0.10	0.20	0.25	0.30	0.50	0.75	0.99
all mechanisms	0.4743	0.5869	0.6302	0.6684	0.7909	0.9066	0.9966
standard mechanisms	0.4137	0.5236	0.5684	0.6092	0.7471	0.8853	0.9958
SPA with random reserve	0.3918	0.5045	0.5517	0.5951	0.7424	0.8849	0.9958
SPA with no reserve	0.3586	0.4933	0.5457	0.5923	0.7424	0.8849	0.9958

Table 3: Maximin ratio for each mechanism class as a function of a/b, with $n = 4$ buyers.

Results extend to positively dependent and i.i.d. regular distribution classes. So far, we have assumed that the class of valuation distributions is n i.i.d. distributions. We further extend our results to other distribution classes. We show that for any mechanism class we have considered (except the class of standard mechanisms), the minimax λ-regret against i.i.d. valuations is the same as that against mixture of i.i.d. valuations, and exchangeable and affiliated distributions, two distribution classes that capture positive dependence, as well as i.i.d. regular distributions.

1.2 Related Work

Our work is related to several streams of literature.

Auction Design and Mechanism Design Vickrey [1961], Myerson [1981] and Riley and Samuelson [1981] pioneered a long line of work on the design of auctions and other economic mechanisms with strategic agents. In particular, Myerson [1981] shows that if agent valuation distributions are known, i.i.d. and *regular*, then the optimal (expected-revenue-maximizing) mechanism is a *second-price auction with reserve.* This is the classical paradigm of Bayesian mechanism design. However, in practice, we most often do not know the environment precisely, motivating the design of auctions that are *robust* to various environments.

Robust Mechanism Design We refer the reader to a recent survey Carroll [2019] for robustness in mechanism design and contracting. In particular, by requiring that the mechanism performs well in a wide range of environments, we often gain structural insights: the resulting robustly optimal mechanism is often (but not always) "simple" and "detail-free" because, in some sense, it extracts the salient features of the mechanism that are common across a range of environments.

We will highlight the work on robustness to distributions throughout the rest of this subsection because they are most related to our work, but we want to note that there are other forms of robustness as well, e.g., robustness to higher-order beliefs [Bergemann and Morris, 2005, 2013], robustness to collusion and renegotiation [Che and Kim, 2006, 2009, Carroll and Segal, 2019], and robustness to strategic behavior that is weaker than dominant strategy [Chung and Ely, 2007, Babaioff et al., 2009].

Robust mechanism design also has conceptual links to robust optimization and distributionally robust optimization; see Bertsimas et al. [2011] and Rahimian and Mehrotra [2019] for an overview.

Structures of Robust Mechanisms The closest line of work to ours is how to robustly sell an item with non-Bayesian uncertainty on valuation distributions. The one-agent case reduces to a pricing problem; Bergemann and Schlag [2008] and Eren and Maglaras [2010] provide exact characterization for minimax regret and maximin ratio pricing, respectively. Koçyiğit et al. [2020], Koçyigit et al. [2022] analyze minimax regret against any number n of agents whose valuation distributions are arbitrarily correlated with a known upper bound on the support. They show that their problem reduces to the one-agent case because Nature can choose the worst-case distribution to only have one effective bidder. Anunrojwong et al. [2022] shows that the second-price auction is robustly optimal for any number n of agents when only the upper bound of the valuations are known, for a wide range of distribution classes with positive dependence (including i.i.d.).

Optimal Mechanisms with Partial Information Our work is also related to the design of robustly optimal pricing and mechanisms with partial information about the distribution. Some works assume access to samples drawn from the i.i.d. distribution [Cole and Roughgarden, 2014, Dhangwatnotai et al., 2015, Allouah et al., 2022, Feng et al., 2021, Fu et al., 2021] while others assume that summary statistics of distributions are known [Azar et al., 2013, Suzdaltsev, 2020a,b, Bachrach and Talgam-Cohen, 2022].

SPA as a focal robust mechanism in previous works Previous works that study robust mechanism design tend to identify second-price auctions (SPA) as optimal. Anunrojwong et al. [2022] show that SPA minimizes worst-case regret for an auction with n bidders and minimal support information for a wide range of distribution classes. Bachrach and Talgam-Cohen [2022] show that SPA maximizes worst-case revenue for an auction with two i.i.d. bidders when only the mean and the upper bound of the support are known. Koçyigit et al. [2020] and Zhang [2022a] show that separate SPAs minimize worst-case regret for an auction with multiple goods and multiple bidders when only the upper bounds are known. Zhang [2022b] shows that if the seller knows only the marginal distribution of each bidder but not the joint distribution, then SPA maximizes worst-case revenue over all DSIC mechanisms in the case of two bidders and over all standard DSIC mechanisms in the case of $n \geq 3$ bidders. Che [2022] assumes only the upper bound and the mean are known but considers a different notion of robustness and shows that SPA is optimal among a class of mechanisms he calls competitive. Allouah and Besbes [2018] shows that SPA (without reserve) achieves exactly the optimal worst-case fraction of the second best benchmark revenue when the distribution has monotone hazard rate.

The upshot of this discussion is that SPA has been the focal candidate for a robust mechanism. One of the main contributions of this paper is to show why SPA fails to be optimal when we have sufficient relative support information and propose a new *building block* for robust mechanism design, the *pooling auction* mechanism (POOL). Other than the fact that the optimal mechanism in our setting is composed of these new mechanisms, this new class of mechanisms may also be of independent interest and can be useful in other robust mechanism design problems as well.

Mechanism Design and Approximation There is a long line of literature that derives proves approximation ratio guarantees for mechanisms [Hartline and Roughgarden, 2009, 2014, Feng et al., 2021, Allouah and Besbes, 2018, Hartline et al., 2020, Hartline and Johnsen, 2021]. See Roughgarden and Talgam-Cohen [2019], Hartline [2020] for surveys.

2 Problem Formulation

The seller wants to sell an indivisible object to one of n buyers. The n buyers have valuations drawn from a joint cumulative distribution $\boldsymbol{F}$. The seller does not know $\boldsymbol{F}$, and only knows the lower bound a and the upper bound b of the valuation of each buyer. That is, the seller only knows that the support of the buyers' valuations belongs to $[a,b]^n$. As discussed in the introduction, the seller chooses a mechanism to minimize the worst-case "gap" (either absolute or relative) between the mechanism revenue and benchmark revenue; we will formalize this later in this section.

Seller's Problem. We model our problem as a game between the seller and Nature, in which the seller first selects a selling mechanism from a given class $\mathcal{M}$ and then Nature may counter such a mechanism with any distribution from a given class $\mathcal{F}$. Buyers' valuations are then drawn from the distribution chosen by Nature and they participate in the seller's mechanism.

We will now consider the choice of the mechanism class $\mathcal{M}$. A selling mechanism $m = (\boldsymbol{x}, \boldsymbol{p})$ is characterized by an allocation rule $\boldsymbol{x}$ and a payment rule $\boldsymbol{p}$, where $\boldsymbol{x} : \mathbb{R}^n \to [a,b]^n$ and $\boldsymbol{p} : \mathbb{R}^n \to \mathbb{R}$. Given buyers' valuations $\boldsymbol{v} \in [a,b]^n$, $x_i(\boldsymbol{v})$ gives the probability that the item is allocated to buyer i, and $p_i(\boldsymbol{v})$ his expected payment to the seller. In our main result, we will consider the class $\mathcal{M}_{\text{all}}$ of all dominant strategy incentive compatible (DSIC) direct mechanisms. A mechanism is DSIC if and only if it is optimal for every buyer to report her true valuation (IR) and participate in the mechanism (IC), regardless of the realization of valuations of the other buyers, and the seller can allocate at most one item (AC). More formally, we require that the mechanism $m = (\boldsymbol{x}, \boldsymbol{p})$ satisfies the following constraints:

$$v_i x_i(v_i, \boldsymbol{v}_{-i}) - p_i(v_i, \boldsymbol{v}_{-i}) \geq 0, \quad \forall i, v_i, \boldsymbol{v}_{-i} \tag{IR}$$

$$v_i x_i(v_i, \boldsymbol{v}_{-i}) - p_i(v_i, \boldsymbol{v}_{-i}) \geq v_i x_i(\hat{v}_i, \boldsymbol{v}_{-i}) - p_i(\hat{v}_i, \boldsymbol{v}_{-i}) \quad \forall i, v_i, \boldsymbol{v}_{-i}, \hat{v}_i \tag{IC}$$

$$\sum_{i=1}^{n} x_i(v_i, \boldsymbol{v}_{-i}) \leq 1 \quad \forall \boldsymbol{v} \tag{AC}$$

Note that we allow the seller's mechanism to be randomized. We can now define the class of all

DSIC mechanisms

$$\mathcal{M}_{\text{all}} = \{(\boldsymbol{x}, \boldsymbol{p}) : \text{(IR), (IC), (AC)}\}\,. \tag{1}$$

Seller's Objective. Informally, the seller seeks to minimize the "gap" between the expected revenue $\mathbb{E}_{\boldsymbol{v}\sim\boldsymbol{F}}\left[\sum_{i=1}^{n} p_i(\boldsymbol{v})\right]$ relative to the benchmark associated with the revenues that could be collected when the valuations of the buyers are known $\mathbb{E}_{\boldsymbol{v}\sim\boldsymbol{F}}\left[\max(\boldsymbol{v})\right]$. We consider two notions of gaps. First is the *absolute gap*, or *regret*, defined by

$$\text{Regret}(m, \boldsymbol{F}) = \mathbb{E}_{\boldsymbol{v}\sim\boldsymbol{F}}\left[\max(\boldsymbol{v}) - \sum_{i=1}^{n} p_i(\boldsymbol{v})\right]. \tag{2}$$

Second is the *relative gap* or *approximation ratio*, defined by

$$\text{Ratio}(m, \boldsymbol{F}) = \frac{\mathbb{E}_{\boldsymbol{v}\sim\boldsymbol{F}}\left[\sum_{i=1}^{n} p_i(\boldsymbol{v})\right]}{\mathbb{E}_{\boldsymbol{v}\sim\boldsymbol{F}}\left[\max(\boldsymbol{v})\right]}. \tag{3}$$

After the seller chooses a mechanism m, Nature then chooses a distribution $\boldsymbol{F}$ from a given class of distributions $\mathcal{F}$ such that the valuation of the n agents $\boldsymbol{v} \in \mathbb{R}_+^n$ are drawn from $\boldsymbol{F}$. The seller aims to select the mechanism m to either minimize the worst-case regret or maximize the worst-case approximation ratio. Our goal, therefore, is to characterize the minimax regret and maximin ratio for different classes of mechanisms $\mathcal{M}$ and classes of distributions $\mathcal{F}$:

$$\text{MinimaxRegret}(\mathcal{M}, \mathcal{F}) := \inf_{m\in\mathcal{M}} \sup_{\boldsymbol{F}\in\mathcal{F}} \text{Regret}(m, \boldsymbol{F})\,, \tag{4}$$

$$\text{MaximinRatio}(\mathcal{M}, \mathcal{F}) := \sup_{m\in\mathcal{M}} \inf_{\boldsymbol{F}\in\mathcal{F}} \text{Ratio}(m, \boldsymbol{F}). \tag{5}$$

To unify the minimax regret and maximin ratio objectives, we will focus on the minimax λ-regret defined by

$$R_\lambda(\mathcal{M}, \mathcal{F}) := \min_{(x,p)\in\mathcal{M}} \max_{\boldsymbol{F}\in\mathcal{F}} \mathbb{E}_{\boldsymbol{v}\sim\boldsymbol{F}}\left[\lambda \max(\boldsymbol{v}) - \sum_{i=1}^{n} p_i(\boldsymbol{v})\right]. \tag{6}$$

where $\lambda \in (0, 1]$ is a constant. The following proposition, whose proof is given in Appendix A, formalizes that the values of problems (4) and (5) can be obtained from a characterization of the problem in (6).

Proposition 1. $\text{MinimaxRegret}(\mathcal{M}, \mathcal{F}) = R_1(\mathcal{M}, \mathcal{F})$ *and* $\text{MaximinRatio}(\mathcal{M}, \mathcal{F})$ *is the largest constant* $\lambda \geq 0$ *such that* $R_\lambda(\mathcal{M}, \mathcal{F}) \leq 0$.

Admissible distributions. Lastly, we consider the choice of the class of admissible distributions $\mathcal{F}$. This class can be seen as capturing the "power" of Nature: the larger the class, the more powerful/adversarial Nature becomes. To simplify exposition, we will assume for most of the paper that $\mathcal{F}$ is a class of independently and identically distributed (i.i.d.) distributions $\mathcal{F}_{\text{iid}}$, defined formally as follows.

Definition 1. *The class* $\mathcal{F}_{\text{iid}}$ *consists of all distributions such that there exists a distribution* F *with support on* $[a, b]$*, referred to as the marginal, such that* $\boldsymbol{F}(\boldsymbol{v}) = \prod_{i=1}^{n} F(v_i)$ *for every* $\boldsymbol{v} \in [a, b]^n$.

However, we will also show in Appendix G.2 that our results extend to other classes that capture positive dependence (mixtures of i.i.d. valuations $\mathcal{F}_{\text{mix}}$, and exchangeable and affiliated valuations $\mathcal{F}_{\text{aff}}$).

Saddle point approach. For each $\mathcal{M}$ and $\mathcal{F}$ considered in this work, we establish the optimality of a mechanism and characterize the associated performance via a saddle point approach.

Definition 2. $(m^*, \boldsymbol{F}^*)$ *is a* saddle point *of* $R_\lambda(m, \boldsymbol{F})$ *defined in (6) if and only if*

$$R_\lambda(m^*, \boldsymbol{F}) \leq R_\lambda(m^*, \boldsymbol{F}^*) \leq R_\lambda(m, \boldsymbol{F}^*) \quad \text{for all } m \in \mathcal{M}, \boldsymbol{F} \in \mathcal{F}.$$

Note that if $(m^*, \boldsymbol{F}^*)$ is a saddle point then $\inf_{m \in \mathcal{M}} \sup_{\boldsymbol{F} \in \mathcal{F}} R_\lambda(m, \boldsymbol{F}) \leq \sup_{\boldsymbol{F} \in \mathcal{F}} R_\lambda(m^*, \boldsymbol{F}) \leq R_\lambda(m^*, \boldsymbol{F}^*) \leq \inf_{m \in \mathcal{M}} R_\lambda(m, \boldsymbol{F}^*) \leq \inf_{m \in \mathcal{M}} \sup_{\boldsymbol{F} \in \mathcal{F}} R_\lambda(m, \boldsymbol{F})$. Therefore, $R_\lambda(m^*, \boldsymbol{F}^*) = \inf_{m \in \mathcal{M}} \sup_{\boldsymbol{F} \in \mathcal{F}} R_\lambda(m, \boldsymbol{F})$ is the minimax λ-regret (the optimal performance), m^* is an optimal mechanism, and $\boldsymbol{F}^*$ is a corresponding worst-case distribution. Therefore, it is sufficient to exhibit a saddle point and verify the inequalities in Definition 2 to obtain an optimal mechanism and its associated performance.

3 Optimal mechanisms over the class of all DSIC mechanisms

In this section, we characterize the optimal mechanism over the class of all DSIC mechanisms $\mathcal{M}_{\text{all}}$ for the minimax λ-regret problem against i.i.d. distributions $\mathcal{F}_{\text{iid}}$ for any $\lambda \in (0,1]$, support $[a,b]$, and number of bidders n. Our main theorem presents an optimal mechanism for each uncertainty regime. We will then use the main theorem to gain insights into the structure and performance of the optimal mechanism.

As discussed in the introduction, we consider the mechanism classes SPA and POOL, which are defined formally as follows.

Definition 3 (second-price and pooling auctions)**.** *A* second-price auction *with reserve r, denoted* SPA(r)*, is defined by the allocation rule $x : [a,b]^n \to [0,1]^n$ given by, for each $i \in [n]$,*

$$x_i(\boldsymbol{v}) = \begin{cases} \frac{1}{k}\mathbf{1}(v_i = \max(\boldsymbol{v})) & \textit{if } \max(\boldsymbol{v}) \geq r \textit{ and there are } k \textit{ entries in } \boldsymbol{v} \textit{ equal to } \max(\boldsymbol{v})\,, \\ 0 & \textit{if } \max(\boldsymbol{v}) < r\,. \end{cases}$$

A pooling auction *with threshold τ, denoted* POOL(τ)*, is defined by the allocation rule $x : [a,b]^n \to [0,1]^n$ given by, for each $i \in [n]$,*

$$x_i(\boldsymbol{v}) = \begin{cases} \frac{1}{k}\mathbf{1}(v_i = \max(\boldsymbol{v})) & \textit{if } \max(\boldsymbol{v}) \geq \tau \textit{ and there are } k \textit{ entries in } \boldsymbol{v} \textit{ equal to } \max(\boldsymbol{v})\,, \\ 1/n & \textit{if } \max(\boldsymbol{v}) < \tau\,. \end{cases}$$

The payment rules $p : [a,b]^n \to \mathbb{R}^n_+$ of both SPA(r) *and* POOL(τ) *are determined uniquely from Myerson's envelope formula and $p_i(v_i = a, \boldsymbol{v}_{-i}) = a x_i(v_i = a, \boldsymbol{v}_{-i})$ for every $\boldsymbol{v}_{-i}$, such that the resulting mechanism (x,p) is dominant strategy incentive compatible.*

Theorem 1 in the introduction is stated in a succinct way to highlight the structural features of our optimal mechanism. It is a corollary of the following Theorem 2 which fully characterizes the saddle point (the optimal mechanism and the worst-case distribution) and the corresponding optimal performance of the minimax λ-regret problem all in closed form.

Theorem 2 (Main Theorem, Full Saddle Characterization)**.** *Fix n and $\lambda \in (0,1]$. Define $k_l \in (0,1)$*

as a unique solution to

$$\lambda \int_{t=k_l}^{t=1} \frac{(t-k_l)^{n-1}}{t^n} dt = (1-k_l)^{n-1}$$

and $k_h \in (0,1)$ *as a unique solution to*

$$\int_{t=k_h}^{t=1} \left[\frac{(t-k_h)^{n-1}}{t^n} - (1-\lambda) \frac{(t-k_h)^n}{t^{n+1}} \right] dt = (1-k_h)^n \, .$$

(We define $k_h = 1$ *if* $n = 1$*.)*

Then we have $R_\lambda(m, \boldsymbol{F}^*) \leq R_\lambda(m^*, \boldsymbol{F}^*) \leq R_\lambda(m^*, \boldsymbol{F})$ *for any* $m \in \mathcal{M}_{\text{all}}$ *and* $\boldsymbol{F} \in \mathcal{F}_{\text{iid}}$*, where* m^* *and* $\boldsymbol{F}^*$ *(which is* n *i.i.d. with marginal* F^**) is defined depending on the value of* a/b *as follows.*

- *For* $a/b \leq k_l$*, we define* m^* *as* $\text{SPA}(\Phi^*)$ *with*

$$\Phi^*(v) = \lambda \frac{v^{n-1}}{(v-r^*)^{n-1}} \int_{t=r^*}^{t=v} \frac{(t-r^*)^{n-1}}{t^n} dt \, ,$$

and

$$F^*(v) = \begin{cases} 0 & \text{if } v \in [a, r^*] \, , \\ 1 - \frac{r^*}{v} & \text{if } v \in [r^*, b) \, , \\ 1 & \text{if } v = b \, , \end{cases}$$

with $r^* = k_l b$*. The minimax* λ*-regret is given by*

$$-(1-\lambda)b + \left[(1-k_l)^n - \lambda \int_{t=k_l}^{t=1} \left(1 - \frac{k_l}{t}\right)^n dt \right] b \, .$$

- *For* $a/b \geq k_h$*, we define* m^* *as* $\text{POOL}(\Psi^*)$ *with*

$$\Psi^*(v) = \frac{n\lambda}{n-1} \left(\frac{v - \phi_0}{v-a} \right)^n \int_{t=a}^{t=v} \frac{(t-a)^n}{(t-\phi_0)^{n+1}} dt \, ,$$

and

$$F^*(v) = \begin{cases} 1 - \frac{a-\phi_0}{v-\phi_0} & \text{if } v \in [a,b)\,, \\ 1 & \text{if } v = b\,, \end{cases}$$

where $\phi_0 = (a - k_h b)/(1-k_h) \in [0,a]$*. The minimax* λ*-regret is given by*

$$-(1-\lambda)b + \left[(1-k_h)^n + \int_{t=0}^{t=1} \frac{t^{n-1}(nk_h'^2 - \lambda t(t+k_h'))}{(t+k_h')^{n+1}} dt\right](b-a)\,.$$

where $k_h' = k_h/(1-k_h)$*.*

- *For* $k_l \le a/b \le k_h$*, there is a* unified threshold *distribution* $\mathcal{D}$ *such that if we draw a sample* $r \sim \mathcal{D}$*, if* $r \le v^*$ *the mechanism is* SPA(r)*, otherwise the mechanism is* POOL(r)*. The CDF of* $\mathcal{D}$ *is given by*

$$\mathcal{D}(v) = \begin{cases} \lambda \left(\frac{v}{v-a}\right)^{n-1} \int_{t=a}^{t=v} \frac{(t-a)^{n-1}}{t^n} dt & \text{for } v \in [a, v^*]\,, \\ -\frac{1}{n-1} + \frac{n}{n-1}\left(\frac{v}{v-a}\right)^n \left[\left(\frac{b-a}{b}\right)^n - \int_{t=v}^{t=b}\left[\frac{(t-a)^{n-1}}{t^n} - (1-\lambda)\frac{(t-a)^n}{t^{n+1}}\right] dt\right] & \text{for } v \in [v^*, b]\,, \end{cases}$$

and

$$F^*(v) = \begin{cases} 1 - \frac{a}{v} & \text{if } v \in [a,b)\,, \\ 1 & \text{if } v = b\,, \end{cases}$$

where (v^*, α) *is the unique solution to*

$$\frac{(v^*-a)^{n-1}}{(v^*)^{n-1}}(1-n\alpha) = \lambda \int_{t=a}^{t=v^*} \frac{(t-a)^{n-1}}{t^n} dt\,,$$
$$\frac{(b-a)^n}{b^n} - \frac{(v^*-a)^n}{(v^*)^n}(1-(n-1)\alpha) = \int_{t=v^*}^{t=b}\left[\frac{(t-a)^{n-1}}{t^n} - (1-\lambda)\frac{(t-a)^n}{t^{n+1}}\right] dt\,.$$

In particular, $\mathcal{D}$ *does not have a point mass, and the probability measure below* v^* *(weight of* SPA*) and above* v^* *(weight of* POOL*) are* $1 - n\alpha$ *and* $n\alpha$*, respectively. The minimax* λ*-regret is given by*

$$-(1-\lambda)b + b\left(1-\frac{a}{b}\right)^n + (b-a)\alpha\left(1 - \left(1-\frac{a}{b}\right)^{n-1}\left(1+\frac{(n-1)a}{b}\right)\right) - \lambda\int_{t=a}^{t=b}\left(1-\frac{a}{t}\right)^n dt.$$

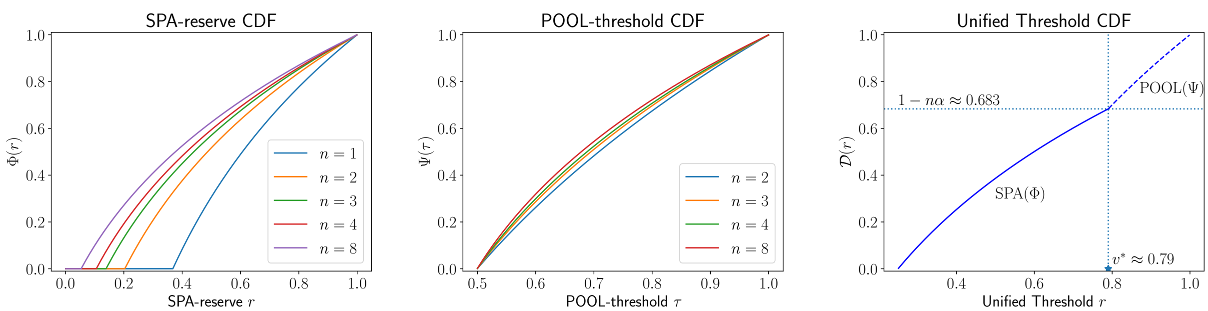
We will present the proof of Theorem 1 in Section 3.3. We now use our characterization to gain insights into the structure and performance of the optimal mechanism.

3.1 Structure of Optimal DSIC Mechanisms

3.1.1 Minimax Regret Objective

The case of minimax regret is obtained by setting $\lambda = 1$ in Theorem 2. In Figure 2, we depict optimal mechanisms for different choices of a and $b = 1$.

In Figure 2a, we depict the CDF of the optimal SPA-reserve distribution Φ as a function of n; in Figure 2b, we plot the optimal POOL-threshold distribution Ψ as a function of n.



(a) low support information $a = 0, b = 1$

(b) high support information $a = 0.5, b = 1$

(c) moderate support information $a = 0.25, b = 1$

Figure 2: Structure of the optimal mechanism in three support information regimes.

Theorem 2 states that in the moderate support information regime ($k_l \leq a/b \leq k_h$), the optimal mechanism is a randomization over SPAs with random reserves (thresholds) in $[a, v^*]$ and POOLs with random thresholds in $[v^*, b]$. We collectively call both types of thresholds "unified thresholds" such that the mechanism is characterized by the CDF of the random unified threshold $\mathcal{D}$. Figure 2c shows the example of such a mechanism (the CDF $\mathcal{D}$) for $n = 2, \lambda = 1, a = 0.25$, which is between $k_l \approx 0.2032$ and $k_h \approx 0.3162$. The unified threshold distribution $\mathcal{D}$ can be decomposed into a measure Φ over SPA-reserves on $[a, v^*]$ with weight $1 - n\alpha$ and a measure Ψ over IRON-thresholds on $[v^*, b]$ with weight $n\alpha$ for some $\alpha \in [0, 1/n]$ as in Theorem 2. The SPA and POOL parts are shown with solid and dashed lines, respectively. The two parts meet at boundary v^* with CDF value $1 - n\alpha$, the total measure of Φ.

3.1.2 Maximin Ratio Objective

Unlike the minimax regret case where $\lambda = 1$ is set exogenously, here λ is obtained from bisection search to find the value of λ such that the minimax λ-regret is zero (cf. Proposition 1) and $\lambda = \lambda^*(k, n)$ is a function of $k \equiv a/b$, i.e., the maximin ratio given k that we computed earlier. As a result, the regime is determined endogeously. For each n, we then compare $k = a/b$ with $k_l = k_l(\lambda^*(k, n), n)$ and $k_h = k_h(\lambda^*(k, n), n)$ to determine the regime. We can prove that the pure SPA regime is *never* possible for the maximin ratio objective; we formalize this in Proposition 2, which is proven in Appendix B.1. For reasonable values of a/b,[5] the optimal mechanisms are in the regime of pure pooling auctions. For a POOL-threshold τ, we define the *normalized* threshold as $\tilde{\tau} = (\tau - a)/(b - a)$. As $\tau \in [a, b]$, we have $\tilde{\tau} \in [0, 1]$, so this normalization allows us to compare the shape of threshold distributions on the same scale. For $n \in \{2, 4\}$ and $a/b \in \{0.10, 0.25, 0.50, 0.75, 0.99\}$, we plot the normalized POOL-threshold CDFs in Figure 3. We see that for low a/b, the distribution

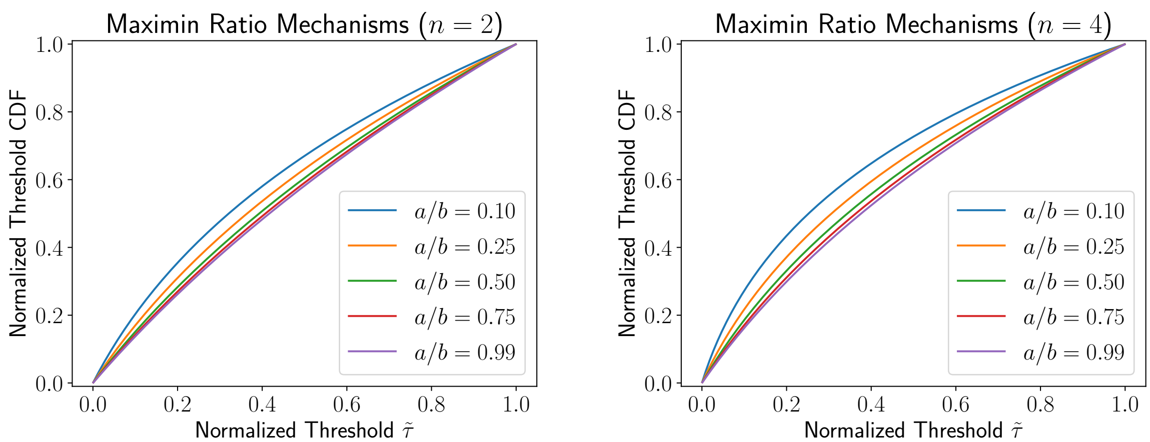


Figure 3: Normalized threshold distributions for $n \in \{2, 4\}$ and $a/b \in \{0.10, 0.25, 0.50, 0.75, 0.99\}$

puts more weight on lower thresholds, and vice versa.

3.2 Remark on the case $n = 1$

An important corollary of Theorem 2 is the pricing case (one-bidder / no-competition). Applying the result with $n = 1$ directly recovers the minimax regret result of Bergemann and Schlag [2008]

[5] For $a/b \geq 0.0978$ for $n = 2$, $a/b \geq 0.0155$ for $n = 3$, and $a/b \geq 0.0035$ for $n = 4$ we are in the pure POOL regime.

and the maximin ratio result of Eren and Maglaras [2010] as special cases.[6] We give the proof of this corollary in Appendix D.4.

Corollary 1 (Pricing)**.** *Suppose $n = 1$. The minimax λ-regret is $\lambda e^{-1/\lambda} b$ for $a/b \leq e^{-1/\lambda}$ and $-a + \lambda a + \lambda \log(b/a)$ for $a/b \geq e^{-1/\lambda}$. In particular, the minimax regret is b/e if $a/b \leq 1/e$ and $a \log(b/a)$ if $a/b \geq 1/e$, while the maximin ratio is $1/(1 + \log(b/a))$.*

3.3 Outline of the Proof of the Main Theorem

The key idea of proof of our main theorem (Theorem 2) is to explicitly exhibit a saddle point of the zero-sum game between seller and Nature, as discussed at the end of Section 2, that is, to find $m^* \in \mathcal{M}_{\text{all}}$ and $\boldsymbol{F}^* \in \mathcal{F}_{\text{iid}}$ such that $R_\lambda(m^*, \boldsymbol{F}) \leq R_\lambda(m^*, \boldsymbol{F}^*) \leq R_\lambda(m, \boldsymbol{F}^*)$ for all $m \in \mathcal{M}_{\text{all}}$ and $\boldsymbol{F} \in \mathcal{F}_{\text{iid}}$. We proceed in two steps. First, we assume that the mechanism is a randomization over SPA and POOL. Then, under certain regularity assumptions, we can pin down the weights of each component SPA(r) and POOL(τ) of the mechanism, and thus the mechanism itself. We can also pin down the worst-case distribution $\boldsymbol{F}^*$. Second, we formally verify, with no additional assumptions, that the pair $(m^*, \boldsymbol{F}^*)$ we derived is indeed a saddle point.

We now outline the first step in more detail. In the full proof, it will be more convenient mathematically to consider the class of (g_u, g_d) mechanisms, parameterized by two functions $g_u, g_d : [a, b] \to [0, 1]$, where if the valuation vector is $\boldsymbol{v}$ with highest $v_{\max} := \max_{i \in [n]} v_i$, then the highest bidder is allocated with probability $g_u(v_{\max})$ and the other bidders are allocated with probability $g_d(v_{\max})$ each. Every convex combination over SPA and POOL mechanisms with random thresholds can be expressed in terms of (g_u, g_d) (see Propositions 3 and 8). Therefore, the (g_u, g_d) formalism will allow us to treat all three regimes in a unified manner. However, for concreteness, in the main text below we will focus on the high support information regime and work directly with the threshold distribution Ψ for illustration; the proof of the moderate support information is more challenging but follows a similar outline.

Assume that $\boldsymbol{F}$ is i.i.d. with marginal F. We can compute the λ-regret $R_\lambda(\Psi, F)$ of POOL(Ψ) and show that it has two representations using an ad-hoc integration by par result that handles distributions that are not necessarily smooth (cf. Proposition 4). Suppose that Ψ is arbitrary,

[6] More precisely, Eren and Maglaras [2010] derives the maximin ratio against the second-best benchmark, whereas our result is against the first-best benchmark. However, in Appendix I we show that for the $n = 1$ case, the maximin ratio for two benchmarks are the same.

while F has a density in the interior with point masses $F(a)$ and $f_b = 1 - F(b^-)$ at a and b, then

$$\begin{aligned} R_\lambda(\Psi, F) &= \lambda b - aF(a)^n - b + b(1-f_b)^n \\ &+ \int_{v=a}^{v=b} \Bigg[\left\{-\lambda F(v)^n + nF(v)^{n-2}F'(v)(v-a) - nF(v)^{n-1}vF'(v)\right\} \\ &+ \underbrace{nF(v)^{n-2}\left\{F(v) - F(v)^2 - (v-a)F'(v)\right\}}_{\text{coefficient of } \Psi(v)} \Psi(v) \Bigg] dv \end{aligned} \qquad \text{(Regret-}\Psi\text{)}$$

Note that (Regret-Ψ) depends on Ψ only through $\Psi(v)$ and is linear in Ψ. This representation is useful for seller's saddle $\inf_m R_\lambda(m, F^*)$, because we do not make any assumption on seller's choice Ψ. By the first-order conditions, under the worst-case distribution F^*, the coefficient of each $\Psi^*(v)$ is zero wherever Ψ^* is interior. Otherwise, the seller could decrease his regret by changing the distribution of reserves. Therefore,

$$F^*(v) - F^*(v)^2 - (v-a)(F^*)'(v) = 0 \Rightarrow \frac{d}{dv}\left(v - \frac{v-a}{F^*(v)}\right) = 0 \Rightarrow v - \frac{v-a}{F^*(v)} = \phi_0\,.$$

This pins down Nature's candidate distribution as $F^*(v) = (v-a)/(v-\phi_0)$, a distribution with constant virtual value ϕ_0.

Alternatively, let F be arbitrary but assume that Ψ is differentiable, then

$$\begin{aligned} R_\lambda(\Psi, F) &= a(\lambda - 1) + \int_{v=a}^{v=b} \left[\lambda - \Psi(v) - (v-a)\Psi'(v)\right] dv \\ &+ \int_{v=a}^{v=b} \underbrace{\left(-\lambda - (n-1)\Psi(v)\right) F(v)^n + \left(n\Psi(v) + (v-a)\Psi'(v)\right) F(v)^{n-1}}_{:=I(F(v),v)} dv \end{aligned} \qquad \text{(Regret-}F\text{)}$$

Note that (Regret-F) depends on F only through $F(v)$ and is "separable" (different $F(v)$ terms do not interact). This representation is useful for Nature's saddle $\sup_F R_\lambda(m^*, F)$, because we do not make any assumption on Nature's choice F and we can optimize the integrand pointwise as an expression in $F(v)$, i.e., we can equivalently solve $\sup_{F(v)} I(F(v), v)$ for each v. Taking derivatives with respect to $F(v)$, the first-order condition on $F(v)$ gives

$$\left(-\lambda - (n-1)\Psi(v)\right) nF(v)^{n-1} + \left(n\Psi(v) + (v-a)\Psi'(v)\right)(n-1)F(v)^{n-2} = 0\,,$$

because, at the saddle, this first derivative must be zero. Substituting $F^*(v) = (v-a)/(v-\phi_0)$ gives

$$(\Psi^*)'(v) + \frac{n(a-\phi_0)}{(v-\phi_0)(v-a)}\Psi^*(v) = \frac{n\lambda}{(n-1)(v-\phi_0)}, \tag{ODE-Ψ}$$

or

$$\frac{d}{dv}\left[\left(\frac{v-a}{v-\phi_0}\right)^n \Psi^*(v)\right] = \frac{n\lambda}{n-1}\frac{(v-a)^n}{(v-\phi_0)^{n+1}}.$$

Because $\left(\frac{v-a}{v-\phi_0}\right)^n \Psi^*(v)$ is zero at $v=a$, integrating from a to v gives

$$\left(\frac{v-a}{v-\phi_0}\right)^n \Psi^*(v) = \frac{n\lambda}{n-1}\int_{t=a}^{t=v}\frac{(t-a)^n}{(t-\phi_0)^{n+1}}dt.$$

Performing the integration, we can rewrite $\Psi^*(v)$ as

$$\Psi^*(v) = \frac{n\lambda}{n-1}\sum_{k=n+1}^{\infty}\frac{(v-a)^{k-n}}{k(v-\phi_0)^{k-n}}.$$

This expression makes it clear that $\Psi^*(a) = 0$, so Ψ^* is well-behaved and does not have a point mass at a, and that $\Psi^*(v)$ is increasing in v. For Ψ^* to be feasible, the only additional condition we need is $\Psi^*(b) \leq 1$. By Nature's saddle, Ψ^* cannot have a point mass at b, so $\Psi^*(b) = 1$, which gives the following equation that determines ϕ_0:

$$\left(\frac{b-a}{b-\phi_0}\right)^n = \frac{n\lambda}{n-1}\int_{t=a}^{t=b}\frac{(t-a)^n}{(t-\phi_0)^{n+1}}dt\,.$$

By definition of k_h, we see by inspection that this equation has an explicit solution

$$\phi_0 = \frac{a-k_h b}{1-k_h}\,.$$

We need $\phi_0 \geq 0$ for Nature's saddle to hold: this is why we need $a/b \geq k_h$ in the pure POOL regime.

Now that we have identified a candidate saddle-point, we need to formally verify its optimality. Seller's saddle (optimize over m with fixed F^*) is a standard Bayesian mechanism design problem and the optimality of POOL(Ψ^*) follows because every mechanism that always allocates is optimal since F^* has positive, constant virtual value. For Nature's saddle (optimize over F with fixed m^*), we need to check that F^* maximizes the regret given the mechanism POOL(Ψ^*). Beyond the first-order conditions that we have already checked, we also need that the second-order derivative is nonpositive, namely,

$$(-\lambda - (n-1)\Psi^*(v))\, n(n-1)F^*(v)^{n-2} + \big(n\Psi(v) + (v-a)(\Psi^*)'(v)\big)\,(n-1)(n-2)F^*(v)^{n-3} \le 0\,,$$

where we interpret the term $(n-2)F^*(v)^{n-3}$ as zero for $n = 2$. Substituting the expression for $(\Psi^*)'(v)$ from (ODE-Ψ) reduces the above inequality to

$$(-\lambda - (n-1)\Psi^*(v))\, nF^*(v) + \frac{n(v-a)}{(v-\phi_0)}\Psi^*(v)(n-2) \le 0\,.$$

Substituting $F^*(v) = (v-a)/(v-\phi_0)$ reduces the above inequality further to

$$\frac{n(v-a)}{(v-\phi_0)}(-\lambda - \Psi^*(v)) \le 0\,,$$

which is trivially true. We therefore have established a saddle point (POOL(Ψ^*), F^*) of the problem in the high relative support information case.

We give a more detailed outline of *all cases* of the proof of the main theorem in Appendix C. Details omitted from Appendix C are given in Appendix D.

4 Minimax λ-Regret across Mechanism Classes

Our main theorem (Theorem 2) gives a complete characterization of the optimal robust performance when Nature's distribution is i.i.d. ($\boldsymbol{F} \in \mathcal{F}_{\text{iid}}$) and the seller can choose any DSIC mechanism ($m \in \mathcal{M}_{\text{all}}$). It turns out that the optimal mechanism is generally a randomization over SPA and POOL mechanisms. This optimal mechanism has interesting features, and we would like to quantify how much each feature contributes to the performance. That is, without that feature, how much (robust) performance, if any, we will lose. Equivalently, our results quantify the "cost of simplicity" or the performance loss if the seller is restricted to simpler classes of mechanisms. We formalize this problem by solving minimax λ-regret problems, $\lambda \in (0,1]$, when the mechanism classes $\mathcal{M}$ are successively smaller, omitting one feature at a time. The subclasses under consideration are shown in Figure 4.

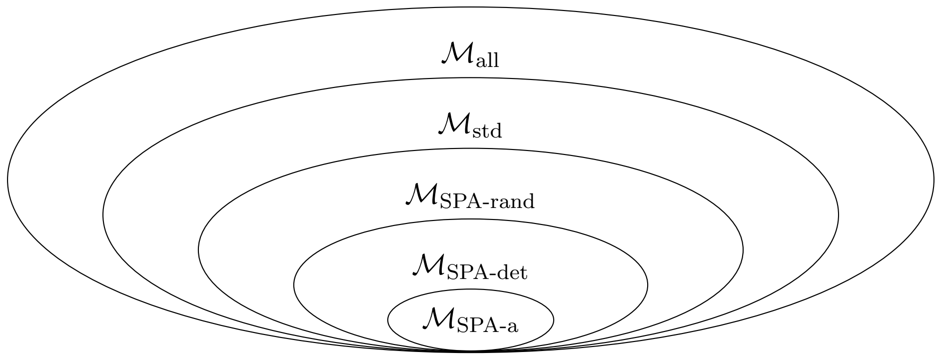


Figure 4: Nested mechanism subclasses we consider, from biggest to smallest: DSIC mechanisms ($\mathcal{M}_{\text{all}}$), standard mechanisms ($\mathcal{M}_{\text{std}}$), SPA with random reserve ($\mathcal{M}_{\text{SPA-rand}}$), SPA with deterministic reserve ($\mathcal{M}_{\text{SPA-det}}$), SPA with no reserve ($\mathcal{M}_{\text{SPA-a}}$)

First, our optimal mechanism is not standard because POOL might allocate to a bidder who is not the highest. To isolate the role of the pooling feature, we study the class of standard mechanisms that only allocate to the maximum bidder. Second, we study the need to deviate from SPAs in standard mechanisms, and hence study SPAs with randomized reserves. Lastly, we quantify the power of randomness and the power of using a reserve by computing minimax regret under the class $\mathcal{M}_{\text{SPA-det}}$ of SPA with a deterministic reserve and the class $\mathcal{M}_{\text{SPA-a}}$ of SPA with no reserve. Interestingly, we show that there are strict separations in terms of maximin ratio between $\mathcal{M}_{\text{all}}$, $\mathcal{M}_{\text{std}}$, $\mathcal{M}_{\text{SPA-rand}}$, and $\mathcal{M}_{\text{SPA-a}}$ (but not between $\mathcal{M}_{\text{SPA-det}}$ and $\mathcal{M}_{\text{SPA-a}}$). In other words, pooling and deviations from SPAs are critical for robust performance, and so is the randomization of reserve prices.

4.1 Minimax λ-Regret Over Standard Mechanisms

A mechanism is said to be *standard* if it never allocates to an agent that does not have the highest value. Formally, it satisfies the following constraint:

$$x_i(v_i, \boldsymbol{v}_{-i}) = 0 \quad \forall i, v_i, \boldsymbol{v}_{-i} \text{ such that } v_i < \max(\boldsymbol{v})\,. \tag{STD}$$

We can now define the class of all standard mechanisms.

Definition 4. *The class of all standard mechanisms is given by*

$$\mathcal{M}_{\text{std}} = \{(\boldsymbol{x}, \boldsymbol{p}) : \text{(IR), (IC), (AC), (STD)}\} . \tag{7}$$

It is clear that any second-price auction (SPA) with random reserve is standard, and intuitively, SPAs seem like "natural" and "typical" elements of this class, but as it turns out, other standard mechanisms are more robust than SPAs when relative support information is high. We now introduce the following mechanism class.

Definition 5 (Generous SPA)**.** *A* generous SPA with reserve distribution Φ*, denoted* $\text{GenSPA}(\Phi)$*, is defined by the allocation rule x given by, for each $i \in [n]$,*

$$x_i(\boldsymbol{v}) = \begin{cases} \Phi(v^{(1)}) & \text{if } v_i \text{ is the highest and } v^{(2)} > a\,, \\ 1 & \text{if } v_i \text{ is the highest and } v^{(2)} = a\,, \end{cases}$$

and zero otherwise, breaking ties uniformly at random. The payment rule $p : [a, b]^n \to \mathbb{R}_+^n$ is determined uniquely from Myerson's formula such that the resulting mechanism (x, p) is dominant strategy incentive compatible.

We call this mechanism *generous* SPA because it behaves like SPA, except in the case when all other non-highest agents have the lowest possible value a, then it always allocates ("generously"). We now state the main theorem of this section.

Theorem 3 (Optimal Standard Mechanism)**.** *Fix n and $\lambda \in (0, 1]$, and let $\tilde{a} = a/b \in [0, 1)$. Define k_l as in Theorem 2. Then, the problem admits an optimal minimax λ-regret standard mechanism m^*, depending on a/b as follows.*

- *(Low Relative Support Information) For $a/b \le k_l$, $m^* = \text{SPA}(\Phi)$ is the same as in Theorem 2.*
- *(High Relative Support Information) For $a/b \ge k_l$, there is a probability distribution Φ such that $m^* = \text{GenSPA}(\Phi)$.*

Note that by Theorem 2, if $a/b \le k_l$, then SPA with random reserve is optimal in $\mathcal{M}_{\text{all}}$, and it is also standard, so it is immediate that it is also optimal in the class $\mathcal{M}_{\text{std}}$. Similar to Theorem 1, Theorem 3 highlights the structural features of our optimal mechanism and is a corollary of Theorem 6 in the Appendix which fully characterizes the saddle point in closed form.

The proof of Theorem 3 follows a similar outline to that of Theorem 2, although the calculations are nontrivial. In particular, we need to derive the expressions of *conditional* distributions of order statistics for arbitrary F, taking into account potential ties, which complicate the calculations.[7] In contrast, the regret of any (g_u, g_d) mechanism (whose class contains all other mechanisms in this paper) depends only on the marginal distributions of the first- and second-order statistics, which are simpler (cf. Proposition 4). However, the hardest part is *coming up with the right structural class* GenSPA that contains the optimal mechanism and is tractable, because the ODE techniques can pin down the candidate mechanism only once we fix the mechanism up to a one-dimensional functional parameter. We discuss key technical challenges and give the full proof in Appendix E.1.

4.2 Minimax λ-Regret over SPA with random reserve and deterministic reserve

We can characterize the minimax λ-regret mechanism and its corresponding worst-case distribution and performance in the following theorem.

Theorem 4 (Optimal SPA with Random Reserve)**.** *Fix n and $\lambda \in (0, 1]$. Define k_l as in Theorem 2 and $k'_h = \lambda n/((1+\lambda)n-1)$. Then, the problem admits a minimax λ-regret $m^* = \mathrm{SPA}(\Phi^*)$, depending on a/b, as follows.*

- *(Low Relative Support Information) For $a/b \leq k_l$, $m^* = \mathrm{SPA}(\Phi^*)$ is the same as in Theorem 2.*
- *(High Relative Support Information) For $a/b \geq k'_h$, Φ^* is point mass only at a, i.e., $m^* = \mathrm{SPA}(a)$ is a SPA with no reserve.*
- *(Moderate Relative Support Information) For $k_l \leq a/b \leq k'_h$, there is $r^* \in [a, b]$ such that Φ^* has a point mass at a and a density on $[r^*, b]$.*

The second bullet point of Theorem 4 formalizes the intuition highlighted in the introduction that in the high scale information regime (a/b is close enough to 1), the optimal SPA with random

[7]The existing results on conditional distributions of order statistics assume that F has a density, see e.g. David and Nagaraja [2003]. These results do not apply because we do not make any assumptions on F. In fact, the worst case F has point masses.

reserve sets no reserve at all. Similar to Theorem 1, Theorem 4 highlights the structural features of our optimal mechanism and is a corollary of Theorem 7 in Appendix E.2 which fully characterizes the saddle point in closed form. The proof of the moderate information regime of Theorem 4 is the most challenging. It is different from previous saddle problems because in this case, the increasing condition on Φ is *binding*; if we optimize pointwise, the resulting distribution is not increasing, which is infeasible. We characterize an optimal distribution of reserves using a Lagrangian approach that involves introducing a Lagrange multiplier for the monotonicity constraint and then designing a primal-dual pair that satisfies complementary slackness and Lagrangian optimality. We discuss key technical challenges and give the full proof in Appendix E.2.

Lastly, we characterize the optimal SPA with deterministic reserve $\mathcal{M}_{\text{SPA-det}}$ and SPA with no reserve $\mathcal{M}_{\text{SPA-a}}$. Proposition 11 in the Appendix gives the the minimax λ-regret for SPA(r) with a fixed deterministic reserve r. In particular, it subsumes the problem of choosing the regret-minimizing reserve r as well as computing worst-case regret of SPA without reserve ($r = a$).

4.3 Performance Separation Between Mechanism Classes

Table 3 in the introduction shows the maximin ratio as a function of a/b of all mechanism classes we consider for $n = 4$. Figure 5 graphically depicts the same information for $n \in \{2, 4\}$. This metric captures the performance of the optimal mechanism. We can see that while $\mathcal{M}_{\text{SPA-det}}$ and $\mathcal{M}_{\text{SPA-a}}$ have the same maximin ratios (so a fixed reserve does not improve over no reserve), there are *strict separations* between $\mathcal{M}_{\text{all}}$, $\mathcal{M}_{\text{std}}$, $\mathcal{M}_{\text{SPA-rand}}$, and $\mathcal{M}_{\text{SPA-a}}$.

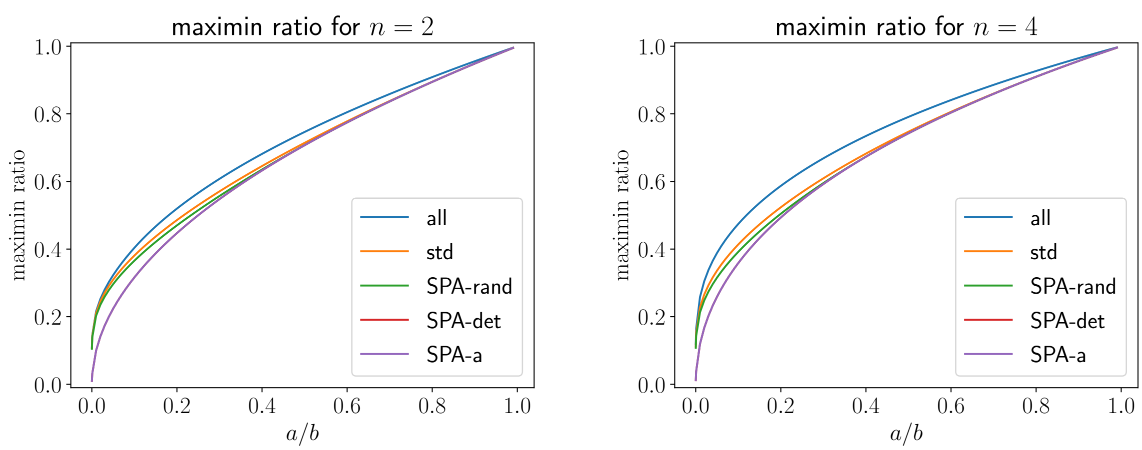


Figure 5: Maximin ratio as a function of a/b for $n \in \{2, 4\}$.

The gap between $\mathcal{M}_{\text{std}}$ and $\mathcal{M}_{\text{SPA-rand}}$ shows that no SPA is optimal within the class of standard mechanisms, even though the gap is quantitatively small. In contrast, the gap between $\mathcal{M}_{\text{all}}$ and $\mathcal{M}_{\text{std}}$ is significant. This means that in robust settings, *it is important to sometimes allocate to non-highest bidders*. We can see from the plots with $n = 2$ and $n = 4$ that the non-standard gap becomes bigger and dominates all other gaps as n gets large, so this becomes more important with more bidders.

Echoing discussions in the introduction, our work shows that there are interesting mechanism classes in DSIC mechanisms beyond SPA in the sense that they are robustly optimal in natural settings. Interestingly, SPA is not optimal even within the class of standard mechanisms; GenSPA is. It is an open question whether GenSPA will also be useful in other settings as well.

We can also use analytical results derived in this section to gain insights into the structure of optimal mechanisms within subclasses. See Appendix F.1 for details.

5 Extensions and Conclusion

In this paper, we give an explicit characterization of a *robustly* optimal mechanism to sell an item to n buyers knowing only the lower bound and the upper bound of the support, where the seller's performance is evaluated in the worst case. Our general framework is broadly applicable to an arbitrary number n of buyers and several mechanism classes $\mathcal{M}$ and captures both regret and ratio objectives.

Furthermore, we provide the following extensions of the framework. First, in Appendix G, we establish that the minimax λ-regret we have obtained for the case of i.i.d. distributions (and the corresponding optimal mechanism) does not change if Nature optimizes over broader classes of distributions capturing positive correlation: exchangeable and affiliated values, a common class considered with knowledge of the distributions [Milgrom and Weber, 1982]; and mixtures of i.i.d. distributions, another common class. The results also do not change if Nature optimizes over the smaller class of i.i.d. regular distributions. Second, our results also have sharp implications on the worst-case gap between first-best and second-best revenues with full information on the distribution of values. We formalize this in Appendix H.

There are many avenues for future work. This work is a step in the more general agenda of *robust mechanism design with partial information*, and it would be interesting to investigate how other forms of side information (such as moments, samples, and shapes of distributions) impact the structure and performance of mechanisms, and the value of such information. Another direction is to consider other benchmarks, especially the second-best benchmark rather than the first-best benchmark considered in this paper.

References

Amine Allouah and Omar Besbes. Prior-independent optimal auctions. In *Proceedings of the 2018 ACM Conference on Economics and Computation*, EC '18, New York, NY, USA, 2018. Association for Computing Machinery.

Amine Allouah, Achraf Bahamou, and Omar Besbes. Pricing with samples. *Operations Research*, 2022.

Jerry Anunrojwong, Santiago R. Balseiro, and Omar Besbes. On the robustness of second-price auctions in prior-independent mechanism design. In *Proceedings of the 23rd ACM Conference on Economics and Computation*, pages 151–152. ACM, 2022.

Pablo Azar, Constantinos Daskalakis, Silvio Micali, and S. Matthew Weinberg. Optimal and efficient parametric auctions. *Proceedings of the 24th annual ACM-SIAM symposium on Discrete algorithms*, 2013.

Moshe Babaioff, Ron Lavi, and Elan Pavlov. Single-value combinatorial auctions and algorithmic implementation in undominated strategies. *Journal of the ACM*, 2009.

Nir Bachrach and Inbal Talgam-Cohen. Distributional robustness: From pricing to auctions. In *Proceedings of the 23rd ACM Conference on Economics and Computation*, pages 182–183. ACM, 2022.

Dirk Bergemann and Stephen Morris. Robust mechanism design. *Econometrica*, 2005.

Dirk Bergemann and Stephen Morris. An introduction to robust mechanism design. *Foundations and Trends in Microeconomics*, 2013.

Dirk Bergemann and Karl H. Schlag. Pricing without priors. *Journal of the European Economic Association*, 2008.

Dimitris Bertsimas, David B. Brown, and Constantine Caramanis. Theory and applications of robust optimization. *SIAM Review*, 2011.

Gabriel Carroll. Robustness in mechanism design and contracting. *Annual Review of Economics*, 2019.

Gabriel Carroll and Ilya Segal. Robustly optimal auctions with unknown resale opportunities. *The Review of Economic Studies*, 2019.

Ethan Che. Robustly optimal auction design under mean constraints. In *Proceedings of the 23rd ACM Conference on Economics and Computation*, pages 153–181. ACM, 2022.

Yeon-Koo Che and Jinwoo Kim. Robustly collusion-proof implementation. *Econometrica*, 2006.

Yeon-Koo Che and Jinwoo Kim. Optimal collusion-proof auctions. *Journal of Economic Theory*, 2009.

Kim-Sau Chung and J. C. Ely. Foundations of dominant-strategy mechanism. *The Review of Economic Studies*, 2007.

Richard Cole and Tim Roughgarden. The sample complexity of revenue maximization. *Proceedings of the 46th Annual ACM Symposium on Theory of Computing*, 2014.

Herbert A. David and Haikady N. Nagaraja. *Order Statistics 3rd ed.* Wiley-Interscience, 2003. ISBN 978-0471389262.

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

Serkan S Eren and Costis Maglaras. Monopoly pricing with limited demand information. *Journal of revenue and pricing management*, 9(1-2):23–48, 2010.

Yiding Feng, Jason D Hartline, and Yingkai Li. Revelation gap for pricing from samples. *Proceedings of the 53rd Annual ACM SIGACT Symposium on Theory of Computing*, 2021.

Hu Fu, Nima Haghpanah, Jason Hartline, and Robert Kleinberg. Full surplus extraction from samples. *Journal of Economic Theory*, 2021.

Jason Hartline. *Mechanism Design and Approximation.* 2020.

Jason Hartline and Aleck Johnsen. Lower bounds for prior independent algorithms. Technical report, 2021.

Jason Hartline and Tim Roughgarden. Optimal platform design. Technical report, 2014.

Jason Hartline, Aleck Johnsen, and Yingkai Li. Benchmark design and prior-independent optimization. *Proceedings of the 2020 IEEE 61st Annual Symposium on Foundations of Computer Science (FOCS)*, 2020.

Jason D. Hartline and Tim Roughgarden. Simple versus optimal mechanisms. *Proceedings of the 11th ACM Conference on Economics and Computation*, 2009.

Çağıl Koçyiğit, Garud Iyengar, Daniel Kuhn, and Wolfram Wiesemann. Distributionally robust mechanism design. *Management Science*, 2020.

Çagil Koçyigit, Daniel Kuhn, and Napat Rujeerapaiboon. Regret minimization and separation in multi-bidder multi-item auctions. Technical report, 2020.

Çagil Koçyigit, Napat Rujeerapaiboon, and Daniel Kuhn. Robust multidimensional pricing: separation without regret. *Mathematical Programming*, 196(1):841–874, 2022.

Kristen McCormick. 2022 Google Ads and Microsoft Ads Benchmarks for Every Industry (New Data!). https://www.wordstream.com/blog/ws/2022/05/18/search-advertising-benchmarks, 2022. Accessed: 2022-10-03.

Paul R. Milgrom and Robert J. Weber. A theory of auctions and competitive bidding. *Econometrica*, 1982.

Paulo Klinger Monteiro and Benar Fux Svaiter. Optimal auction with a general distribution: Virtual valuation without densities. *Journal of Mathematical Economics*, 2010.

Roger B. Myerson. Optimal auction design. *Mathematics of Operations Research*, 1981.

Hamed Rahimian and Sanjay Mehrotra. Distributionally robust optimization: A review. Technical report, 2019.

John G. Riley and William F. Samuelson. Optimal auctions. *American Economic Review*, 1981.

Tim Roughgarden and Inbal Talgam-Cohen. Approximately optimal mechanism design. *Annual Review of Economics*, 11(1):355–381, 2019. doi: 10.1146/annurev-economics-080218-025607. URL https://doi.org/10.1146/annurev-economics-080218-025607.

Alex Suzdaltsev. An optimal distributionally robust auction. Technical report, 2020a.

Alex Suzdaltsev. Distributionally robust pricing in independent private value auctions. Technical report, 2020b.

William Vickrey. Counterspeculation, auctions, and competitive sealed tenders. *The Journal of Finance*, 1961.

Robert Wilson. Game-theoretic analyses of trading processes in advanced in economic theory. In Truman Fassett Bewley, editor, *Advances in Economic Theory Fifth World Congress*, chapter 2, pages 33–70. Cambridge University Press, 1987.

Wanchang Zhang. Auctioning multiple goods without priors. Technical report, 2022a.

Wanchang Zhang. Correlation-robust optimal auctions. Technical report, 2022b.

A Proofs and Discussions from Section 2

Proof of Proposition 1. The definition of MaximinRatio$(\mathcal{M}, \mathcal{F})$ says that it is a solution to

$$\max_{(x,p)\in\mathcal{M}} \lambda \text{ s.t.}$$
$$\frac{\mathbb{E}_{\boldsymbol{v}\sim\boldsymbol{F}}\left[\sum_{i=1}^n p_i(\boldsymbol{v})\right]}{\mathbb{E}_{\boldsymbol{v}\sim\boldsymbol{F}}\left[\max(\boldsymbol{v})\right]} \leq \lambda \quad \forall \boldsymbol{F} \in \mathcal{F}$$

or

$$\max_{(x,p)\in\mathcal{M}} \lambda \text{ s.t.}$$
$$\mathbb{E}_{\boldsymbol{v}\sim\boldsymbol{F}}\left[\lambda \max(\boldsymbol{v}) - \sum_{i=1}^n p_i(\boldsymbol{v})\right] \leq 0 \quad \forall \boldsymbol{F} \in \mathcal{F}$$

or

$$\max_{(x,p)\in\mathcal{M}} \lambda \text{ s.t.}$$
$$\max_{\boldsymbol{F}\in\mathcal{F}} \mathbb{E}_{\boldsymbol{v}\sim\boldsymbol{F}}\left[\lambda \max(\boldsymbol{v}) - \sum_{i=1}^n p_i(\boldsymbol{v})\right] \leq 0$$

That is, the maximin ratio is the highest value of λ such that *there exists* $(x,p) \in \mathcal{M}$ that $\max_{\boldsymbol{F}\in\mathcal{F}} \mathbb{E}_{\boldsymbol{v}\sim\boldsymbol{F}}\left[\lambda \max(\boldsymbol{v}) - \sum_{i=1}^n p_i(\boldsymbol{v})\right] \leq 0$. Equivalently,

$$R_{\lambda,n}(\mathcal{M}, \mathcal{F}) = \min_{(x,p)\in\mathcal{M}} \max_{\boldsymbol{F}\in\mathcal{F}} \mathbb{E}_{\boldsymbol{v}\sim\boldsymbol{F}}\left[\lambda \max(\boldsymbol{v}) - \sum_{i=1}^n p_i(\boldsymbol{v})\right] \leq 0.$$

□

B Proofs and Discussions from Section 3

We note that the worst-case distributions identified in our main theorem (Theorem 2) belong to either the isorevenue or the constant virtual value family, which are defined as follows.

Definition 6 (Isorevenue and Constant Virtual Value Distributions)**.** *For* $r^* \in [a, b)$*, an* isorevenue distribution *on* $[r^*, b]$ *with constant* c*, denoted* IsoRev$(r^*; c)$*, is a distribution* F *supported on* $[r^*, b]$

with $F(v) = 1 - c/v$ *for* $v \in [r^*, b)$ *and* $F(b) = 1$*. If* $c = r^*$*, we slightly abuse the notation and write* $\text{IsoRev}(r^*) := \text{IsoRev}(r^*; r^*)$*. For a constant* ϕ_0*, a* constant virtual value distribution *on* $[a, b]$ *with virtual value* ϕ_0*, denoted* $\text{ConstVirt}(\phi_0)$*, is a distribution* F *supported on* $[a, b]$ *with* $F(v) = 1 - (a - \phi_0)/(v - \phi_0)$ *for* $v \in [a, b)$ *and* $F(b) = 1$*.*

Recall that whenever F has a density f at v, the virtual value of F at v is defined as $v - (1 - F(v))/f(v)$. It is immediate that $\text{ConstVirt}(\phi_0)$ has virtual value ϕ_0 for all $v \in [a, b)$, as expected. $\text{IsoRev}(r^*; c)$ is called isorevenue because if this is the valuation distribution, then for any posted price $p \in [r^*, b)$, the revenue is $p(1 - F(p)) = c$, a constant. Note that an isorevenue is simply a distribution with zero virtual value: $\text{ConstVirt}(0) = \text{IsoRev}(a)$.

Throughout this section, we will use the following technical lemma.

Lemma 1. *Let* ϕ_0 *be a constant, then for any positive integer* n *and* $v \geq a$ *we have the identity*

$$\int_{t=a}^{t=v} \frac{(t-a)^{n-1}}{(t-\phi_0)^n} dt = \log\left(\frac{v-\phi_0}{a-\phi_0}\right) - \sum_{k=1}^{n-1} \frac{(v-a)^k}{k(v-\phi_0)^k} = \sum_{k=n}^{\infty} \frac{(v-a)^k}{k(v-\phi_0)^k}$$

Proof of Lemma 1. We first check the equality of the first and the second expressions. Note that both expressions are zero when $v = a$. It is then sufficient to check that the derivatives of the two expressions agree. The derivative of the second expression is

$$\frac{1}{v-\phi_0} - \sum_{k=1}^{n-1} \frac{1}{k} k \left(\frac{v-a}{v-\phi_0}\right)^{k-1} \frac{(a-\phi_0)}{(v-\phi_0)^2}$$

$$= \frac{1}{v-\phi_0} - \frac{(a-\phi_0)}{(v-\phi_0)^2} \frac{1 - \left(\frac{v-a}{v-\phi_0}\right)^{n-1}}{1 - \frac{v-a}{v-\phi_0}} = \frac{1}{v-\phi_0}\left(\frac{v-a}{v-\phi_0}\right)^{n-1} = \frac{(v-a)^{n-1}}{(v-\phi_0)^n}$$

which is the derivative of the first expression.

Now we check the third expression. We have the Taylor Series

$$-\log(1-x) = \sum_{k=1}^{\infty} \frac{x^k}{k}$$

Substituting $x = \frac{v-a}{v-\phi_0}$ gives

$$\log\left(\frac{v-\phi_0}{a-\phi_0}\right) = -\log\left(1 - \frac{v-a}{v-\phi_0}\right) = \sum_{k=1}^{\infty} \frac{1}{k}\left(\frac{v-a}{v-\phi_0}\right)^k$$

We see that the first $n-1$ terms of k cancel out, and we get the third expression. $\square$

Lemma 2.

$$\int_{t=a}^{t=v}\left[\frac{(t-a)^{n-1}}{(t-\phi_0)^n}-(1-\lambda)\frac{(t-a)^n}{(t-\phi_0)^{n+1}}\right]dt=\lambda\log\left(\frac{v-\phi_0}{a-\phi_0}\right)-\lambda\sum_{k=1}^{n-1}\frac{(v-a)^k}{k(v-\phi_0)^k}+(1-\lambda)\frac{(v-a)^n}{n(v-\phi_0)^n}$$
$$=\frac{(v-a)^n}{n(v-\phi_0)^n}+\lambda\sum_{k=n+1}^{\infty}\frac{(v-a)^k}{k(v-\phi_0)^k}$$

Proof of Lemma 2. We apply Lemma 1:

$$\int_{t=a}^{t=v}\frac{(t-a)^{n-1}}{(t-\phi_0)^n}dt=\log\left(\frac{v-\phi_0}{a-\phi_0}\right)-\sum_{k=1}^{n-1}\frac{(v-a)^k}{k(v-\phi_0)^k}=\sum_{k=n}^{\infty}\frac{(v-a)^k}{k(v-\phi_0)^k}$$
$$\int_{t=a}^{t=v}\frac{(t-a)^{n}}{(t-\phi_0)^{n+1}}dt=\log\left(\frac{v-\phi_0}{a-\phi_0}\right)-\sum_{k=1}^{n}\frac{(v-a)^k}{k(v-\phi_0)^k}=\sum_{k=n+1}^{\infty}\frac{(v-a)^k}{k(v-\phi_0)^k}$$

Therefore,

$$\int_{t=a}^{t=v}\left[\frac{(t-a)^{n-1}}{(t-\phi_0)^n}-(1-\lambda)\frac{(t-a)^n}{(t-\phi_0)^{n+1}}\right]dt=\lambda\log\left(\frac{v-\phi_0}{a-\phi_0}\right)-\lambda\sum_{k=1}^{n-1}\frac{(v-a)^k}{k(v-\phi_0)^k}+(1-\lambda)\frac{(v-a)^n}{n(v-\phi_0)^n}$$
$$=\frac{(v-a)^n}{n(v-\phi_0)^n}+\lambda\sum_{k=n+1}^{\infty}\frac{(v-a)^k}{k(v-\phi_0)^k}$$

$\square$

B.1 Proofs and Discussions from Section 3.1

Proposition 2. *The optimal maximin ratio mechanism is never in the pure* SPA *regime.*

Proof of Proposition 2. Suppose for the sake of contradiction that the SPA regime is possible. By Theorem 2, the λ-regret

$$-1+(1-k_l)^n+\lambda\left(1-\int_{t=k_l}^{t=1}\left(1-\frac{k_l}{t}\right)^n dt\right)$$

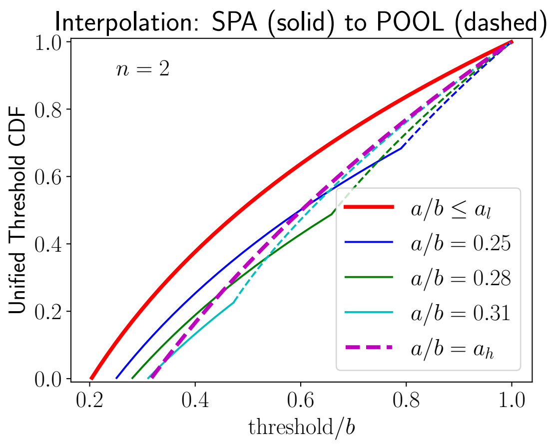


Figure 6: Unified threshold CDFs with SPA-reserve (solid lines) and POOL-thresholds (dashed lines) with $n = 2, \lambda = 1$. The figure shows the continuous interpolation of the mechanisms for $a/b \in \{k_l, 0.25, 0.28, 0.31, k_h\}$.

is zero, while the corresponding λ satisfies

$$\lambda = \frac{(1-k_l)^{n-1}}{\int_{t=k_l}^{t=1} \frac{(t-k)^{n-1}}{t^n} dt} .$$

Substituting the expression of λ gives

$$-1 + (1-k_l)^n + (1-k_l)^{n-1} \frac{1 - \int_{t=k_l}^{t=1} \left(1 - \frac{k_l}{t}\right)^n dt}{\int_{t=k_l}^{t=1} \frac{(t-k_l)^{n-1}}{t^n} dt} = 0 . \tag{8}$$

Let $f(k_l)$ be the left hand side of (8). We will derive a contradiction by showing that $f(k_l) > 0$ for all $0 < k_l \le 1$. We will prove this by viewing $k_l \in (0, 1]$ as a free variable. Let

$$I(k_l) = \int_{t=k_l}^{t=1} \frac{(t-k_l)^{n-1}}{t^n} dt = \sum_{i=n}^{\infty} \frac{1}{i} (1-k_l)^i .$$

Note that

$$I'(k_l) = \sum_{i=n}^{\infty} (1-k_l)^{i-1}(-1) = -\frac{(1-k_l)^{n-1}}{1-(1-k_l)} = -\frac{(1-k_l)^{n-1}}{k_l}$$

Also, by integration by part,

$$\int_{t=k_l}^{t=1} \left(1 - \frac{k_l}{t}\right)^n dt = \left[\left(1 - \frac{k_l}{t}\right)^n t\right]_{t=k_l}^{t=1} - \int_{t=k_l}^{t=1} tn \left(1 - \frac{k_l}{t}\right)^{n-1} \frac{k}{t^2} dt = (1 - k_l)^n - nkI(k_l)\,.$$

Therefore, both integrals in $f(k_l)$ can be written in terms of $I(k_l)$. We want to show that

$$-1 + (1 - k_l)^n + (1 - k_l)^{n-1} \frac{1 - (1 - k_l)^n + nkI(k_l)}{I(k_l)} > 0\,.$$

This is equivalent to

$$I(k_l) < \frac{1 - (1 - k_l)^n}{(1 - k_l)^{-(n-1)} - (1 + (n-1)k_l)}\,.$$

Note that $(1 - k_l)^{-(n-1)} > (1 + (n-1)k_l)$, so the above manipulation is valid, and both sides of the inequality are positive. Let

$$g(k_l) = I(k_l) - \frac{1 - (1 - k_l)^n}{(1 - k_l)^{-(n-1)} - (1 + (n-1)k_l)}\,.$$

It is clear from the integral definition of $I(k_l)$ that $\lim_{k_l \uparrow 1} I(k_l) = 1$. We now compute, by L'Hopital's rule,

$$\lim_{k_l \uparrow 1} \frac{1 - (1 - k_l)^n}{(1 - k_l)^{-(n-1)} - (1 + (n-1)k_l)} = \lim_{k_l \uparrow 1} \frac{-n(1 - k_l)^{n-1}(-1)}{(-n+1)(1 - k_l)^{-n}(-1) - (n-1)} = 0\,.$$

Therefore, $\lim_{k_l \uparrow 1} g(k_l) = 0$ To prove that $g(k_l) < 0$ it is sufficient to prove that $g(k_l)$ is strictly increasing in k_l, i.e., $g'(k_l) > 0$. This is very convenient because $I'(k_l)$ does not involve an integral.

We compute

$$
\begin{aligned}
g'(k_l) &= I'(k_l) - \frac{d}{dk_l}\left[\frac{1-(1-k_l)^n}{(1-k_l)^{-(n-1)} - (1+(n-1)k_l)}\right] \\
&= -\frac{(1-k_l)^{n-1}}{k_l} - \frac{1}{[(1-k_l)^{-n+1} - (1+(n-1)k_l)]^2} \times \Bigg\{ \\
&\left[(1-k_l)^{-n+1} - (1+(n-1)k_l)\right]\left[-n(1-k_l)^{n-1}(-1)\right] \\
&- [1-(1-k_l)^n]\left[(-n+1)(1-k_l)^{-n}(-1) - (n-1)\right]\Bigg\} \\
&= -\frac{(1-k_l)^{n-1}}{k_l} + \frac{\frac{(n-1)(1-(1-k_l)^n)^2}{(1-k_l)^n} - n\left[1-(1+(n-1)k_l)(1-k_l)^{n-1}\right]}{\frac{1}{(1-k_l)^{2n-2}}\left[1-(1+(n-1)k_l)(1-k_l)^{n-1}\right]^2}\,.
\end{aligned}
$$

Then $g'(k_l) > 0$ is equivalent to

$$
\begin{aligned}
(n-1)k_l(1-(1-k_l)^n)^2 - n(1-k_l)^n k_l\left[1-(1+(n-1)k_l)(1-k_l)^{n-1}\right] \\
-(1-k_l)\left[1-(1+(n-1)k_l)(1-k_l)^{n-1}\right]^2 > 0\,.
\end{aligned}
$$

Let $x = 1-k_l \in [0,1)$, algebraic simplification gives that the left hand side is

$$
\begin{aligned}
&(n-1)(1-x)(1-x^n)^2 - nx^n(1-x)\left[1-(1+(n-1)(1-x))x^{n-1}\right] \\
&\quad - \left[1-(1+(n-1)(1-x))x^{n-1}\right]^2 \\
&= (1-x^n)(n-1-nx+x^n)\,.
\end{aligned}
$$

It is clear that $1-x^n > 0$. We also have

$$
n-1-nx+x^n = n(1-x) - (1-x^n) = (1-x)\left(n - \sum_{i=0}^{n-1} x^i\right) > 0\,.
$$

We conclude that the optimal maximin ratio mechanism is never in the SPA regime. $\square$

C More Detailed Outline of the Proof of the Main Theorem from Section 3.3

C.1 Direct Mechanism Representation of Random-Thresholds SPA and POOL

Our main theorem shows that the optimal mechanism is a convex combination over the basis mechanisms $\{\mathrm{SPA}(r), \mathrm{POOL}(\tau)\}$. To prove the necessary saddle inequalities, we need to be able to compute the expected regret of any such mechanism against some distribution. It is more mathematically convenient to "flatten" the randomness over thresholds and give an almost equivalent "direct" representation that gives the allocation rule $x(\boldsymbol{v})$ and payment rule $p(\boldsymbol{v})$ for any valuation vector $\boldsymbol{v}$ as follows.

Definition 7 ((g_u, g_d) mechanisms)**.** *Let $g_u, g_d : [a,b] \to [0,1]$ be given functions. A mechanism (g_u, g_d) is defined by the allocation rule $x : [a,b]^n \to [0,1]^n$ given by, for each $i \in [n]$,*

$$x_i(\boldsymbol{v}) = \begin{cases} \frac{1}{k} g_u(v_{\max}) + \frac{k-1}{k} g_d(v_{\max}) & \textit{if } v_i = \max(\boldsymbol{v}) := v_{\max} \textit{ and there are } k \textit{ entries in } \boldsymbol{v} \textit{ equal to } v_{\max} \\ g_d(v_{\max}) & \textit{if } v_i < \max(\boldsymbol{v}) := v_{\max} \end{cases}$$

and the payment rule $p : [a,b]^n \to \mathbb{R}^n_{++}$ is determined uniquely from Myerson's formula such that the resulting mechanism (x, p) is dominant strategy incentive compatible.

In other words, the mechanism allocates $g_u(v_{\max})$ to the highest bidder(s) and $g_d(v_{\max})$ to the non-highest bidder(s). If there are k highest bidders, select one of them to be the "winner" with g_u uniformly at random.

We can prove the following proposition, wh and 8 that (g_u, g_d) mechanisms and convex combination of $\{\mathrm{SPA}(r), \mathrm{POOL}(\tau)\}$ are *almost* equivalent representations of the same mechanism class in the sense that one can be converted to another.

Proposition 3. *We have the following correspondence between the (g_u, g_d) mechanisms arising in our main theorem (Theorem 5) and convex combinations of SPAs and POOLs.*

(1) A mechanism is a (g_u, g_d) mechanism with $g_u(v) \in [0,1]$ increasing in v, $g_u(a) = 0$, and $g_d(v) = 0$ for all v if and only if it is $\mathrm{SPA}(\Phi)$, Φ has measure 1, and $\Phi(v) = g_u(v)$.

(2) *A mechanism is a (g_u, g_d) mechanism with $g_u(v) \in [0,1]$ increasing in v, $g_u(a) = g_d(a) = 1/n$, and $g_u(v) + (n-1)g_d(v) = 1$ for all v if and only if is* POOL(Ψ), *Ψ has measure 1, and $\Psi(v) = 1 - ng_d(v)$.*

(3) *A mechanism is a (g_u, g_d) mechanism with $g_u(v) \in [0,1]$ increasing in v for $v \in [a,b]$, $g_u(a) = g_d(a) := \alpha$, $g_d(v) = \alpha$ for $v \in [a, v^*]$ for some constant $v^* \in [a,b]$, $g_u(v) + (n-1)g_d(v) = 1$ for $v \in [v^*, b]$ if and only if it is a randomization over* SPA(Φ), *with Φ supported on $[a, v^*]$ and* POOL(Ψ) *with Ψ supported on $[v^*, b]$. Furthermore, their cumulative probabilities are given by $\Phi(v) = g_u(v) - \alpha$ for $v \in [a, v^*]$ and $\Psi(v) = n(\alpha - g_d(v))$ for $v \in [v^*, b]$.*

Proposition 3 follows immediately from the more general Proposition 8 about the near-equivalence between the classes of SPA-POOL mechanisms and (g_u, g_d) mechanisms. We state and prove Proposition 8 in Appendix D.1.

We choose to highlight $\{\text{SPA}(r), \text{POOL}(\tau)\}$ in the main text because the basis mechanisms SPA and POOL are intuitive and interpretable. However, the (g_u, g_d) representation is more mathematically convenient because it directly gives the allocation probabilities and payments for each valuation vector $\boldsymbol{v}$ which are ingredients for the expected regret calculation, and also because it allows the three cases to be treated in a unified manner.

We now derive the regret expressions for a given (g_u, g_d) mechanism. This will be a key ingredient in all subsequent saddle point calculations.

(Regret-$\boldsymbol{F}$)

Proposition 4 (Expected Regret of a (g_u, g_d) mechanism)**.** *Let $R_{\lambda,n}(\boldsymbol{g}, \boldsymbol{F}) := R_{\lambda,n}((g_u, g_d), \boldsymbol{F})$ be the expected regret of a (g_u, g_d) mechanism and valuation distribution $\boldsymbol{F}$.*

If we assume that g_u and g_d are continuous everywhere and differentiable everywhere except a finite number of points, and $\boldsymbol{F}$ is arbitrary, then the expected regret is

$$\begin{aligned} R_{\lambda,n}(\boldsymbol{g}, \boldsymbol{F}) &= a(\lambda - g_u(a) - (n-1)g_d(a)) + \int_{v=a}^{v=b} (\lambda - g_u(v) + g_d(v) - vg_u'(v) + (v - na)g_d'(v))dv \\ &+ \int_{v=a}^{v=b} (-\lambda + vg_u'(v) + (n-1)ag_d'(v))\boldsymbol{F}_n^{(1)}(v)dv + (g_u(v) - g_d(v) - (v-a)g_d'(v))\boldsymbol{F}_n^{(2)}(v)dv \end{aligned} \tag{Regret-$\boldsymbol{F}$}$$

where $\boldsymbol{F}_n^{(1)}$ and $\boldsymbol{F}_n^{(2)}$ are the first and second order statistics (highest and second-highest) among n

agents whose valuations are drawn from $\boldsymbol{F}$*.*

If we further assume that $\boldsymbol{F}$ *is i.i.d. with marginal* F*, then*

$$
\begin{aligned}
R_{\lambda,n}(\boldsymbol{g},\boldsymbol{F}) &= a(\lambda - g_u(a) - (n-1)g_d(a)) + \int_{v=a}^{v=b} (\lambda - g_u(v) + g_d(v) - vg_u'(v) + (v-na)g_d'(v)) \\
&\quad + \int_{v=a}^{v=b} \left(-\lambda - (n-1)(g_u(v) - g_d(v)) + v(g_u'(v) + (n-1)g_d'(v))\right) F(v)^n dv \\
&\quad + \int_{v=a}^{v=b} n(g_u(v) - g_d(v) - (v-a)g_d'(v))F(v)^{n-1} dv
\end{aligned}
\tag{Regret-F}
$$

If, instead, we let g_u *and* g_d *be arbitrary but we assume that* $\boldsymbol{F}$ *is i.i.d. with marginal* F *such that* F *has a density* $f = F'$ *on* $[a, b)$ *(so it potentially has point masses only at* a *and* b *of size* $F(a)$ *and* $F(\{b\}) := f_b$ *respectively), then we have the (Regret-*$\boldsymbol{g}$*) expression*

$$
\begin{aligned}
R_{\lambda,n}(\boldsymbol{g},\boldsymbol{F}) &= \lambda b - a\left(g_u(a) + (n-1)g_d(a)\right) F(a)^n - \left(bg_u(b) + (n-1)ag_d(b)\right)\left(1 - (1-f_b)^n\right) \\
&\quad + (b-a)g_d(b)\left(1 - (1-f_b)^{n-1}(1 + (n-1)f_b)\right) \\
&\quad + \int_{v=a}^{v=b} -\lambda F(v)^n + g_u(v)nF(v)^{n-1}(1 - F(v) - vF'(v))dv \\
&\quad + \int_{v=a}^{v=b} g_d(v)n(n-1)F(v)^{n-2}F'(v)\left\{(v-a)(1-F(v)) - aF(v)\right\} dv
\end{aligned}
\tag{Regret-$\boldsymbol{g}$}
$$

The above expression is valid for $n \geq 1$ *if we take the expression* $n(n-1)F(v)^{n-2}$ *to be zero for* $n = 1$*.*

We prove Proposition 4 in Appendix D.1.

C.2 Deriving Candidate Worst-Case Distribution from Seller's Saddle

Proposition 3 also shows that g_d can be written in terms of the function $g_u \equiv g$ (exact formulae depending on the 3 cases). We can substitute g_u and g_d with g in the (Regret-$\boldsymbol{g}$) expression to get $R_{\lambda,n}(g, F)$, which is explicitly linear in g; the resulting expression (Regret-g) is given in Lemma 4 in Appendix D.2. Seller's saddle then requires that the coefficient of $g(v)$ is zero; this gives an ODE for F^*, which has different forms depending on whether $v \in [a, v^*]$ or $v \in [v^*, b]$. The formal result is as follows.

Proposition 5. *Suppose that the problem* $\inf_g \sup_F R_{\lambda,n}(g, F)$ *has a solution* g^* *with a positive density over* (r^*, b) *for some* $r^* \leq v^*$*, and is such that the problem* $\sup_F R_{\lambda,n}(g^*, F)$ *has a unique solution* F^**. Then,* F^* *satisfies the ODE* $1 - F^*(v) - v(F^*)'(v) = 0$ *for* $v \in (r^*, v^*)$ *and* $F^*(v) - F^*(v)^2 - (v-a)(F^*)'(v) = 0$ *for* $v \in (v^*, b)$*.*

The solution to the first ODE is the isorevenue: $F(v) = 1 - r^*/v$, if we also require $F(r^*) = 0$. The second ODE is trickier: it is equivalent to $\frac{d}{dv}\left(v - \frac{v-a}{F^*(v)}\right) = 0$, so the expression in the parenthesis is a constant. If this constant is ϕ_0, then we get $F^*(v) = (v-a)/(v-\phi_0)$, a distribution with virtual value ϕ_0. For $a/b \leq k_l$, $v^* = b$, only the first ODE matters, and we let $r^* = k_l b$ so that the solution is feasible anywhere in the regime. For the other two regimes we can let $r^* = a$. In the case $k_l \leq a/b \leq k_h$, continuity of F^* at v^* requires that $\phi_0 = 0$, so $F^*(v) = 1 - a/v$ for $v \in [a, b)$. In the case, $a/b \geq k_h$, $v^* = a$, only the second ODE matters, and $F^*(v) = (v-a)/(v-\phi_0)$, where ϕ_0 will later be determined by the boundary condition required to make the mechanism feasible (cf. Section C.3).

C.3 Deriving Candidate Mechanism from Nature's Saddle

As discussed in the introduction to this section, because the (Regret-F) expression is a separable function of $F(v)$, if there is a solution that maximizes each term in the integrand pointwise, then it will solve Nature's saddle. It is sufficient that the first derivative of this term, evaluated at F^*, is zero, and the second derivative of this term, evaluated at F^*, is negative. This is stated formally in the following proposition.

Proposition 6. *Suppose that* m^* *is a* (g_u^*, g_d^*) *mechanism and* $F^*(v)$ *is an increasing function that satisfies the following conditions:*

$$\begin{aligned}\left(-\lambda - (n-1)(g_u^*(v) - g_d^*(v)) + v(g_u'(v) + (n-1)g_d'^*(v))\right) F^*(v) \\ +(n-1)(g_u^*(v) - g_d^*(v) - (v-a)g_d'^*(v)) = 0 \qquad \text{(FOC)}\\ g_u^*(v) - g_d^*(v) - (v-a)g_d'^*(v) > 0 \qquad \text{(SOC)}\end{aligned}$$

Then $R(m^*, F^*) \leq R(m^*, F)$ *for any* F*.*

Proposition 6 is useful both for deriving the candidate mechanism and to prove that Nature's saddle holds unconditionally. For the latter, it is sufficient to check (FOC) and (SOC) for a particular $((g_u^*, g_d^*), F^*)$. For the former, we parametrize both g_u^* and g_d^* in terms of a single function g^* just like Appendix C.2 and use (FOC) to derive an ODE that g^* must satisfy. The result is summarized in the following proposition.

Proposition 7. *Let m^* be a (g_u^*, g_d^*) mechanism such that $g_u^*(v) = g^*(v)$, $g_d^*(v) = \alpha$ for $v \in [a, v^*]$ and $g_d^*(v) = (1 - g^*(v))/(n-1)$ for $v \in [v^*, b]$, g^* has positive density on (r^*, b), and F^* is given by Proposition 5.*[8] *Then g^* satisfies the ODE*

$$(g^*)'(v) + \frac{(n-1)r^*}{v(v-r^*)}(g^*(v) - \alpha) = \frac{\lambda}{v} \qquad \textit{for } v \in (r^*, v^*) \qquad \text{(ODE-}g\text{-1)}$$

$$(g^*)'(v) + \frac{n(a-\phi_0)}{(v-\phi_0)(v-a)}g^*(v) = \frac{1}{v-a} - \frac{1-\lambda}{v-\phi_0} \qquad \textit{for } v \in (v^*, b) \qquad \text{(ODE-}g\text{-2)}$$

We are able to solve for solutions to the above ODEs in closed form; both are first order linear ODEs, so the general solution of each has one additional unknown constant. For $a/b \le k_l$, we have $\alpha = 0$ and only (ODE-g-1) is involved; we can pin down both the general constant and the unknown r^* are determined by the boundary conditions $g^*(r^*) = 0, g^*(b) = 1$, so we have a defining equation for r^* which does not depend on a, so this solution should work if and only if $r^* \ge a$. We therefore can characterize the threshold $k_l = r^*/b$. For $a/b \le k_l$, we have $\alpha = 1/n$ and only (ODE-g-2) is involved; we can pin down both the general constant and the unknown ϕ_0 are determined by the boundary conditions that $g^*(b) = 1$ and $\lim_{v \downarrow a} g^*(v)$ is nondegenerate.[9] The boundary k_h of the regime $a/b \ge k_h$ by the fact that the equation for ϕ_0 must have a solution. Also, it turns out that we can write ϕ_0 in terms of k_h as $\phi_0 = (a - k_h b)/(1 - k_h)$, which makes it clear that $\phi_0 = 0$ for $a/b = k_h$ (as it should be, for continuity with the intermediate regime) and $\phi_0 \uparrow 1$ for $a/b \uparrow 1$. Lastly, for $k_l \le a/b \le k_h$, we solve the two ODEs on (a, v^*) and (v^*, b) separately, and the constants α and $v^* := \tilde{v}^* b$ are determined by the boundary conditions on g^* at a, v^*, b. Lastly, we check that the resulting g^* is well-defined and increasing. When we transform g^* into (Φ^*, Ψ^*) using Proposition 3, we get the mechanism stated in Theorem 1.

C.4 Bringing everything together

We now have all the ingredients to prove Theorem 2. We first check that all mechanisms as stated are well-defined. For Nature's saddle, it is sufficient to verify that (FOC) and (SOC) holds for

[8] More precisely, F^* is given by the discussion right after Proposition 5, the last of Appendix C.2.

[9] It turns out we can derive $g^*(a) = \lim_{v \downarrow a} g(v) = 1/n$ from the equation itself.

the particular (g_u^*, g_d^*) in each case. (FOC) holds because we specifically construct the functions to satisfy the ODE. We can check (SOC) by writing all of g_u^*, g_d^*, and $(g_d^*)'$ in terms of g^* (we use the ODE for g^* for the last one), and the condition reduces to an easy inequality in every case. For seller's saddle, we fix the distribution F^* and want to show that our mechanism m^* minimizes expected regret. Because F^* pins down the benchmark, this is equivalent to showing that our mechanism maximizes expected revenue, given a known distribution F^*. This is a Bayesian mechanism design problem. We cannot directly use the standard result from Myerson [1981] because F^* has both a discrete and a continuous component, but we can apply the more general result from Monteiro and Svaiter [2010]. We give full details in Appendix D.4.

D Proofs and Discussions from Appendix C

D.1 Proofs and Discussions from Appendix C.1

Proposition 8 (Which (g_u, g_d) mechanisms correspond to a convex combination of different SPAs and POOLs?)**.** *A (g_u, g_d) mechanism can be written as a convex combination of* SPA(r)*, $r \sim \Phi$, and* POOL(τ)*, $\tau \sim \Psi$ (assuming that Φ and Ψ does not have a point mass at a), if and only if $g_u(a) = g_d(a) = \alpha$ for some α, $g_u(v) + (n-1)g_d(v)$ is increasing, and $g_d(v)$ is decreasing.*

Furthermore, there is a one-to-one correspondence between (g_u, g_d) and (Φ, Ψ). Given a valid (g_u, g_d), Φ has measure $|\Phi| := 1 - n\alpha$, Ψ has measure $|\Psi| := n\alpha$, and

$$\Psi(v) = n(\alpha - g_d(v))$$
$$\Phi(v) = g_u(v) + (n-1)g_d(v) - n\alpha$$

Here, we abuse notation and write $\Phi(v) = \int \mathbf{1}(r \leq v) d\Phi(r)$ and $\Psi(v) = \int \mathbf{1}(\tau \leq v) d\Psi(\tau)$ as cumulative probabilities. Conversely, given (Φ, Ψ), we have

$$g_u(v) = \Phi(v) + \frac{1}{n}|\Psi| + \frac{n-1}{n}\Psi(v)$$
$$g_d(v) = \frac{1}{n}(|\Psi| - \Psi(v))$$

Proof. Note that both SPA(r) and POOL(τ) are members of the (g_u, g_d) class of mechanisms. The functions of SPA(r) are given by Table 4.

	SPA(r)	RAND(τ)
$g_u(v)$	$\mathbf{1}(v \geq r)$	$\frac{1}{n} + \frac{n-1}{n}\mathbf{1}(v \geq \tau)$
$g_d(v)$	0	$\frac{1}{n} - \frac{1}{n}\mathbf{1}(v \geq \tau)$

Table 4: Both SPA(r) and POOL(τ) can be written as (g_u, g_d) mechanisms. Here are their associated g functions.

Note also that because the (g_u, g_d) mechanism allocates $g_u(v_{\max})$ to the highest, and $g_d(v_{\max})$ to the non-highest, breaking ties randomly, this tie-breaking is the same for the convex combination of mechanisms as well, and a convex combination of (g_u, g_d) mechanisms is still a (g_u, g_d) mechanisms, with the new functions being convex combinations of the old ones with the same weight. Therefore, A mechanism that is SPA(r), $r \sim \Phi$, and POOL(τ), $\tau \sim \Psi$, is a (g_u, g_d) mechanism with

$$\begin{aligned} g_u(v) &= \int \mathbf{1}(v \geq r) d\Phi(r) + \int \left(\frac{1}{n} + \frac{n-1}{n}\mathbf{1}(v \geq \tau)\right) d\Psi(r) &&= \Phi(v) + \frac{1}{n}|\Psi| + \frac{n-1}{n}\Psi(v) \\ g_d(v) &= \int 0 d\Phi(r) + \int \left(\frac{1}{n} - \frac{1}{n}\mathbf{1}(v \geq \tau)\right) d\Psi(v) &&= \frac{1}{n}|\Psi| - \frac{1}{n}\Psi(v) \end{aligned}$$

We therefore have a formula that transforms (Φ, Ψ) to (g_u, g_d). From these formula, we immediately see that $g_u(a) = g_d(a)$; we let this be α. We also see that $g_u(v) + (n-1)g_d(v) = \Phi(v) + |\Psi|$ is increasing in v, while $g_d(v) = \frac{1}{n}(|\Psi| - \Psi(v))$ is decreasing in v, because Φ and Ψ are increasing functions.

Conversely, assume that (g_u, g_d) has these properties. We will show that we can invert these formulas and find the corresponding (Φ, Ψ). From $g_u(v) = \Phi(v) + \frac{1}{n}|\Psi| + \frac{n-1}{n}\Psi(v)$, setting $v = a$ gives $\alpha = g_u(a) = \frac{1}{n}|\Psi|$, so $|\Psi| = n\alpha$, and $|\Phi| = 1 - |\Psi| = 1 - n\alpha$. From $g_d(v) = \frac{1}{n}(|\Psi| - \Psi(v))$, we get $\Psi(v) = |\Psi| - ng_d(v) = n(\alpha - g_d(v))$, and from $g_u(v) = \Phi(v) + \frac{1}{n}|\Psi| + \frac{n-1}{n}\Psi(v) = \Phi(v) + \frac{1}{n}(n\alpha) + \frac{n-1}{n} \cdot n(\alpha - g_d(v)) = \Phi(v) + n\alpha - (n-1)g_d(v)$, we get $\Phi(v) = g_u(v) + (n-1)g_d(v) - n\alpha$. $\square$

Proof of Proposition 4. We first derive the (Regret-$\boldsymbol{F}$) expression, assuming that g_u and g_d are continuous everywhere and differentiable everywhere except a finite number of points.

From Myerson's lemma,

$$p_i(\boldsymbol{v}) = v_i x_i(\boldsymbol{v}) - \int_{\tilde{v}_i = a}^{\tilde{v}_i = v_i} x_i(\tilde{v}_i, \boldsymbol{v}_{-i}) d\tilde{v}_i$$

the allocation rule (g_u, g_d) gives

$$p_i(\boldsymbol{v}) = \begin{cases} v_i g_u(v_i) - (v^{(2)} - a)g_d(v^{(2)}) - \int_{t=v^{(2)}}^{t=v_i} g_u(t)dt & \text{if } v_i \text{ is the highest and } v^{(2)} \text{ is the second-highest} \\ a g_d(v^{(1)}) & \text{if } v^{(1)} \text{ is the highest and is NOT } v_i \end{cases}$$

so the pointwise regret is

$$v^{(1)}(\lambda - g_u(v^{(1)})) - (n-1)a g_d(v^{(1)}) + (v^{(2)} - a)g_d(v^{(2)}) + \int_{t=v^{(2)}}^{t=v^{(1)}} g_u(t)dt$$

We will now prove the following technical lemma.

Lemma 3. *If h is a differentiable function, then*

$$\int_{w\in[a,b]} h(w)dG(w) = h(a) + \int_{w\in[a,b]} h'(w)(1-G(w))dw$$

Proof of Lemma 3.

$$\begin{aligned}\int_{w\in[a,b]} h(w)dG(w) &= \int_{w\in[a,b]} \left(h(a) + \int_{\tilde{w}=a}^{\tilde{w}=w} h'(\tilde{w})d\tilde{w}\right) dG(w) = h(a) + \int_{\tilde{w}=a}^{\tilde{w}=b} h'(\tilde{w}) \int_{w\in(\tilde{w},b]} dG(w)dw \\ &= h(a) + \int_{\tilde{w}=a}^{\tilde{w}=b} h'(\tilde{w})(G(b) - G(\tilde{w}))d\tilde{w} = h(a) + \int_{\tilde{w}=a}^{\tilde{w}=b} h'(\tilde{w})(1-G(\tilde{w}))d\tilde{w}\end{aligned}$$

□

Therefore, by Lemma 3,

$$\begin{aligned}&\mathbb{E}[v^{(1)}(\lambda - g_u(v^{(1)})) - (n-1)a g_d(v^{(1)})] \\ &= a(\lambda - g_u(a) - (n-1)g_d(a)) + \int_{v\in[a,b]} (\lambda - g_u(v) - v g_u'(v) - (n-1)a g_d'(v))(1 - \boldsymbol{F}_n^{(1)}(v))dv\end{aligned}$$

Now we compute the second term.

$$\begin{aligned}\mathbb{E}[(v^{(2)} - a)g_d(v^{(2)})] &= (a-a)g_d(a) + \int_{v\in[a,b]} (g_d(v) + (v-a)g_d'(v))(1 - \boldsymbol{F}_n^{(2)}(v))dv \\ &= \int_{v\in[a,b]} (g_d(v) + (v-a)g_d'(v))(1 - \boldsymbol{F}_n^{(2)}(v))dv\end{aligned}$$

Lastly, we compute the third term

$$\begin{aligned}
\mathbb{E}\left[\int_{t=v^{(2)}}^{t=v^{(1)}} g_u(t)dt\right] &= \mathbb{E}\left[\int_{v\in[a,b]} g_u(v)\mathbf{1}(v^{(2)} < v \le v^{(1)})dv\right] \\
&= \int_{v\in[a,b]} g_u(v)\Pr(v^{(2)} < v \le v^{(1)})dv \\
&= \int_{v\in[a,b]} g_u(v)(\boldsymbol{F}_n^{(2)}(v) - \boldsymbol{F}_n^{(1)}(v))dv
\end{aligned}$$

Therefore, the regret is

$$\begin{aligned}
&a(\lambda - g_u(a) - (n-1)g_d(a)) + \int_{v\in[a,b]} (\lambda - g_u(v) - vg_u'(v) - (n-1)ag_d'(v))(1 - \boldsymbol{F}_n^{(1)}(v))dv \\
&\quad + \int_{v\in[a,b]} (g_d(v) + (v-a)g_d'(v))(1 - \boldsymbol{F}_n^{(2)}(v))dv + \int_{v\in[a,b]} g_u(v)(\boldsymbol{F}_n^{(2)}(v) - \boldsymbol{F}_n^{(1)}(v))dv
\end{aligned}$$

Rearranging this gives the (Regret-$\boldsymbol{F}$) expression.

$$\begin{aligned}
R(g, \boldsymbol{F}) = a(\lambda - g_u(a) - (n-1)g_d(a)) &+ \int_{v\in[a,b]} (\lambda - g_u(v) + g_d(v) - vg_u'(v) + (v - na)g_d'(v)) \\
&+ \int_{v\in[a,b]} \left(-\lambda - (n-1)(g_u(v) - g_d(v)) + v(g_u'(v) + (n-1)g_d'(v))\right) \boldsymbol{F}_n^{(1)}(v)dv \\
&+ \int_{v\in[a,b]} n(g_u(v) - g_d(v) - (v-a)g_d'(v))\boldsymbol{F}_{n-1}^{(1)}(v)dv
\end{aligned}$$

Now we derive the (Regret-g) expression, assuming that g_u and g_d are arbitrary but $\boldsymbol{F}$ is i.i.d. with marginal F with a density F' in the interior. We start with the expected pointwise regret expression

$$\mathbb{E}\left[v^{(1)}(\lambda - g_u(v^{(1)})) - (n-1)ag_d(v^{(1)}) + (v^{(2)} - a)g_d(v^{(2)}) + \int_{t=v^{(2)}}^{t=v^{(1)}} g_u(t)dt\right]$$

The third term is still the same

$$\begin{aligned}
\mathbb{E}\left[\int_{t=v^{(2)}}^{t=v^{(1)}} g_u(t)dt\right] &= \int_{v\in[a,b]} g_u(v)(\boldsymbol{F}_n^{(2)}(v) - \boldsymbol{F}_n^{(1)}(v))dv \\
&= \int_{v\in[a,b]} g_u(v)(nF(v)^{n-1} - nF(v)^n)dv
\end{aligned}$$

We will write the point mass of F at b as $f_b := F(\{b\}) = 1 - F(b^-)$ for convenience. Now for the first term, we know that the probability that $v^{(1)} = b$ is

$$\Pr(v^{(1)} = b) = 1 - \Pr(v^{(1)} < b) = 1 - \prod_{i=1}^{n} \Pr(v_i < b) = 1 - F(b^-)^n = 1 - (1 - f_b)^n$$

The probability that $v^{(1)} = a$ is

$$\Pr(v^{(1)} = a) = \prod_{i=1}^{n} \Pr(v_i = a) = F(a)^n$$

For $v \in (a, b)$, $F^{(1)}$ has a density given by

$$f^{(1)}(v) = nF(v)^{n-1}F'(v)$$

Now,

$$\begin{aligned} \Pr(v^{(2)} = b) = \Pr(\text{ at least 2 of the } n\ v\text{'s are } b\ ) \\ = 1 - \Pr(\text{ exactly 0 of the } n\ v\text{'s are } b\ ) - \Pr(\text{ exactly 1 of the } n\ v\text{'s are } b\ ) \\ = 1 - (1 - f_b)^n - nf_b(1 - f_b)^{n-1} \\ = 1 - (1 - f_b)^{n-1}(1 + (n-1)f_b) \end{aligned}$$

and

$$\begin{aligned} \Pr(v^{(2)} = a) = \Pr(\text{ all are } a\ ) + \Pr(\ n-1 \text{ are } a, \text{ 1 are } > a) \\ = F(a)^n + nF(a)^{n-1}(1 - F(a)) = F(a)^{n-1}(1 + (n-1)F(a)) \end{aligned}$$

For $v \in (a, b)$, $F^{(2)}$ has a density given by

$$f^{(2)}(v) = n(n-1)F(v)^{n-2}(1 - F(v))F'(v)$$

and CDF given by

$$F^{(2)}(v) = nF(v)^{n-1} - (n-1)F(v)^n$$

We then have

$$
\begin{aligned}
&\mathbb{E}\left[v^{(1)}(\lambda - g_u(v^{(1)})) - (n-1)ag_d(v^{(1)})\right] \\
&= (a(\lambda - g_u(a)) - (n-1)ag_d(a)) \Pr(v^{(1)} = a) + (b(\lambda - g_u(b)) - (n-1)ag_d(b)) \Pr(v^{(1)} = b) \\
&+ \int_{v\in[a,b]} (v(\lambda - g_u(v)) - (n-1)ag_d(v)) f^{(1)}(v)dv \\
&= (a(\lambda - g_u(a)) - (n-1)ag_d(a))F(a)^n + (b(\lambda - g_u(b)) - (n-1)ag_d(b))(1-(1-f_b)^n) \\
&+ \int_{v\in[a,b]} (v(\lambda - g_u(v)) - (n-1)ag_d(v)) nF(v)^{n-1}F'(v)dv
\end{aligned}
$$

and

$$
\begin{aligned}
&\mathbb{E}\left[(v^{(2)} - a)g_d(v^{(2)})\right] \\
&= (a-a)g_d(a)\Pr(v^{(2)} = a) + (b-a)g_d(b)\Pr(v^{(2)} = b) + \int_{v=a}^{v=b} (v-a)g_d(v)f^{(2)}(v)dv \\
&= (b-a)g_d(b)(1-(1-f_b)^{n-1}(1+(n-1)f_b))) + \int_{v=a}^{v=b} (v-a)g_d(v)n(n-1)F(v)^{n-2}(1-F(v))F'(v)dv
\end{aligned}
$$

Therefore,

$$
\begin{aligned}
R(g,F) &= (a(\lambda - g_u(a)) - (n-1)ag_d(a))F(a)^n + (b(\lambda - g_u(b)) - (n-1)ag_d(b))(1-(1-f_b)^n) \\
&\quad + (b-a)g_d(b)(1-(1-f_b)^{n-1}(1+(n-1)f_b))) \\
&\quad + \int_{v=a}^{v=b} \left[(n-1)(v-a)g_d(v)(1-F(v)) + (v(\lambda - g_u(v)) - (n-1)ag_d(v))F(v)\right] nF(v)^{n-2}F'(v)dv \\
&\quad + \int_{v=a}^{v=b} g_u(v)(nF(v)^{n-1} - nF(v)^n)dv
\end{aligned}
$$

or

$$
\begin{aligned}
R(g,F) &= (a(\lambda - g_u(a)) - (n-1)ag_d(a))F(a)^n + (b(\lambda - g_u(b)) - (n-1)ag_d(b))(1-(1-f_b)^n) \\
&\quad + (b-a)g_d(b)(1-(1-f_b)^{n-1}(1+(n-1)f_b))) \\
&\quad + \int_{v=a}^{v=b} \lambda v nF(v)^{n-1}F'(v) - g_u(v)nvF(v)^{n-1}F'(v) \\
&\quad + \int_{v=a}^{v=b} \left\{(v-a)(1-F(v)) - aF(v)\right\} g_d(v)n(n-1)F(v)^{n-2}F'(v) + g_u(v)(nF(v)^{n-1} - nF(v)^n)dv
\end{aligned}
$$

By integration by part, the first term of the third line is λ times

$$\int_{v=a}^{v=b} vnF(v)^{n-1}F'(v)dv = bF(b^-)^n - aF(a)^n - \int_{v=a}^{v=b} F(v)^n dv$$

where $F(b^-) = (1 - f_b)$. Substituting this in gives

$$\begin{aligned}
R(g,F) = \lambda b &- a\left(g_u(a) + (n-1)g_d(a)\right)F(a)^n - \left(bg_u(b) + (n-1)ag_d(b)\right)\left(1-(1-f_b)^n\right)\\
&+ (b-a)g_d(b)\left(1-(1-f_b)^{n-1}(1+(n-1)f_b)\right)\\
&+ \int_{v=a}^{v=b} -\lambda F(v)^n + g_u(v)nF(v)^{n-1}(1-F(v)-vF'(v))\\
&+ \int_{v=a}^{v=b} g_d(v)n(n-1)F(v)^{n-2}F'(v)\left\{(v-a)(1-F(v)) - aF(v)\right\}dv
\end{aligned}$$

□

D.2 Proofs and Discussions from Section C.2

Lemma 4. *If we set $g_u(v) = g(v)$, $g_d(v) = \alpha$, for $v \in [a, v^*]$ and $g_u(v) = g(v), g_d(v) = (1 - g(v))/(n-1)$ for $v \in [v^*, b]$, $g(a) = \alpha, g(b) = 1$ in (Regret-$\boldsymbol{g}$), then we get the following expression*

$$\begin{aligned}
&\lambda b - an\alpha F(a)^n - b + b(1-f_b)^n\\
&+ \int_{v=a}^{v=v^*} -\lambda F(v)^n + g(v)nF(v)^{n-1}(1-F(v)-vF'(v))\\
&+ \int_{v=v^*}^{v=b} \left\{-\lambda F(v)^n + nF(v)^{n-2}F'(v)(v-a) - nF(v)^{n-1}vF'(v)\right\} + nF(v)^{n-2}\left\{F(v) - F(v)^2 - (v-a)F'(v)\right\}g(v)d
\end{aligned} \tag{Regret-g}$$

Proof of Lemma 4. We start from

$$\begin{aligned}
R_{\lambda,n}(\boldsymbol{g},\boldsymbol{F}) = \lambda b &- a\left(g_u(a) + (n-1)g_d(a)\right)F(a)^n - \left(bg_u(b) + (n-1)ag_d(b)\right)\left(1-(1-f_b)^n\right)\\
&+ (b-a)g_d(b)\left(1-(1-f_b)^{n-1}(1+(n-1)f_b)\right)\\
&+ \int_{v=a}^{v=b} -\lambda F(v)^n + g_u(v)nF(v)^{n-1}(1-F(v)-vF'(v))dv\\
&+ \int_{v=a}^{v=b} g_d(v)n(n-1)F(v)^{n-2}F'(v)\left\{(v-a)(1-F(v)) - aF(v)\right\}dv
\end{aligned} \tag{Regret-$\boldsymbol{g}$}$$

The first line is

$$\lambda b - a(n\alpha)F(a)^n - (b \cdot 1 + (n-1)a \cdot 0)(1-(1-f_b)^n)$$
$$= \lambda b - an\alpha F(a)^n - b + b(1-f_b)^n$$

The second line is zero because $g_d(b) = 0$. The third and fourth line (with integrals) are

$$\int_{v=a}^{v=v^*} -\lambda F(v)^n + g(v)nF(v)^{n-1}(1-F(v)-vF'(v))$$
$$+ \int_{v=v^*}^{v=b} -\lambda F(v)^n + g(v)nF(v)^{n-1}(1-F(v)-vF'(v)) + (1-g(v))nF(v)^{n-2}F'(v)\{(v-a)(1-F(v)) - aF(v)\}\, dv$$

The integral from a to v^* is already simplified. The integral from v^* to b is

$$\int_{v=v^*}^{v=b} -\lambda F(v)^n + g(v)nF(v)^{n-1}(1-F(v)-vF'(v)) + \left\{nF(v)^{n-2}F'(v)(v-a) - nF(v)^{n-1}vF'(v)\right\}(1-g(v))dv$$
$$= \int_{v=v^*}^{v=b} -\lambda F(v)^n + nF(v)^{n-2}F'(v)(v-a) - nF(v)^{n-1}vF'(v)$$
$$+ \left\{nF(v)^{n-1}(1-F(v)-vF'(v)) - nF(v)^{n-2}F'(v)(v-a) + nF(v)^{n-1}vF'(v)\right\} g(v)dv$$
$$= \int_{v=v^*}^{v=b} \left\{-\lambda F(v)^n + nF(v)^{n-2}F'(v)(v-a) - nF(v)^{n-1}vF'(v)\right\} + nF(v)^{n-2}\left\{F(v) - F(v)^2 - (v-a)F'(v)\right\} g(v)d$$

$\square$

Proof of Proposition 5. The essential core of the proof, arguing that the coefficients of every $g(v)$ in the interior must be zero, is essentially the same as the proof of Proposition 2 of Anunrojwong et al. [2022]. We essentially repeat that argument here for completeness.

Define $h(g) = \sup_F R(g, F)$ to be the worst-case regret for the function g using the (Regret-g) expression and $Z_0(g) = \{\tilde{F} : R(g, \tilde{F}) = \sup_F R(g, F)\}$ the set of optimal worst-case distributions for g. From the (Regret-g) expression given in Lemma 4, $R(g, F)$ is linear. Because the pointwise supremum of linear functions is convex, we have that $h(g)$ is convex in g. Since we assumed that $Z_0(g^*) = \{F^*\}$, Danskin's theorem then implies that h is differentiable at g^*, and the derivative of h at g^* is given by $\frac{\partial h}{\partial g}(g^*) = \frac{\partial R}{\partial g}(g^*, F^*)$. Therefore, for any Lipschitz-continuous function δ :

$[0,1] \to \mathbb{R}$, the directional derivative is given by

$$\partial_\Phi h(g^*)[\delta] = \int_{v=r^*}^{v=v^*} nF(v)^{n-1}(1-F(v)-vF'(v))\delta(v)dv + \int_{v=v^*}^{v=b} nF(v)^{n-2}\left\{F(v)-F(v)^2-(v-a)F'(v)\right\}\delta(v)dv.$$

A function δ is a *feasible variation* at g^* if there exists some $\epsilon_0 > 0$ such that $g^*(v) + \epsilon\delta(v)$ is increasing and in $[0,1]$ for all $v \in [a,b]$ and $\epsilon < \epsilon_0$. Because g^* has a positive density over (r^*, b) we have that every δ satisfying $\delta(v) = 0$ for $v \notin (r^*, b)$ is a feasible variation. Because h has a minimum at g^*, the first-order conditions imply $\partial_g h(g^*)[\delta] = 0$ for every feasible variation δ. Therefore, the fundamental lemma of calculus of variation implies that $1 - F^*(v) - v(F^*)'(v) = 0$ for $v \in (r^*, v^*)$ and $F^*(v) - F^*(v)^2 - (v-a)(F^*)'(v) = 0$ for $v \in (v^*, b)$ as desired.

□

D.3 Proofs and Discussions from Section C.3

Proof of Proposition 6. We use the following expression for λ-regret

$$\begin{aligned} R(g,F) = a(\lambda - g_u(a) - (n-1)g_d(a)) + \int_{v\in[a,b]} (\lambda - g_u(v) + g_d(v) - vg_u'(v) + (v-na)g_d'(v)) \\ + \int_{v\in[a,b]} \left(-\lambda - (n-1)(g_u(v) - g_d(v)) + v(g_u'(v) + (n-1)g_d'(v))\right) F(v)^n dv \\ + \int_{v\in[a,b]} n(g_u(v) - g_d(v) - (v-a)g_d'(v))F(v)^{n-1}dv \end{aligned}$$

In Nature's saddle, we fix the mechanism (g_u, g_d) and optimize over F. The integral expression is separable over $F(v)$ for $v \in (a,b)$. Here we will assume that the optimization is done pointwise.

The first order condition on F on the regret pointwise is

$$\begin{aligned} \left(-\lambda - (n-1)(g_u^*(v) - g_d^*(v)) + v(g_u'^*(v) + (n-1)g_d'^*(v))\right) \cdot nF(v)^{n-1} \\ + n(g_u(v) - g_d(v) - (v-a)g_d'(v)) \cdot (n-1)F(v)^{n-2} = 0 \end{aligned}$$

Nature's saddle states that over all F, F^* maximizes the λ-regret. If pointwise optimization is valid, then F^* must satisfy the above FOC equation. Dividing both sides by $nF^*(v)^{n-2}$ gives (FOC) as required.

For F^* to be maximizing, we also need the second-order conditions to hold, namely, that the second derivative with respect to $F(v)$ evaluated at $F^*(v)$ is negative[10]:

$$\begin{aligned}\left(-\lambda-(n-1)(g_u^*(v)-g_d^*(v))+v(g_u'^*(v)+(n-1)g_d'^*(v))\right)\cdot n(n-1)F^*(v)^{n-2}\\ +n(g_u(v)-g_d(v)-(v-a)g_d'(v))\cdot(n-1)(n-2)F^*(v)^{n-3}<0\end{aligned}$$

or

$$\left(-\lambda-(n-1)(g_u^*(v)-g_d^*(v))+v(g_u'^*(v)+(n-1)g_d'^*(v))\right)F^*(v)+(n-2)(g_u(v)-g_d(v)-(v-a)g_d'(v))<0$$

but from the (FOC) equality that we have just derived,

$$\begin{aligned}&\left(-\lambda-(n-1)(g_u^*(v)-g_d^*(v))+v(g_u'^*(v)+(n-1)g_d'^*(v))\right)F^*(v)+(n-2)(g_u(v)-g_d(v)-(v-a)g_d'(v))\\ =&\left(-\lambda-(n-1)(g_u^*(v)-g_d^*(v))+v(g_u'^*(v)+(n-1)g_d'^*(v))\right)F^*(v)+(n-1)(g_u^*(v)-g_d^*(v)-(v-a)g_d'^*(v))\\ &-(g_u^*(v)-g_d^*(v)-(v-a)g_d'^*(v))\\ &=-(g_u^*(v)-g_d^*(v)-(v-a)g_d'^*(v))\end{aligned}$$

Therefore, our condition reduces to (SOC), as required.

□

Proof of Proposition 7. We take (FOC) and substitute F^* from Proposition 5.

For $v\in(r^*,v^*)$, we let $F^*(v)=1-r^*/v$, $g_u^*(v)=g^*(v)$, $g_d^*(v)=\alpha$ to get

$$\left(-\lambda-(n-1)g^*(v)+v(g^*)'(v)\right)\left(1-\frac{r^*}{v}\right)+(n-1)g^*(v)=0$$

which is equivalent to (ODE-g-1)

For $v\in(v^*,b)$, we let $F^*(v)=1-(a-\phi_0)/(v-\phi_0)=(v-a)/(v-\phi_0)$, $g_u^*(v)=g(v)$,

[10] Note that if $n=2$ the last term disappear, so we can write $F^*(v)^{n-3}$ there with the understanding that the entire term becomes zero for $n=2$.

$g_d^*(v) = (1 - g(v))/(n-1)$ to get

$$\left(-\lambda - (n-1)\left(g^*(v) - \frac{1-g^*(v)}{n-1}\right)\right)\left(\frac{v-a}{v-\phi_0}\right) + (n-1)\left(g^*(v) - \frac{1-g^*(v)}{n-1} + (v-a)\frac{(g^*)'(v)}{n-1}\right)$$
$$(1-\lambda-ng^*(v))\left(\frac{v-a}{v-\phi_0}\right) + ng^*(v) - 1 + (v-a)(g^*)'(v) = 0$$
$$(1-\lambda)\frac{(v-a)}{(v-\phi_0)} + n\frac{(a-\phi_0)}{(v-\phi_0)}g^*(v) - 1 + (v-a)(g^*)'(v) = 0$$
$$(g^*)'(v) + \frac{n(a-\phi_0)}{(v-\phi_0)(v-a)}g^*(v) = \frac{1}{v-a} - \frac{(1-\lambda)}{(v-\phi_0)}$$

The last ODE is (ODE-g-2).

□

D.4 Proofs and Discussions from Section C.4

Proof of Theorem 2. We will convert the main theorem (Theorem 2) into the equivalent (g_u, g_d) representation Theorem 5 which we will prove.

□

Theorem 5 (Main Theorem in (g_u, g_d))**.** *Fix $\lambda \in (0,1]$ and a positive integer n. Define $k_l \in (0,1)$ as a unique solution to*

$$\lambda\left[\frac{\log(1/k_l)}{(1-k_l)^{n-1}} - \sum_{k=1}^{n-1}\frac{1}{k(1-k_l)^{n-1-k}}\right] = \lambda\sum_{k=n}^{\infty}\frac{1}{k}(1-k_l)^{k-(n-1)} = 1$$

and $k_h \in (0,1)$ as a unique solution to

$$\frac{1-\lambda}{n} + \lambda\frac{\log(1/k_h) - \sum_{k=1}^{n-1}\frac{(1-k_h)^k}{k}}{(1-k_h)^n} = \frac{1}{n} + \lambda\sum_{k=n+1}^{\infty}\frac{(1-k_h)^{k-n}}{k} = 1$$

(We define $k_h = 1$ if $n = 1$.)

Then we have $R_{\lambda,n}(m, \boldsymbol{F}^*) \leq R_{\lambda,n}(m^*, \boldsymbol{F}^*) \leq R_{\lambda,n}(m^*, \boldsymbol{F})$ for any $m \in \mathcal{M}_{\text{all}}$ and $\boldsymbol{F} \in \mathcal{F}_{\text{iid}}$, where m^* and $\boldsymbol{F}^*$ (which is n i.i.d. with marginal F^*) is defined depending on the value of a/b as follows.

- *For $a/b \leq k_l$, we define m^* as SPA(Φ^*), or a (g_u^*, g_d^*) mechanism with $g_u^*(v) = \Phi^*(v), g_d^*(v) = 0$, where*

$$\Phi^*(v) = g_u^*(v) = \lambda \frac{v^{n-1}}{(v-r^*)^{n-1}} \int_{t=r^*}^{t=v} \frac{(t-r^*)^{n-1}}{t^n} dt$$

and

$$F^*(v) = \begin{cases} 0 & \text{if } v \in [a, r^*] \\ 1 - \frac{r^*}{v} & \text{if } v \in [r^*, b) \\ 1 & \text{if } v = b \end{cases}$$

with $r^* = k_l b$.

- *For* $a/b \geq k_h$, *we define* m^* *as a* (g_u^*, g_d^*) *mechanism with*

$$g_u^*(v) = \frac{1-\lambda}{n} + \lambda \left(\frac{v-\phi_0}{v-a}\right)^n \log\left(\frac{v-\phi_0}{a-\phi_0}\right) - \lambda \sum_{k=1}^{n-1} \frac{1}{k} \left(\frac{v-\phi_0}{v-a}\right)^{n-k}$$
$$g_d^*(v) = \frac{1-g_u^*(v)}{n-1}$$

and

$$F^*(v) = \begin{cases} 1 - \frac{a-\phi_0}{v-\phi_0} & \text{if } v \in [a, b) \\ 1 & \text{if } v = b \end{cases}$$

where

$$\phi_0 = \frac{a-k_h b}{1-k_h} \in [0, a]$$

- *For* $k_l \leq a/b \leq k_h$, *we define* m^* *as a* (g_u^*, g_d^*) *mechanism with*

$$g_u^*(v) = \begin{cases} \alpha + \lambda \left(\frac{v}{v-a}\right)^{n-1} \int_{t=a}^{t=v} \frac{(t-a)^{n-1}}{t^n} dt & \text{for } v \in [a, v^*] \\ \frac{v^n}{(v-a)^n} \left[\frac{(b-a)^n}{b^n} - \int_{t=v}^{t=b} \left[\frac{(t-a)^{n-1}}{t^n} - (1-\lambda)\frac{(t-a)^n}{t^{n+1}}\right] dt\right] & \text{for } v \in [v^*, b] \end{cases}$$

and

$$g_d^*(v) = \begin{cases} \alpha & \text{if } v \in [a, r^*] \\ \frac{1-g_u^*(v)}{n-1} & \text{if } v \in [r^*, b] \end{cases}$$

and

$$F^*(v) = \begin{cases} 1 - \frac{a}{v} & \text{if } v \in [a, b) \\ 1 & \text{if } v = b \end{cases}$$

where (v^*, α) *is the unique solution to*

$$\frac{(r^* - a)^{n-1}}{(r^*)^{n-1}}(1 - n\alpha) = \lambda \int_{t=a}^{t=r^*} \frac{(t-a)^{n-1}}{t^n} dt$$
$$\frac{(b-a)^n}{b^n} - \frac{(r^*-a)^n}{(r^*)^n}(1-(n-1)\alpha) = \int_{t=r^*}^{t=b} \left[\frac{(t-a)^{n-1}}{t^n} - (1-\lambda)\frac{(t-a)^n}{t^{n+1}} \right] dt$$

In particular, this implies that against $\mathcal{M}_{\text{all}}$ *and* $\mathcal{F}_{\text{iid}}$, m^* *is the optimal mechanism,* $\boldsymbol{F}^*$ *is the corresponding worst-case distribution, and the optimal performance is* $\inf_{m \in \mathcal{M}_{\text{all}}} \sup_{\boldsymbol{F} \in \mathcal{F}_{\text{iid}}} R_{\lambda,n}(m, \boldsymbol{F}) = R_{\lambda,n}(m^*, \boldsymbol{F}^*)$ *with the* m^* *and* $\boldsymbol{F}^*$ *just defined.*

Proof of Theorem 5. The Seller's Saddle is a Bayesian mechanism design problem, and we can check optimality through the conditions of Monteiro and Svaiter [2010]. (In particular, we see that the mechanism and the distribution have the same support, and the distribution has constant nonnegative virtual value on the support.) Therefore, for the rest of this proof we will focus on Nature's saddle.

By Proposition 6, it is sufficient to exhibit (g_u^*, g_d^*) that satisfies (FOC) and (SOC). We will solve the ODEs in Proposition 7) and show that there is a feasible solution, and that solution satisfies (FOC) by definition, and we will then verify (SOC), finishing the proof.

Low Information ($a/b \le k_l$) Regime. Here, only (ODE-g-1) is relevant and we set $\alpha = 0$. Then we can write

$$\frac{d}{dv}\left[\frac{(v-r^*)^{n-1}}{v^{n-1}}g(v)\right]=\lambda\frac{(v-r^*)^{n-1}}{v^n}$$

Because $\frac{(v-r^*)^{n-1}}{v^{n-1}}g(v)$ is 0 for $v=r^*$, integrating from r^* to v gives

$$\frac{(v-r^*)^{n-1}}{v^{n-1}}g(v)=\lambda\int_{t=r^*}^{t=v}\frac{(t-r^*)^{n-1}}{t^n}dt$$

or

$$g(v)=\lambda\frac{v^{n-1}}{(v-r^*)^{n-1}}\int_{t=r^*}^{t=v}\frac{(t-r^*)^{n-1}}{t^n}dt$$

We can then use Lemma 1 to evaluate this integral. We determine r^* from $g(b)=1$, so

$$\lambda\left[\left(\frac{b}{b-r^*}\right)^{n-1}\log\left(\frac{b}{r^*}\right)-\sum_{k=1}^{n-1}\frac{1}{k}\left(\frac{b}{b-r^*}\right)^{n-1-k}\right]=\lambda\sum_{k=n}^{\infty}\frac{1}{k}\left(\frac{b-r^*}{b}\right)^{k-(n-1)}=1$$

Note that r^*/b satisfies the same equation as k_l, so we can set $r^*=k_lb$. This equation has a unique solution in r^* if and only if as $r^*\downarrow a$, the left hand side is ≥ 1, that is,

$$\lambda\sum_{k=n}^{\infty}\frac{1}{k}\left(1-\frac{a}{b}\right)^{k-(n-1)}\geq 1$$

or $a/b\leq k_l$ where $k_l\in(0,1)$ is a unique solution to

$$\lambda\left[\frac{\log(1/k_l)}{(1-k_l)^{n-1}}-\sum_{k=1}^{n-1}\frac{1}{k(1-k_l)^{n-1-k}}\right]=\lambda\sum_{k=n}^{\infty}\frac{1}{k}(1-k_l)^{k-(n-1)}=1\,.$$

High Information ($a/b\geq k_l$) Regime. Here, only (ODE-g-2) is relevant and we can write the ODE as

$$\frac{d}{dv}\left[\left(\frac{v-a}{v-\phi_0}\right)^n g(v)\right]=\frac{(v-a)^{n-1}}{(v-\phi_0)^n}-(1-\lambda)\frac{(v-a)^n}{(v-\phi_0)^{n+1}}$$

Because $\left(\frac{v-a}{v-\phi_0}\right)^n g(v)$ is zero when $v = a$, integrating from a to v gives

$$\frac{(v-a)^n}{(v-\phi_0)^n} g(v) = \int_{t=a}^{t=v} \left[\frac{(t-a)^{n-1}}{(t-\phi_0)^n} - (1-\lambda)\frac{(t-a)^n}{(t-\phi_0)^{n+1}}\right] dt$$
$$g(v) = \frac{(v-\phi_0)^n}{(v-a)^n} \int_{t=a}^{t=v} \left[\frac{(t-a)^{n-1}}{(t-\phi_0)^n} - (1-\lambda)\frac{(t-a)^n}{(t-\phi_0)^{n+1}}\right] dt$$

In order for this to be valid, $g(v)$ must be an increasing function in v and $\lim_{v\downarrow a} g(v)$ must be a constant.

Note here that we want $0 < \phi_0 < a$.

Using Lemma 2, we can write

$$\begin{aligned} g(v) &= \frac{(v-\phi_0)^n}{(v-a)^n} \left(\lambda \log\left(\frac{v-\phi_0}{a-\phi_0}\right) - \lambda \sum_{k=1}^{n-1} \frac{(v-a)^k}{k(v-\phi_0)^k} + (1-\lambda)\frac{(v-a)^n}{n(v-\phi_0)^n}\right) \\ &= \frac{(v-\phi_0)^n}{(v-a)^n} \left[\lambda \sum_{k=n}^{\infty} \frac{(v-a)^k}{(v-\phi_0)^k} + (1-\lambda)\frac{(v-a)^n}{n(v-\phi_0)^n}\right] \\ &= \frac{1}{n} + \lambda \sum_{k=n+1}^{\infty} \frac{(v-a)^{k-n}}{k(v-\phi_0)^{k-n}} \end{aligned}$$

This expression immediately tells us both that $\lim_{v\downarrow a} g(v) = 1/n$ and $g(v)$ is increasing in v, as desired.

Lastly, we need $g(b) = 1$. That is, ϕ_0 is a solution to

$$\frac{(b-\phi_0)^n}{(b-a)^n} \left(\lambda \log\left(\frac{b-\phi_0}{a-\phi_0}\right) - \lambda \sum_{k=1}^{n-1} \frac{(b-a)^k}{k(b-\phi_0)^k} + (1-\lambda)\frac{(b-a)^n}{n(b-\phi_0)^n}\right) = \frac{1-\lambda}{n} + \lambda \sum_{k=n+1}^{\infty} \frac{(b-a)^{k-n}}{k(b-\phi_0)^{k-n}} = 1$$

To solve this equation, consider the LHS as a function of $\phi_0 \in [0, a]$. It is clear from the expression that it is increasing in ϕ_0. As $\phi_0 \uparrow a$, we have $LHS \uparrow \frac{1}{n} + \lambda \sum_{k=n+1}^{\infty} \frac{1}{k} = \infty$, so there is a unique solution if and only if that value of LHS at $\phi_0 = 0$ is ≤ 1, that is,

$$\frac{b^n}{(b-a)^n} \left(\lambda \log\left(\frac{b}{a}\right) - \lambda \sum_{k=1}^{n-1} \frac{(b-a)^k}{kb^k} + (1-\lambda)\frac{(b-a)^n}{nb^n}\right) = \frac{1-\lambda}{n} + \lambda \sum_{k=n+1}^{\infty} \frac{(b-a)^{k-n}}{kb^{k-n}} \leq 1$$

That is, $a/b \geq k_h$ where $k_h \in (0,1)$ is a solution to

$$\frac{1}{(1-k_h)^n}\left(\lambda \log\left(\frac{1}{k_h}\right) - \lambda \sum_{k=1}^{n-1} \frac{(1-k_h)^k}{k} + (1-\lambda)\frac{(1-k_h)^n}{n}\right)$$
$$= \frac{1-\lambda}{n} + \lambda \frac{\log(1/k_h) - \sum_{k=1}^{n-1} \frac{(1-k_h)^k}{k}}{(1-k_h)^n} = \frac{1}{n} + \lambda \sum_{k=n+1}^{\infty} \frac{(1-k_h)^{k-n}}{k} = 1$$

Because the expression of k_h is increasing (this is clear from the infinite sum expression), k_h is unique.

Furthermore, we can check that the above solution in ϕ_0 in fact has an explicit solution:

$$\phi_0 = \frac{a - k_h b}{1 - k_h}$$

With this ϕ_0, we have $a - \phi_0 = \frac{k_h(b-a)}{1-k_h}$, $b - \phi_0 = \frac{(b-a)}{1-k_h}$, so the above equation reduces to the identity

$$\frac{1}{(1-k_h)^n}\left(\lambda \log\left(\frac{1}{k_h}\right) - \lambda \sum_{k=1}^{n-1} \frac{(1-k_h)^k}{k} + (1-\lambda)\frac{(1-k_h)^n}{n}\right) = 1$$

as desired.

Note also that from the above calculation, an alternative specification for k_h, from the original expression is that k_h is the value of a/b such that

$$\int_{t=a}^{t=b} \left[\frac{(t-a)^{n-1}}{t^n} - (1-\lambda)\frac{(t-a)^n}{t^{n+1}}\right] dt = \frac{(b-a)^n}{b^n}$$

or, equivalently,

$$\int_{t=k_h}^{t=1} \left[\frac{(t-a)^{n-1}}{t^n} - (1-\lambda)\frac{(t-a)^n}{t^{n+1}}\right] dt = (1-k_h)^n$$

Moderate Information ($k_l \leq a/b \leq k_h$) Regime. Here, we consider the intermediate a/b regime. Here, we set $r^* = a$, so there are two regimes, (ODE-g-1) for $v \in (a, v^*)$ and (ODE-g-2) for $v \in (v^*, b)$ with the requirement that g is continuous at v^*.

We can write (ODE-g-1) as

$$\frac{d}{dv}\left[\frac{(v-a)^{n-1}}{v^{n-1}}(g^*(v)-\alpha)\right]=\lambda\frac{(v-a)^{n-1}}{v^n}\text{ for } v\in[a,v^*]$$

or

$$\frac{(v-a)^{n-1}}{v^{n-1}}(g^*(v)-\alpha)=\lambda\int_{t=a}^{t=v}\frac{(t-a)^{n-1}}{t^n}dt=\lambda\left[\log\left(\frac{v}{a}\right)-\sum_{k=1}^{n-1}\frac{(v-a)^k}{kv^k}\right]$$

For $v\in(v^*,b)$ we have (ODE-g-2) with $\phi_0=0$, so

$$\frac{d}{dv}\left[\left(\frac{v-a}{v}\right)^n g^*(v)\right]=\frac{(v-a)^{n-1}}{v^n}-(1-\lambda)\frac{(v-a)^n}{v^{n+1}}$$

or

$$\frac{(b-a)^n}{b^n}g^*(b)-\frac{(v-a)^n}{v^n}g^*(v)=\int_{t=v}^{t=b}\left[\frac{(t-a)^{n-1}}{t^n}-(1-\lambda)\frac{(t-a)^n}{t^{n+1}}\right]dt$$

Using the condition $g(b)=1$ gives the formula for $g(v)$ for $v\in[v^*,b]$.

We assume that $g_u(a)=g_d(a)=\alpha$, so $g(a)=\alpha$. Because $g_u(v)+(n-1)g_d(v)=\lambda$ at $v=v^*$, and $g_d(v^*)=\alpha$, we have $g(v^*)=1-(n-1)\alpha$. We assume that $g(b)=1$.

Integrating the ODE that describes g from a to v^* and from v^* to b gives

$$\frac{(v^*-a)^{n-1}}{(v^*)^{n-1}}(g^*(v^*)-\alpha)=\lambda\int_{t=a}^{t=v^*}\frac{(t-a)^{n-1}}{t^n}dt$$
$$\frac{(b-a)^n}{b^n}g^*(b)-\frac{(v^*-a)^n}{(v^*)^n}g^*(v^*)=\int_{t=v^*}^{t=b}\left[\frac{(t-a)^{n-1}}{t^n}-(1-\lambda)\frac{(t-a)^n}{t^{n+1}}\right]dt$$

This gives the two equations that pin down the two constants v^* and α as follows.

$$\frac{(v^*-a)^{n-1}}{(v^*)^{n-1}}(1-n\alpha)=\lambda\int_{t=a}^{t=v^*}\frac{(t-a)^{n-1}}{t^n}dt$$
$$\frac{(b-a)^n}{b^n}-\frac{(v^*-a)^n}{(v^*)^n}(1-(n-1)\alpha)=\int_{t=v^*}^{t=b}\left[\frac{(t-a)^{n-1}}{t^n}-(1-\lambda)\frac{(t-a)^n}{t^{n+1}}\right]dt$$

Eliminating α from the two equations gives

$$n - \frac{n(v^*)^n}{(v^*-a)^n}\left(\frac{(b-a)^n}{b^n} - \int_{t=v^*}^{t=b}\left[\frac{(t-a)^{n-1}}{t^n} - (1-\lambda)\frac{(t-a)^n}{t^{n+1}}\right]dt\right)$$
$$= (n-1) - \frac{(n-1)(v^*)^{n-1}}{(v^*-a)^{n-1}}\lambda\int_{t=a}^{t=v^*}\frac{(t-a)^{n-1}}{t^n}dt$$

or

$$\frac{n(b-a)^n}{b^n} = \frac{(r^*-a)^n}{(r^*)^n} + \frac{(n-1)(r^*-a)}{(r^*)}\lambda\int_{t=a}^{t=r^*}\frac{(t-a)^{n-1}}{t^n}dt + n\int_{t=r^*}^{t=b}\left[\frac{(t-a)^{n-1}}{t^n} - (1-\lambda)\frac{(t-a)^n}{t^{n+1}}\right]dt$$

Let $\mathrm{fn}(v)$ denote the RHS with v^*, so

$$\mathrm{fn}(v) := \frac{(v-a)^n}{v^n} + \frac{(n-1)(v-a)}{v}\lambda\int_{t=a}^{t=v}\frac{(t-a)^{n-1}}{t^n}dt + n\int_{t=v}^{t=b}\left[\frac{(t-a)^{n-1}}{t^n} - (1-\lambda)\frac{(t-a)^n}{t^{n+1}}\right]dt$$

We claim that $\mathrm{fn}(v)$ is a decreasing function. That is, we want to show that $d\mathrm{fn}(v)/dv = \mathrm{fn}'(v) \leq 0$. We compute

$$\mathrm{fn}'(v) = n\left(\frac{v-a}{v}\right)^{n-1}\frac{a}{v^2} + (n-1)\lambda\frac{d}{dv}\left[\frac{(v-a)}{v}\int_{t=a}^{t=v}\frac{(t-a)^{n-1}}{t^n}dt\right] - n\left(\frac{(v-a)^{n-1}}{v^n} - (1-\lambda)\frac{(v-a)^n}{v^{n+1}}\right)$$

Note that

$$n\left(\frac{v-a}{r}\right)^{n-1}\frac{a}{v^2} = \frac{na(v-a)^{n-1}}{v^{n+1}} = \frac{n(v-(v-a))(v-a)^{n-1}}{v^{n+1}} = \frac{n(v-a)^{n-1}}{v^n} - \frac{n(v-a)^n}{v^{n+1}}$$

We then have

$$\begin{aligned}\mathrm{fn}'(v) &= -n\lambda\frac{(v-a)^n}{v^{n+1}} + (n-1)\lambda\frac{d}{dv}\left[\frac{(v-a)}{v}\int_{t=a}^{t=v}\frac{(t-a)^{n-1}}{t^n}dt\right]\\ &= -n\lambda\frac{(v-a)^n}{v^{n+1}} + (n-1)\lambda\left[\frac{(v-a)}{v}\frac{(v-a)^{n-1}}{v^n} + \frac{a}{v^2}\int_{t=a}^{t=v}\frac{(t-a)^{n-1}}{t^n}dt\right]\\ &= \lambda\left[-\frac{(v-a)^n}{v^{n+1}} + \frac{(n-1)a}{v^2}\int_{t=a}^{t=v}\frac{(t-a)^{n-1}}{t^n}dt\right]\end{aligned}$$

Therefore, we have $\mathrm{fn}'(v) \le 0$ if and only if

$$\int_a^v \frac{(t-a)^{n-1}}{t^n} dt \le \frac{1}{(n-1)a} \frac{(v-a)^n}{v^{n-1}}$$

We can prove this inequality as follows. Both sides are zero for $v = a$, so it is sufficient to show that the derivative of the LHS is $\le$ the derivative of the RHS. This is true because the derivative of the LHS is $(v-a)^{n-1}/v^n$ and the derivative of the RHS is

$$\begin{aligned}\frac{1}{(n-1)a} \frac{v^{n-1} n (v-a)^{n-1} - (v-a)^n (n-1) v^{n-2}}{v^{2n-2}} &= \frac{1}{(n-1)a} \frac{(v-a)^{n-1}}{v^n} (nv - (n-1)(v-a)) \\ &= \frac{(v-a)^{n-1}}{v^n} + \frac{(v-a)^{n-1}}{(n-1)a v^{n-1}} \ge \frac{(v-a)^{n-1}}{v^n}\end{aligned}$$

Therefore, we have proved that $\mathrm{fn}(v)$ is decreasing in v.

Therefore, to show that the equation $\mathrm{fn}(v^*) = n \left(\frac{b-a}{b}\right)^n$ has a unique solution $v^* \in [a, b]$, it is sufficient to show that $\mathrm{fn}(a) \ge n \left(\frac{b-a}{b}\right)^n \ge \mathrm{fn}(b)$, or

$$n \int_a^b \left[\frac{(t-a)^{n-1}}{t^n} - (1-\lambda) \frac{(t-a)^n}{t^{n+1}} \right] dt \ge n \left(\frac{b-a}{b}\right)^n \ge \left(\frac{b-a}{b}\right)^n + \frac{(n-1)(b-a)}{b} \lambda \int_a^b \frac{(t-a)^{n-1}}{t^n} dt$$

The first inequality

$$\int_a^b \left[\frac{(t-a)^{n-1}}{t^n} - (1-\lambda) \frac{(t-a)^n}{t^{n+1}} \right] dt \ge \left(\frac{b-a}{b}\right)^n$$

is true by the definition of k_h and $a/b \le k_h$. The second inequality is equivalent to

$$\lambda \int_{t=a}^{t=b} \frac{(t-a)^{n-1}}{t^n} dt \le \frac{(b-a)^{n-1}}{b^{n-1}}$$

which is true by the definition of k_l and $a/b \ge k_l$.

We also conclude from the above that as $a/b \uparrow k_h$ we have $r^* \downarrow a$, while as $a/b \downarrow k_l$, we have $r^* \uparrow b$.

Now we will show that the resulting α is in $[0, 1]$, so it is valid. In fact, we will show something stronger: that $\alpha \in [0, 1/n]$ and as $a/b \uparrow k_h$ we have $\alpha \uparrow 1/n$, while as $a/b \downarrow k_l$, we have $\alpha \downarrow 0$.

We will first show that for any $v \in [a, b]$, we have $\int_a^v \frac{(t-a)^{n-1}}{t^n} dt \le \frac{1}{\lambda} \frac{(v-a)^{n-1}}{v^{n-1}}$. Let $\mathrm{fn}(v) := \int_a^v \frac{(t-a)^{n-1}}{t^n} dt - \frac{1}{\lambda} \frac{(v-a)^{n-1}}{v^{n-1}}$. (We overload the fn notation here — it has nothing to do with the earlier fn; it is just a shorthand that we discard after we finish proving the technical statement.) Note that $\mathrm{fn}(a) = 0$ and $\mathrm{fn}(b) \le 0$ by the definition of $a/b \ge k_l$. We have

$$\mathrm{fn}'(v) = \frac{(v-a)^{n-1}}{v^n} - \frac{1}{\lambda}(n-1)\frac{(v-a)^{n-2}}{v^{n-2}}\frac{a}{v^2} = \frac{(v-a)^{n-2}}{v^n}\left(v - \left(1 + \frac{n-1}{\lambda}\right)a\right)$$

So $\mathrm{fn}(v)$ is decreasing for $v \le \left(1 + \frac{n-1}{\lambda}\right)a$ and increasing for $v \ge \left(1 + \frac{n-1}{\lambda}\right)a$. Regardless of whether $b \le \left(1 + \frac{n-1}{\lambda}\right)a$ or not, we have $\mathrm{fn}(v) \le \max(\mathrm{fn}(a), \mathrm{fn}(b)) \le 0$, and we are done. Note also that as $a/b \downarrow k_l$, we have $r^* \uparrow b$ so every inequality here approaches equality, and $\alpha \downarrow 0$.

From the defining equation we have

$$1 - n\alpha = \lambda\frac{(r^*)^{n-1}}{(r^*-a)^{n-1}}\int_a^{r^*}\frac{(t-a)^{n-1}}{t^n}dt$$

but by the lemma we have just proved,

$$\int_a^{r^*}\frac{(t-a)^{n-1}}{t^n}dt \le \frac{1}{\lambda}\frac{(r^*-a)^{n-1}}{(r^*)^{n-1}}$$

so $1 - n\alpha \le 1$, which implies $\alpha \ge 0$.

From the same equation, it is clear that $\frac{(r^*)^{n-1}}{(r^*-a)^{n-1}}\int_a^{r^*}\frac{(t-a)^{n-1}}{t^n}dt \ge 0$, so $1 - n\alpha \ge 0$, so $\alpha \le 1/n$. Furthermore, as $a/b \uparrow k_h$, we have $r^* \downarrow a$ so

$$1 - n\alpha = \lambda\frac{(r^*)^{n-1}}{(r^*-a)^{n-1}}\int_a^{r^*}\frac{(t-a)^{n-1}}{t^n}dt \le \lambda\frac{(r^*)^{n-1}}{(r^*-a)^{n-1}}\int_a^{r^*}\frac{(t-a)^{n-1}}{a^n}dt = \lambda\frac{(r^*)^{n-1}(r^*-a)}{na^n} \downarrow 0$$

so $\alpha \uparrow 1/n$.

We conclude that for $k_l \le a/b \le k_h$, there is a unique valid solution (r^*, α). Furthermore, as $a/b \uparrow k_h$ we have $r^* \downarrow a$ and $\alpha \uparrow 1/n$, while as $a/b \downarrow k_l$, we have $r^* \uparrow b$ and $\alpha \downarrow 0$. We are done!

Checking (SOC). The following applies whether we are in the regime $a/b \le k_l$, $a/b \ge k_h$, or $k_l \le a/b \le k_h$.

In the case $v \le v^*$, we have $g_u^*(v) = g^*(v)$ and $g_d^*(v) = \alpha$, so (SOC) reduces to $g^*(v) - \alpha > 0$, which is true by definition of g^* in the interior.

In the case $v \ge v^*$, we have $g_u^*(v) = g^*(v)$ and $g_d^*(v) = (1 - g^*(v))/(n-1)$, so (SOC) reduces to $ng^*(v) - 1 + (v-a)(g^*)'(v) > 0$.

Here, g^* satisfies (ODE-g-2) with $\phi_0 = 0$ in the case $k_l \leq a/b \leq k_h$ and $\phi_0 \geq 0$ in the case $a/b \geq k_h$. Substituting $(g^*)'(v) = \frac{1}{v-a} - \frac{1-\lambda}{v-\phi_0} - \frac{n(a-\phi_0)}{(v-\phi_0)(v-a)} g^*(v)$, we get that (SOC) reduces to $ng^*(v) - 1 + (v-a)(g^*)'(v) = \frac{(v-a)}{(v-\phi_0)}(ng^*(v) - 1 + \lambda) > 0$, so we have to prove that $ng^*(v) - 1 + \lambda > 0$. From $v \geq v^*$ we have $g^*(v) \geq 1 - (n-1)\alpha$, we have $ng^*(v) - 1 + \lambda \geq n(1 - (n-1)\alpha) - 1 + \lambda = (n-1)(1-\alpha) + \lambda \geq 0$, because $\alpha \leq 1, \lambda \geq 0$, with strict inequality everywhere but the boundary, as desired.

□

Proof of Corollary 1. We will use the equivalent formulation of Theorem 5. k_l is a solution to $\lambda \log(1/k_l) = 1$, so $k_l = \exp(-1/\lambda)$. k_h is a solution to $(1-\lambda)(1-k_h) + \lambda \log(1/k_h) = (1-k_h)$ or $\log(1/k_h) = (1-k_h)$, so $k_h = 1$.

Therefore, we only need to consider the regime $a/b \leq k_l$ and $k_l \leq a/b \leq k_h$.

For $a/b \leq k_l$, the regret is b/e. For $a/b \geq k_l$, we know that (v^*, α) is a solution to the following system of equations

$$
\begin{aligned}
1 - \alpha &= \lambda \log\left(\frac{v^*}{a}\right) \\
\frac{b-a}{b} - \frac{v^* - a}{v^*} &= \lambda \log\left(\frac{b}{v^*}\right) - \frac{(1-\lambda)a}{b} + \frac{(1-\lambda)a}{v^*}
\end{aligned}
$$

The solution to this is $v^* = b$ and $\alpha = 1 - \lambda \log(b/a)$.

Substituting $n = 1$ to the minimax λ-regret expression gives, for $a/b \leq k_l$,

$$
\begin{aligned}
&-(1-\lambda)b + \left[(1-k_l) - \lambda \int_{\tilde{r}=k_l}^{\tilde{r}=1} \left(1 - \frac{k_l}{\tilde{r}}\right) d\tilde{r}\right] b \\
&= -(1-\lambda)b + \left[(1-k_l) - \lambda(1 - k_l + k_l \log(k_l))\right] b \\
&= (-1 + \lambda + 1 - e^{-1/\lambda} - \lambda + \lambda e^{-1/\lambda} + e^{-1/\lambda}) b \\
&= \lambda e^{-1/\lambda} b
\end{aligned}
$$

and for $a/b \geq k_l$,

$$
\begin{aligned}
&-(1-\lambda)b + b\left(1 - \frac{a}{b}\right) + (b-a)\alpha(1-1) - \lambda \int_{v=a}^{v=b} \left(1 - \frac{a}{v}\right) dv \\
&= -(1-\lambda)b + (b-a) - \lambda(b - a - a\log(b/a)) = -a + \lambda a + \lambda \log(b/a)
\end{aligned}
$$

Therefore, the minimax regret is when $\lambda = 1$, which is b/e for $a/b \le 1/e$ and $a\log(b/a)$ for $a/b \ge 1/e$. The maximin ratio is λ such that the minimax λ-regret is zero. Because $\lambda e^{-1/\lambda} b > 0$ always, the regime $a/b \le k_l$ does not apply. The λ such that $-a + \lambda a + \lambda \log(b/a) = 0$ is $\lambda = 1/(1 + \log(b/a))$. $\square$

E Proofs and Discussions from Section 4

E.1 Proofs and Discussions from Section 4.1

Before we prove the main theorem characterizing the minimax λ-regret standard mechanism (Theorem 3), we first derive the regret expression of generous SPA GenSPA(Φ).

Proposition 9. *The regret of the mechanism* GenSPA(Φ) *under distribution* $\boldsymbol{F}$ *that is* n *i.i.d. with marginal* F *is*

$$a(\lambda - \Phi(a)) + \int_{v=a}^{v=b} (\lambda - \Phi(v) - v\Phi'(v))(1 - F(v)^n) + \Phi(v)(nF(v)^{n-1} - nF(v)^n)dv$$
$$+ F(a)^{n-1}\left[(b-a)(n-(n-1)F(a)) - \int_{v=a}^{v=b} \left((n-(n-1)F(a))\Phi(v) + v(nF(v) - (n-1)F(a))\Phi'(v)\right)dv\right]$$

Proof of Proposition 9. The λ-regret expression, pointwise at $\boldsymbol{v}$, is

$$\lambda v^{(1)} - \left(v^{(1)}\Phi(v^{(1)}) - \int_{t=v^{(2)}}^{t=v^{(1)}} \Phi(t)dt\right)\mathbf{1}(v^{(2)} > a) - \left(v^{(1)} \cdot 1 - \int_{t=v^{(2)}}^{t=v^{(1)}} 1dt\right)\mathbf{1}(v^{(2)} = a)$$
$$= \lambda v^{(1)} - \left(v^{(1)}\Phi(v^{(1)}) - \int_{t=v^{(2)}}^{t=v^{(1)}} \Phi(t)dt\right)\mathbf{1}(v^{(2)} > a) - \left(v^{(2)}\right)\mathbf{1}(v^{(2)} = a)$$

Writing $\mathbf{1}(v^{(2)} > a) = 1 - \mathbf{1}(v^{(2)} = a)$, the regret expression becomes

$$\left(\lambda v^{(1)} - v^{(1)}\Phi(v^{(1)}) + \int_{t=v^{(2)}}^{t=v^{(1)}} \Phi(t)dt\right) + \left(v^{(1)}\Phi(v^{(1)}) - \int_{t=v^{(2)}}^{t=v^{(1)}} \Phi(t)dt - v^{(2)}\right)\mathbf{1}(v^{(2)} = a)$$
$$= \left(\lambda v^{(1)} - v^{(1)}\Phi(v^{(1)}) + \int_{t=v^{(2)}}^{t=v^{(1)}} \Phi(t)dt\right) + \left(v^{(1)}\Phi(v^{(1)}) - \int_{t=a}^{t=v^{(1)}} \Phi(t)dt - a\right)\mathbf{1}(v^{(2)} = a)$$

We will now calculate the distribution of $v^{(1)}|v^{(2)} = a$.

We note that, for any $x \in [a, b]$,

$$\Pr(v^{(1)} > x, v^{(2)} = a) = \Pr(\ n-1\ v\text{'s are } a\text{, one is } > x) = nF(a)^{n-1}F((x,b])) = nF(a)^{n-1}(1 - F(x))$$

and

$$\begin{aligned}\Pr(v^{(2)} = a) &= \Pr(\text{ exactly } n-1 \text{ are } a\ ) + \Pr(\text{ exactly } n \text{ are } a\ ) \\ &= nF(a)^{n-1}(1 - F(a)) + F(a)^n = F(a)^{n-1}(n - (n-1)F(a))\end{aligned}$$

If $\Pr(v^{(2)} = a) > 0$ (that is, if F has an atom at a), then dividing the two equations gives

$$\Pr(v^{(1)} > x | v^{(2)} = a) = \frac{n(1 - F(x))}{n - (n-1)F(a)} \qquad \text{for } x \in [a, b]$$

This gives us the CDF of $v^{(1)}|v^{(2)} = a$ as

$$\tilde{F}(x) := \Pr(v^{(1)} \leq x | v^{(2)} = a) = 1 - \frac{n(1 - F(x))}{n - (n-1)F(a)} = \frac{nF(x) - (n-1)F(a)}{n - (n-1)F(a)} \qquad \text{for } x \in [a, b]$$

(If $\Pr(v^{(2)} = a) = 0$, then the expression we want to evaluate is zero and we won't need this conditional distribution anyway.)

We can now calculate the expected regret. The first term is exactly the expected regret expression from SPA:

$$\begin{aligned}&\mathbb{E}_{\boldsymbol{v} \sim F}\left[\lambda v^{(1)} - \left(v^{(1)}\Phi(v^{(1)}) - \int_{t=v^{(2)}}^{t=v^{(1)}} \Phi(t)dt\right)\right] \\ &= a(\lambda - \Phi(a)) + \int_{v=a}^{v=b} (\lambda - \Phi(v) - v\Phi'(v))(1 - F(v)^n) + \Phi(v)(nF(v)^{n-1} - nF(v)^n)dv\end{aligned}$$

The second term can be written as

$$
\begin{aligned}
&\mathbb{E}_{\boldsymbol{v}\sim F}\left[\left(v^{(1)}\Phi(v^{(1)}) - \int_{t=v^{(2)}}^{t=v^{(1)}} \Phi(t)dt - a\right)\mathbf{1}(v^{(2)}=a)\right] \\
&= \Pr(v^{(2)}=a)\mathbb{E}\left[\left(v^{(1)}\Phi(v^{(1)}) - \int_{t=a}^{t=v^{(1)}} \Phi(t)dt - a\right)\middle| v^{(2)}=a\right] \\
&= F(a)^{n-1}(n-(n-1)F(a))\left[-a + \int_{v\in[a,b]}\left(v\Phi(v) - \int_{t=a}^{t=v}\Phi(t)dt\right)d\tilde{F}(v)\right]
\end{aligned}
$$

We now use the following integration-by-part-like statements.

$$
\int_{w\in[a,b]} h(w)dG(w) = h(b) - \int_{v=a}^{v=b} h'(v)G(v)dv
$$

Therefore,

$$
\begin{aligned}
&\int_{v\in[a,b]} v\Phi(v)d\tilde{F}(v) \\
&= b\Phi(b) - \int_{v=a}^{v=b}(\Phi(v)+v\Phi'(v))\tilde{F}(v)dv \\
&= b - \int_{v=a}^{v=b}(\Phi(v)+v\Phi'(v))\frac{nF(v)-(n-1)F(a)}{n-(n-1)F(a)}dv
\end{aligned}
$$

We also have

$$
\begin{aligned}
&\int_{v\in[a,b]}\int_{t=a}^{t=v}\Phi(t)dtd\tilde{F}(v) \\
&= \int_{t=a}^{t=b}\int_{v\in(t,b]}\Phi(t)d\tilde{F}(v)dt \\
&= \int_{t=a}^{t=b}\Phi(t)(1-\tilde{F}(t))dt \\
&= \int_{v=a}^{v=b}\Phi(v)\frac{n(1-F(v))}{n-(n-1)F(a)}dv
\end{aligned}
$$

Therefore, the second term is

$$F(a)^{n-1}(n-(n-1)F(a))\left[-a+b-\int_{v=a}^{v=b}(\Phi(v)+v\Phi'(v))\frac{nF(v)-(n-1)F(a)}{n-(n-1)F(a)}-\int_{v=a}^{v=b}\Phi(v)\frac{n(1-F(v))}{n-(n-1)F(a)}dv\right]$$
$$=F(a)^{n-1}(n-(n-1)F(a))\left[b-a-\int_{v=a}^{v=b}\left(\Phi(v)+v\Phi'(v)\frac{nF(v)-(n-1)F(a)}{n-(n-1)F(a)}\right)dv\right]$$
$$=F(a)^{n-1}\left[(b-a)(n-(n-1)F(a))-\int_{v=a}^{v=b}\big((n-(n-1)F(a))\Phi(v)+v(nF(v)-(n-1)F(a))\Phi'(v)\big)\,dv\right]$$

We therefore have the regret expression

$$=a(\lambda-\Phi(a))+\int_{v=a}^{v=b}(\lambda-\Phi(v)-v\Phi'(v))(1-F(v)^n)+\Phi(v)(nF(v)^{n-1}-nF(v)^n)dv$$
$$+F(a)^{n-1}\left[(b-a)(n-(n-1)F(a))-\int_{v=a}^{v=b}\big((n-(n-1)F(a))\Phi(v)+v(nF(v)-(n-1)F(a))\Phi'(v)\big)\,dv\right]$$

Note that the extra term (in the second line) is linear in F).

□

We are now ready to state and prove the main theorem.

Theorem 6. *Fix n and $\lambda \in (0,1]$, and let $\tilde{a} = a/b \in [0,1)$. Define k_l as in Theorem 1. The minimax λ-regret problem $R_{\lambda,n}(\mathcal{M}_{\text{std}}, \mathcal{F}_{\text{iid}})$ admits the following saddle point (m^*, F^*), depending on a/b, as follows.*

- *For $a/b \le k_l$, the optimal mechanism m^* and worst-case distribution F^* are the same as those of Theorem 2.*

- *For $a/b \ge k_l$, the optimal mechanism m^* is $\text{GenSPA}(\Phi^*)$ with*

$$\Phi^*(v) = \frac{\int_{t=a}^{t=v} \frac{\lambda(1-c/t)^{n-1}}{t}dt}{(1-c/v)^{n-1}-(1-c/a)^{n-1}} \qquad \text{for } v \in [a,b]$$

where $c \in [0,a]$ is a unique constant such that $\Phi^*(c) = 1$. The worst case distribution F^* is $\text{IsoRev}(a;c)$ defined by $F^*(v) = 1 - c/v$ for $v \in [a,b)$ and $F^*(b) = 1$.

Remark. Note that the worst case distribution $\text{IsoRev}(a;c)$ has two point masses at $v = a$ and $v = b$ (of size $1-c/a$ and c/b respectively), whereas the worst case distributions of $\mathcal{M}_{\text{all}}$ in Theorem 2 each only has one point mass at $v = b$. Also, the distributions Φ and Ψ of the reserve of SPA and threshold of POOL in Theorem 2 do not have any point masses, whereas the distribution Φ of GenSPA in Theorem 6 has a point mass at $v = a$ of size $\Phi^*(a) = \lim_{v \downarrow a} \Phi^*(v) = \lambda(a/c - 1)/(n-1)$.

Proof of Theorem 6. Seller's saddle is that, fixing F^*, the giving mechanism m^* gives the lowest regret over all standard mechanisms. Because F^* is fixed, this is equivalent to that m^* maximizes expected revenue under F^*. This is a standard Bayesian mechanism design problem, and we check with Monteiro and Svaiter [2010] that m^* is indeed optimal, even over all DSIC mechanisms. Henceforth, we will focus on Nature's Saddle.

Note that if $a/b \leq k_l$, then Theorem 1 immediately tells us that the same SPA(Φ^*) is optimal over $\mathcal{M}_{\text{all}}$, and thus over $\mathcal{M}_{\text{std}}$ also. Henceforth we assume $a/b \geq k_l$.

We have the regret expression from Proposition 9:

$$a(\lambda - \Phi(a)) + \int_{v=a}^{v=b} (\lambda - \Phi(v) - v\Phi'(v))(1 - F(v)^n) + \Phi(v)(nF(v)^{n-1} - nF(v)^n)dv$$
$$+ F(a)^{n-1}\left[(b-a)(n-(n-1)F(a)) - \int_{v=a}^{v=b} \big((n-(n-1)F(a))\Phi(v) + v(nF(v) - (n-1)F(a))\Phi'(v)\big)\, dv\right]$$

The additional term is linear in F, so if pointwise optimization gives a global maximum then, it will also give a global maximum now. The first-order condition gives

$$(\lambda - \Phi(v) - v\Phi'(v))(-nF(v)^{n-1}) + \Phi(v)(n(n-1)F(v)^{n-2} - n^2F(v)^{n-1}) - nvF(a)^{n-1}\Phi'(v) = 0$$

or

$$v(F(v)^{n-1} - F(a)^{n-1})\Phi'(v) + (n-1)F(v)^{n-2}(1 - F(v))\Phi(v) = \lambda F(v)^{n-1}$$

We want Φ such that $F(v) = 1 - c/v$ is a solution to that equation. With $F(v) = 1 - c/v$ we have $F'(v) = \frac{c}{v^2} = \frac{1-F(v)}{v}$. So

$$(F(v)^{n-1} - F(a)^{n-1})\Phi'(v) + (n-1)F(v)^{n-2}F'(v)\Phi(v) = \frac{\lambda F(v)^{n-1}}{v}$$

or

$$\frac{d}{dv}\left[(F(v)^{n-1} - F(a)^{n-1})\Phi(v)\right] = \frac{\lambda F(v)^{n-1}}{v}$$

Because $(F(v)^{n-1} - F(a)^{n-1})\Phi(v)$ is zero when $v = a$, we have

$$\Phi(v) = \frac{\int_{t=a}^{t=v} \frac{\lambda F(t)^{n-1}}{t} dt}{F(v)^{n-1} - F(a)^{n-1}} = \frac{\int_{t=a}^{t=v} \frac{\lambda\left(1-\frac{c}{t}\right)^{n-1}}{t} dt}{\left(1-\frac{c}{v}\right)^{n-1} - \left(1-\frac{c}{a}\right)^{n-1}}$$

We will now show that, for any $c \in (0, a]$, $\Phi(v)$ is increasing in v.

Let $\tilde{v} = 1 - c/v$, $\tilde{a} = 1 - c/a$, and use the substitution $u = 1 - c/t$, $t = \frac{c}{1-u}$, $dt = \frac{c}{(1-u)^2}du$ to get

$$\frac{1}{\lambda}\Phi(v) = \frac{\int_{u=\tilde{a}}^{u=\tilde{v}} \frac{u^{n-1}}{1-u} du}{\tilde{v}^{n-1} - \tilde{a}^{n-1}}$$

Showing that $\Phi(v)$ is increasing in v is equivalent to showing that the right hand side is increasing in $\tilde{v} = 1 - c/v$. From $v \geq a$ we have $\tilde{v} \geq \tilde{a}$. The derivative with respect to $\tilde{v}$ of the right hand side is

$$\frac{(\tilde{v}^{n-1} - \tilde{a}^{n-1})\frac{\tilde{v}^{n-1}}{1-\tilde{v}} - \left(\int_{u=\tilde{a}}^{u=\tilde{v}} \frac{u^{n-1}}{1-u} du\right)(n-1)\tilde{v}^{n-2}}{(\tilde{v}^{n-1} - \tilde{a}^{n-1})^2}$$

This expression is nonnegative if and only if

$$\int_{u=\tilde{a}}^{u=\tilde{v}} \frac{u^{n-1}}{1-u} du \leq \frac{\tilde{v}^n - \tilde{a}^{n-1}\tilde{v}}{(n-1)(1-\tilde{v})}$$

It is clear that this is true if it holds for $\tilde{v} = a$ and the derivative of LHS is $\leq$ the derivative of RHS. For $\tilde{v} = a$, both sides are zero, so the inequality holds. The condition that the derivative of LHS is $\leq$ the derivative of RHS is

$$\frac{\tilde{v}^{n-1}}{1-\tilde{v}} \leq \frac{1}{n-1}\frac{(1-\tilde{v})(n\tilde{v}^{n-1} - \tilde{a}^{n-1}) - (\tilde{v}^n - \tilde{a}^{n-1}\tilde{v})(-1)}{(1-\tilde{v})^2}$$

This is equivalent to

$$(n-1)\tilde{v}^{n-1} - (n-1)\tilde{v}^n \leq n\tilde{v}^{n-1} - \tilde{a}^{n-1} - n\tilde{v}^n + \tilde{a}^{n-1}\tilde{v} + \tilde{v}^n - \tilde{a}^{n-1}\tilde{v}$$

or

$$\tilde{a}^{n-1} \leq \tilde{v}^{n-1}$$

which is true because $\tilde{v} \geq \tilde{a}$.

Lastly, we will show that if $a/b \geq k_l$, then there is a $c \in (0, a]$ such that $\Phi(b) = 1$, making this a valid solution. From

$$\Phi(b) = \frac{\int_{t=a}^{t=b} \frac{\lambda\left(1-\frac{c}{t}\right)^{n-1}}{t} dt}{\left(1-\frac{c}{b}\right)^{n-1} - \left(1-\frac{c}{a}\right)^{n-1}}$$

Consider the right hand side as a function of c. The numerator is clearly a decreasing function of c, because $\frac{\lambda\left(1-\frac{c}{t}\right)^{n-1}}{t}$, for each fixed t, is a decreasing function of c. The denominator is an increasing function of c because

$$\left(1-\frac{c}{b}\right)^{n-1} - \left(1-\frac{c}{a}\right)^{n-1} = c\left(\frac{1}{a}-\frac{1}{b}\right)\left(\sum_{k=0}^{n-1}\left(1-\frac{c}{b}\right)^k\left(1-\frac{c}{a}\right)^{n-1-k}\right)$$

and each of c, $1-\frac{c}{b}$, $1-\frac{c}{a}$ are positive and increasing in c, and $\frac{1}{a}-\frac{1}{b} > 0$.

Therefore, the right hand side is decreasing in c. As $c \downarrow 0$, the numerator converges to $\int_{t=a}^{t=b} \frac{\lambda}{t} dt = \lambda \log(b/a) > 0$, and the denominator converges to 0, so the expression converges to $+\infty$. At $c = a$, the expression is

$$\frac{\int_{t=a}^{t=b} \frac{\lambda\left(1-\frac{a}{t}\right)^{n-1}}{t} dt}{\left(1-\frac{a}{b}\right)^{n-1}} \leq 1$$

where the ≤ 1 holds because $a/b \geq k_l$. Therefore, there is a $c \in (0, a]$ such that $\Phi(b) = 1$.

Lastly, we note that

$$\Phi(a) = \lim_{v \downarrow a} \Phi(v) = \lambda \lim_{v \downarrow a} \frac{\frac{\left(1-\frac{c}{v}\right)^{n-1}}{v}}{(n-1)\left(1-\frac{c}{v}\right)^{n-2}\frac{c}{v^2}} = \lambda \frac{(a-c)}{(n-1)c}$$

where the second equality holds by L'Hopital's rule. □

E.2 Proofs and Discussions from Section 4.2

The main theorem in the main text (Theorem 4) is an immediate corollary of the main theorem in this Appendix (Theorem 7). We first outline key challenges of the proof before diving into the full proof of the main theorem.

Key Challenges of the Proof of The Main Theorem. The proof of the moderate information regime is the hardest, and we outline key technical ideas here. Our candidate Φ^* has a point mass Φ_0 at a and a density on $[r^*, b]$, while the candidate worst-case distribution F^* has a point mass F_0, an isorevenue density on $[r^*, b)$, and a point mass at b. The seller's saddle is to find Φ that minimizes the regret $R(\Phi) := R(\Phi, F^*)$ such that $\Phi(v) \in [0, 1]$ is an increasing function. This is different from previous saddle problems because in this case, the increasing condition is *binding*; if we optimize pointwise, the result is nonincreasing, which is infeasible. We write the Lagrangian $\mathcal{L}(\Phi, \mu) = R(\Phi) - \int_a^b \mu(v) d\Phi(v)$. Here, $\mu : [a, b] \to \mathbb{R}_+$ is the dual variable associated with the increasing constraint. Complementary slackness requires that μ^* is zero wherever Φ^* is strictly increasing; this immediately suggests the correct form of Φ^*, and that $\mu^*(v) = 0$ on $[r^*, b]$. Lagrangian optimality requires that Φ^* also minimizes $\mathcal{L}(\Phi, \mu^*)$, which is linear in Φ (as can be made explicit by integration by part on the $d\Phi(v)$ term). We satisfy this by requiring that the coefficients of every $\Phi(v)$ term to be zero. These, together with complementary slackness, pin down μ^*. Nature's saddle is significantly simpler because the condition that F is increasing does not bind here, so pointwise optimization (as before) works. The full proof is given in Appendix E.2. □

Before we proceed to the proof, we derive the regret expression for SPA(Φ), which is just (g_u, g_d) with $g_u = \Phi$ and $g_d \equiv 0$.

Proposition 10. *Let the mechanism be a second-price auction with random reserve CDF Φ and distribution $\boldsymbol{F}$ with regret $R(\Phi, \boldsymbol{F})$.*

Suppose $\Phi : [a,b] \to [0,1]$ is absolutely continuous, while $\boldsymbol{F}$ is arbitrary then we have the (Regret-$\boldsymbol{F}$) expression

$$R(\Phi, \boldsymbol{F}) = a(\lambda - \Phi(a)) + \int_{v\in[a,b]} (\lambda - \Phi(v) - v\Phi'(v))(1 - F^{(1)}(v))dv + \int_{v\in[a,b]} \Phi(v)(F^{(2)}(v) - F^{(1)}(v))dv$$

If we further assume that $\boldsymbol{F}$ is n i.i.d. then

$$R(\Phi, \boldsymbol{F}) = a(\lambda - \Phi(a)) + \int_{v\in[a,b]} (\lambda - \Phi(v) - v\Phi'(v))(1 - F(v)^n)dv + \int_{v\in[a,b]} \Phi(v)(nF(v)^{n-1} - nF(v)^n)dv$$

Suppose instead that $\boldsymbol{F}$ is i.i.d. with marginal F that has a density in (a,b), and we denote by $F(\{b\})$ and $\boldsymbol{F}^{(1)}(b)$ the mass at b. Then we have the (Regret-Φ) expression

$$R(\Phi, F) = \lambda b - a\Phi(a)F(a)^n - (1 - (1 - f_b)^n)b\Phi(b) + \int_{v=a}^{v=b} -\lambda F(v)^n + \Phi(v)nF(v)^{n-1}(1 - F(v) - vF'(v))dv$$

Proof of Proposition 10. These expressions follow immediately by taking $g_u(v) = \Phi(v)$ and $g_d(v) = 0$.

□

We are now ready to state and prove the main theorem.

Theorem 7 (Optimal SPA with Random Reserve)**.** *Let $\tilde{a} = a/b \in [0,1)$. The minimax λ-regret problem $R_{\lambda,n}(\mathcal{M}_{\text{SPA-rand}}, \mathcal{F}_{\text{iid}})$ admits the following saddle point $(m^* = \text{SPA}(\Phi^*), F^*)$, depending on a/b, as follows.*

- *For $a/b \le k_l$, the optimal mechanism m^* and worst-case distribution F^* are the same as those of Theorem 1.*
- *For $a/b \ge \frac{\lambda n}{(1+\lambda)n-1}$, the optimal mechanism is a SPA with no reserve and the worst-case distribution F^* is a two-point distribution with point masses at $v = a$ and $v = b$ with weights $f_a := \frac{n-1}{n-1+\lambda}$ and $f_b := \frac{\lambda}{n-1+\lambda}$.*
- *For $k_l \le a/b \le \frac{\lambda n}{(1+\lambda)n-1}$, let $F_0 \in [0,1)$ be a unique solution to*

$$\frac{\lambda F_0^n}{(n-1)(1-F_0)} - \left(1 - \frac{n-(n-1)F_0}{n}\tilde{a}\right)^{n-1} = \lambda\left[\log\left(\frac{n-(n-1)F_0}{n(1-F_0)}\tilde{a}\right) - \sum_{k=1}^{n-1}\frac{1}{k}F_0^k + \sum_{k=1}^{n-1}\frac{1}{k}\left(1 - \frac{n-(n-1)F_0}{n}\tilde{a}\right)^k\right]$$

and

$$r^* = \frac{n-(n-1)F_0}{n(1-F_0)}a, \quad \Phi_0 = \frac{\lambda F_0}{(n-1)(1-F_0)}, \quad c = \frac{n-(n-1)F_0}{n}a, \quad d = \frac{F_0^n}{(n-1)(1-F_0)} + \log(1-F_0) + \sum_{k=1}^{n-1}\frac{1}{k}F_0^k$$

Then the optimal mechanism (optimal reserve distribution) is

$$\Phi^*(v) = \begin{cases} \Phi_0 & \text{for } v \in [a, r^*] \\ \lambda\left(\frac{v}{v-c}\right)^{n-1}\left[d + \log\left(\frac{v}{c}\right) - \sum_{k=1}^{n-1}\frac{1}{k}\left(\frac{v-c}{v}\right)^k\right] & \text{for } v \in [r^*, b] \end{cases}$$

and the worst-case distribution is given by $F^*(v) = F_0$ *for* $v \in [a, r^*]$, $F^*(v) = 1 - c/v$ *for* $v \in [r^*, b)$ *and* $F^*(b) = 1$.

Proof of Theorem 7. We will prove the 3 cases separately.

Low Information ($a/b \leq k_l$) Regime. We know from Theorem 1 that SPA(Φ^*) is minimax optimal over $\mathcal{M}_{\text{all}}$ and so it is also minimax optimal over $\mathcal{M}_{\text{SPA-rand}}$.

High Information ($a/b \geq \frac{\lambda n}{(1+\lambda)n-1}$) Regime. We claim that the minimax optimal mechanism Φ^* in this regime is SPA without reserve and F^* is a two-point distribution with mass $f_a := \frac{n-1}{n-1+\lambda}$ at a and $f_b := \frac{\lambda}{n-1+\lambda}$ at b. Here, $R(\Phi^*, F^*) = -(1-\lambda)b + \left(\frac{n-1}{n-1+\lambda}\right)^{n-1}(b-a)$.

We want to show that $R(\Phi^*, F) \leq R(\Phi^*, F^*) \leq R(\Phi, F^*)$.

<u>Part 1: $R(\Phi, F^*) \geq R(\Phi^*, F^*)$</u>

We have the regret expression $R(\Phi, F)$ from Proposition 10:

$$R(\Phi, F) = \lambda b - a\Phi(a)F(a)^n - (1-(1-f_b)^n)b\Phi(b) + \int_{v=a}^{v=b} -\lambda F(v)^n + \Phi(v)nF(v)^{n-1}(1 - F(v) - vF'(v))dv$$

Under F^*, we have $F^*(v) = F(a) = f_a \frac{n-1}{n-1+\lambda}$ for all $v \in (a, b)$, which means every appearance of $F(v)$ becomes a constant:

$$R(\Phi, F^*) = \lambda b - af_a^n\Phi(a) - (1-f_a^n)b\Phi(b) + \int_{v=a}^{v=b} -\lambda f_a^n + nf_a^{n-1}(1-f_a)\Phi(v)dv$$

We have $\Phi(b) \le 1$ and $-\lambda f_a^n + n f_a^{n-1}(1-f_a)\Phi(v) \ge -\lambda f_a^n + n f_a^{n-1}(1-f_a)\Phi(a)$, so

$$R(\Phi, F^*) \ge \lambda b - a\Phi(a) f_a^n - (1-f_a^n)b + (b-a)(-\lambda f_a^n + n f_a^{n-1}(1-f_a)\Phi(a))$$
$$= -(1-\lambda)b + (n f_a^{n-1} - (n-1+\lambda)f_a^n)(b-a) + (1-\Phi(a))f_a^{n-1}(a f_a - (b-a)(n-nf_a))$$

The last term is ≥ 0 because we require that $af_a - (b-a)(n-nf_a) \ge 0 \Leftrightarrow \frac{a}{b} \ge \frac{n-nf_a}{n-(n-1)f_a} = \frac{\lambda n}{(1+\lambda)n-1}$. We now see that the choice $f_a = \frac{n-1}{n-1+\lambda}$ is chosen so the second term is maximized, and the bound on $\frac{a}{b}$ that is required to make the third term work follows accordingly. Therefore,

$$R(\Phi, F^*) \ge -(1-\lambda)b + (n f_a^{n-1} - (n-1+\lambda)f_a^n)(b-a) = -(1-\lambda)b + \left(\frac{n-1}{n-1+\lambda}\right)^{n-1}(b-a)$$

Part 2: $R(\Phi^*, F) \le R(\Phi^*, F^*)$

Because Φ^* is an SPA without reserve, we have $\Phi^*(v) = 1$ for every v. This gives

$$\begin{aligned} R(\Phi^*, F) &= a(\lambda-1) + \int_{v=a}^{v=b} (\lambda-1)(1-F(v)^n) + (nF(v)^{n-1} - nF(v)^n)dv \\ &= -(1-\lambda)b + \int_{v=a}^{v=b} nF(v)^{n-1} - (n-1+\lambda)F(v)^n dv \\ &\le -(1-\lambda)b + (b-a)\sup_{z\in[0,1]} nz^{n-1} - (n-1+\lambda)z^n \\ &= -(1-\lambda)b + \left(\frac{n-1}{n-1+\lambda}\right)^{n-1}(b-a) = R(\Phi^*, F^*) \end{aligned}$$

Moderate Information ($k_l \le a/b \le \frac{\lambda n}{(1+\lambda)n-1}$) Regime. We will exhibit a saddle point with the following structure: Φ^* has a point mass $\Phi^*(a)$ at a, then it is flat on $[a, r^*]$ (that is, $\Phi^*(v) = \Phi^*(a)$ on $[a, r^*]$), then it has a density on $[r^*, b]$, but no point mass at b (so $\Phi^*(b) = 1$). F^* has a point mass $F^*(a)$ at a, then it is flat on $[a, r^*]$, then it is $F^*(v) = 1 - c/v$ on $[r^*, b)$. That is, $F^*(v) = \max(1-c/v, F^*(a))$ on $v \in [a, b)$. Importantly, we assume that F^* is continuous at r^*, so $1 - c/r^* = F^*(a)$, and this r^* is the same as r^* of Φ^*.

We use the regret expression Regret-Φ under F^*. The general Regret-Φ expression is

$$R_n(\boldsymbol{\Phi}, \boldsymbol{F}) = \lambda b - F(a)^n a\Phi(a) - (1-(1-f_b)^n)b\Phi(b) + \int_{v=a}^{v=b} -\lambda F(v)^n + \Phi(v)nF(v)^{n-1}(1-F(v) - vF'(v))dv$$

Note that $F^*(b^-) = 1 - f_b = 1 - c/b$. We will write $F_0 := F^*(a)$ for convenience.

We can then write

$$R_n(\Phi, \boldsymbol{F}^*) = \lambda b - F_0^n a\Phi(a) - (1-(1-c/b)^n)b\Phi(b) + \int_{v=a}^{v=r^*} -\lambda F_0^n + \Phi(v)nF_0^{n-1}(1-F_0)dv$$
$$+ \int_{v=r^*}^{v=b} -\lambda \left(1-\frac{c}{v}\right)^n dv$$

Consider the problem

$$R^* := \min_{\Phi} R_n(\Phi, \boldsymbol{F}^*) = \text{Regret}(\Phi) \text{ s.t. } \Phi(v) \in [0,1] \text{ non-decreasing}$$

We dualize the non-decreasing constraint.

Let

$$\mathcal{L}(\Phi, \mu) = \text{Regret}(\Phi) - \int_a^b \mu(v)d\Phi(v)$$

and

$$q(\mu) = \min_{\Phi(v)\in[0,1]} \mathcal{L}(\Phi, \mu) \text{ with } \mu(v) \geq 0$$

Weak duality says that $R^* \geq q(\mu)$ for all $\mu : [a,b] \to \mathbb{R}_+$.

To get to strong duality, we want to choose a specific μ^* such that:

- Complementary Slackness (CS): $\int_a^b \mu^*(v)d\Phi^*(v) = 0$. That is, wherever Φ^* is strictly increasing, μ^* is zero.
- Lagrangian Optimality (LO): $\Phi^* \in \arg\min_{\Phi\in\text{CDF}} \mathcal{L}(\Phi, \mu^*)$

Because the condition that $\Phi(v)$ is non-decreasing doesn't bind (it is strictly increasing) on $[r^*, b]$, by complementary slackness $\mu^*(v) = 0$ on $[r^*, b]$.

We can then use the integration by parts to get (using $\mu^*(r^*) = 0$)

$$\int_{v=a}^{v=b} \mu^*(v)d\Phi(v) = \int_{v=a}^{v=r^*} \mu^*(v)d\Phi(v) = \mu^*(r^*)\Phi(r^*) - \mu^*(a)\Phi(a) - \int_{v=a}^{v=r^*} (\mu^*)'(v)\Phi(v)dv$$
$$= -\mu^*(a)\Phi(a) - \int_{v=a}^{v=r^*} (\mu^*)'(v)\Phi(v)dv$$

We substitute this into the expression for $\mathcal{L}(\Phi, \mu^*)$ to get

$$\mathcal{L}(\Phi, \mu^*) = \lambda b - \lambda(r^* - a)F_0^n + (\mu^*(a) - F_0^n a)\Phi(a) - \left(1 - \left(1 - \frac{c}{b}\right)^n\right) b\Phi(b) - \lambda \int_{v=r^*}^{v=b} \left(1 - \frac{c}{v}\right)^n dv$$
$$+ \int_{v=a}^{v=r^*} \left[(\mu^*)'(v) + nF_0^{n-1}(1 - F_0)\right] \Phi(v)dv$$

We want to choose μ^* such that

$$\mu^*(a) - F_0^n a = 0$$
$$(\mu^*)'(v) + nF_0^{n-1}(1 - F_0) = 0$$

so that the above expression for $\mathcal{L}(\Phi, \mu^*)$ becomes independent of Φ (zero out the coefficient of $\Phi(a)$ and $\Phi(v)$ between a and r^*; there still is $\Phi(b)$ but we will let this be 1). From

$$\mu^*(r^*) = \mu^*(a) + \int_{v=a}^{v=r^*} (\mu^*)'(v)dv$$
$$0 = aF_0^n - (r^* - a)nF_0^{n-1}(1 - F_0)$$
$$aF_0 = n(r^* - a)(1 - F_0)$$

With that μ^*, we have

$$\mathcal{L}(\Phi, \mu^*) = \lambda b - \lambda(r^* - a)F_0^n - \left(1 - \left(1 - \frac{c}{b}\right)^n\right) b\Phi(b) - \lambda \int_{v=r^*}^{v=b} \left(1 - \frac{c}{v}\right)^n dv$$

If we further assume that $\Phi(b) = 1$, then

$$\mathcal{L}(\Phi, \mu^*) = -(1 - \lambda)b - \lambda(r^* - a)F_0^n + b\left(1 - \frac{c}{b}\right)^n - \lambda \int_{v=r^*}^{v=b} \left(1 - \frac{c}{v}\right)^n dv$$

Now we derive conditions from the fact that F maximizes regret given fixed Φ^*, that is, the saddle $R(\Phi^*, F) \leq R(\Phi^*, F^*)$. We use the Regret-$F$ equation

$$R(\Phi, F) = a(\lambda - \Phi(a)) + \int_{v=a}^{v=b} \left[(\lambda - \Phi(v) - v\Phi'(v))(1 - F(v)^n) + \Phi(v)nF(v)^{n-1}(1 - F(v))\right] dv$$

We do pointwise optimization for each v. For $v \in (a, r^*)$. $\Phi^*(v) = \Phi^*(a)$ is a constant, so $F(v)$ that maximizes that is a constant, the same for every v, given by

$$F^*(v) \in \arg\max_z (\lambda - \Phi^*(a))(1 - z^n) + \Phi^*(a) n z^{n-1}(1 - z)$$

Taking the derivative of z gives

$$\begin{aligned} -(\lambda - \Phi^*(a)) n z^{n-1} + n\Phi^*(a)((n-1)z^{n-2} - nz^{n-1}) &= 0 \\ -(\lambda - \Phi^*(a))z + \Phi^*(a)(n - 1 - nz) &= 0 \end{aligned}$$

Write $\Phi^*(a) = \Phi_0$ for convenience. By the first-order condition, $z = F^*(a) = F_0$ satisfies this equation, so

$$-(\lambda - \Phi_0)F_0 + \Phi_0(n - 1 - nF_0) = 0$$

For $v \in (r^*, b)$, $\Phi^*(v)$ is no longer a constant (but this is the regime that we have dealt with before). We have

$$F^*(v) \in \arg\max_z (\lambda - \Phi(v) - v\Phi'(v))(1 - z^n) + \Phi(v) n z^{n-1}(1 - z)$$

The first order condition gives

$$\begin{aligned} -(\lambda - \Phi(v) - v\Phi'(v)) n z^{n-1} + \Phi(v) n((n-1)z^{n-2} - nz^{n-1}) &= 0 \\ -(\lambda - \Phi(v) - v\Phi'(v))z + \Phi(v)((n - 1) - nz) &= 0 \end{aligned}$$

By the first order condition, $z = F^*(v) = 1 - c/v$ satisfies this equation, so

$$-(\lambda - (\Phi^*)(v) - v(\Phi^*)'(v))\left(1 - \frac{c}{v}\right) + \Phi(v)\left((n-1) - n\left(1 - \frac{c}{v}\right)\right) = 0$$

which simplifies to

$$(\Phi^*)'(v) = \frac{\lambda}{v} - \frac{(n-1)c}{v(v-c)}\Phi^*(v)$$

We have seen this ODE before. The solution is

$$\left(\frac{v-c}{v}\right)^{n-1}\Phi^*(v) = \lambda\left[d + \log\left(\frac{v}{c}\right) - \sum_{k=1}^{n-1}\frac{1}{k}\left(\frac{v-c}{v}\right)^k\right]$$

for some constant d.

$\Phi^*(r^*) = \Phi^*(a) = \Phi_0$ is the point mass of Φ^* at a, which is unknown. We get

$$\left(\frac{r^*-c}{r^*}\right)^{n-1}\Phi_0 = \lambda\left[d + \log\left(\frac{r^*}{c}\right) - \sum_{k=1}^{n-1}\frac{1}{k}\left(\frac{r^*-c}{r^*}\right)^k\right]$$

From $1 - c/r^* = F_0 \geq 0$ we have $c \leq r^*$ with strict inequality if $F_0 > 0$. So the summation makes sense (and also tells us it doesn't necessarily go away as zero like before). With $\Phi^*(b) = 1$ we get

$$\left(\frac{b-c}{b}\right)^{n-1} = \lambda\left[d + \log\left(\frac{b}{c}\right) - \sum_{k=1}^{n-1}\frac{1}{k}\left(\frac{b-c}{b}\right)^k\right]$$

Therefore, we have 5 equations for 5 unknowns Φ_0, F_0, c, d, r^*:

$$aF_0 = n(r^*-a)(1-F_0) \tag{9}$$

$$-(\lambda-\Phi_0)F_0 + \Phi_0(n-1-nF_0) = 0 \tag{10}$$

$$\left(\frac{r^*-c}{r^*}\right)^{n-1}\Phi_0 = \lambda\left[d + \log\left(\frac{r^*}{c}\right) - \sum_{k=1}^{n-1}\frac{1}{k}\left(\frac{r^*-c}{r^*}\right)^k\right] \tag{11}$$

$$\left(\frac{b-c}{b}\right)^{n-1} = \lambda\left[d + \log\left(\frac{b}{c}\right) - \sum_{k=1}^{n-1}\frac{1}{k}\left(\frac{b-c}{b}\right)^k\right] \tag{12}$$

$$1 - \frac{c}{r^*} = F_0 \tag{13}$$

For notational convenience, we will write $F^*(a)$ as F_0.

We will write every variable in terms of F_0, so we have a single-variable equation we can solve.

From (9), we get

$$r^* = a\left(\frac{F_0}{n(1-F_0)} + 1\right) = \frac{n-(n-1)F_0}{n(1-F_0)}a \tag{14}$$

From (10) we get

$$\Phi_0 = \frac{\lambda F_0}{(n-1)(1-F_0)} \tag{15}$$

From (13) we get

$$c = r^*(1-F_0) = \frac{n-(n-1)F_0}{n} a \tag{16}$$

Subtracting (11) and (12) gives

$$\left(\frac{r^*-c}{r^*}\right)^{n-1}\Phi_0 - \left(\frac{b-c}{b}\right)^{n-1} = \lambda\left[\log\left(\frac{r^*}{b}\right) - \sum_{k=1}^{n-1}\frac{1}{k}\left(\frac{r^*-c}{r^*}\right)^k + \sum_{k=1}^{n-1}\frac{1}{k}\left(\frac{b-c}{b}\right)^k\right]$$

Substituting the expressions of r^*, Φ_0, c in terms of F_0 from (14), (15), (16) gives (writing $\tilde{a} = a/b$)

$$\begin{aligned} &\frac{\lambda F_0^n}{(n-1)(1-F_0)} - \left(1 - \frac{n-(n-1)F_0}{n}\tilde{a}\right)^{n-1} \\ &= \lambda\left[\log\left(\frac{n-(n-1)F_0}{n(1-F_0)}\tilde{a}\right) - \sum_{k=1}^{n-1}\frac{1}{k}F_0^k + \sum_{k=1}^{n-1}\frac{1}{k}\left(1-\frac{n-(n-1)F_0}{n}\tilde{a}\right)^k\right] \end{aligned}$$

Let $\mathrm{fn}(F_0)$ be the left hand side minus the right hand side (taking $\tilde{a} = a/b$ as fixed):

$$\begin{aligned} \mathrm{fn}(F_0) = \frac{\lambda F_0^n}{(n-1)(1-F_0)} - \left(1-\frac{n-(n-1)F_0}{n}\tilde{a}\right)^{n-1} - \lambda\log\left(\frac{n-(n-1)F_0}{n(1-F_0)}\tilde{a}\right) \\ + \lambda\sum_{k=1}^{n-1}\frac{1}{k}F_0^k - \lambda\sum_{k=1}^{n-1}\frac{1}{k}\left(1-\frac{n-(n-1)F_0}{n}\tilde{a}\right)^k \end{aligned}$$

(When $\tilde{a}$ is not fixed, we will write the above expression instead as $L(F_0, \tilde{a})$, and we will use this notation later in the proof.)

We first note that, because $\tilde{a} \geq k_l$, by definition of k_l we have

$$\mathrm{fn}(0) = -(1-\tilde{a})^{n-1} - \lambda\log(\tilde{a}) - \lambda\sum_{k=1}^{n-1}\frac{1}{k}(1-\tilde{a})^k \leq 0$$

with equality only when $\tilde{a} = k_l$.

We also note that as $F_0 \uparrow 1$, LHS grows as $\frac{1}{1-F_0}$ whereas RHS grows as $\log\left(\frac{1}{1-F_0}\right)$, so $\lim_{F_0\uparrow 1} \mathrm{fn}(F_0) = +\infty$.

Lastly, we compute the derivative of fn as

$$
\begin{aligned}
\mathrm{fn}'(F_0) &= \frac{\lambda F_0^{n-1}(n-(n-1)F_0)}{(n-1)(1-F_0)^2} - (n-1)\left(1-\frac{n-(n-1)F_0}{n}\tilde{a}\right)^{n-2}\cdot\frac{(n-1)}{n}\tilde{a} \\
&+ \lambda\left[\frac{(n-1)}{n-(n-1)F_0} - \frac{1}{1-F_0} + \sum_{k=1}^{n-1} F_0^{k-1} - \sum_{k=1}^{n-1}\left(1-\frac{n-(n-1)F_0}{n}\tilde{a}\right)^{k-1}\cdot\frac{(n-1)}{n}\tilde{a}\right] \\
&= \frac{\lambda F_0^{n-1}(n-(n-1)F_0)}{(n-1)(1-F_0)^2} - (n-1)\left(1-\frac{n-(n-1)F_0}{n}\tilde{a}\right)^{n-2}\cdot\frac{(n-1)}{n}\tilde{a} \\
&+ \lambda\left[\frac{(n-1)}{n-(n-1)F_0} - \frac{1}{1-F_0} + \frac{1-F_0^{n-1}}{1-F_0} - \frac{(n-1)}{n-(n-1)F_0}\left[1-\left(1-\frac{n-(n-1)F_0}{n}\tilde{a}\right)^{n-1}\right]\right] \\
&= \frac{\lambda F_0^{n-1}(n-(n-1)F_0)}{(n-1)(1-F_0)^2} - \frac{(n-1)^2}{n}\tilde{a}\left(1-\frac{n-(n-1)F_0}{n}\tilde{a}\right)^{n-2} \\
&+ \lambda\left[-\frac{F_0^{n-1}}{1-F_0} + \frac{(n-1)}{n-(n-1)F_0}\left(1-\frac{n-(n-1)F_0}{n}\tilde{a}\right)^{n-1}\right] \\
&= F_0^{n-1}\lambda\left(\frac{n-(n-1)F_0}{(n-1)(1-F_0)^2} - \frac{1}{1-F_0}\right) \\
&+ \left(1-\frac{n-(n-1)F_0}{n}\tilde{a}\right)^{n-2}\left[-\frac{(n-1)^2}{n}\tilde{a} + \frac{\lambda(n-1)}{n-(n-1)F_0}\left(1-\frac{n-(n-1)F_0}{n}\tilde{a}\right)\right] \\
&= \frac{\lambda F_0^{n-1}}{(n-1)(1-F_0)^2} + \left(1-\frac{n-(n-1)F_0}{n}\tilde{a}\right)^{n-2}(n-1)\left(\frac{\lambda}{n-(n-1)F_0} - \frac{n-1+\lambda}{n}\tilde{a}\right)
\end{aligned}
$$

We can see from the expression that $\mathrm{fn}'(F_0)$ is increasing in F_0 , so either fn is increasing for all F_0 in range (if $\mathrm{fn}'(0) \geq 0$, or it is decreasing for $F_0 \leq F_0^*$ for some constant F_0^* and increasing for $F_0 \geq F_0^*$. Because, for $a/b > k_l$, $\mathrm{fn}(0) < 0$ and $\lim_{F_0\uparrow 1} \mathrm{fn}(F_0) = +\infty$, we conclude that the equation $\mathrm{fn}(F_0) = 0$ has a unique solution $F_0 \in (0, 1)$.

We will also need to show that this unique solution F_0 leads to feasible values of other parameters as well. The ones that concern us are $\Phi^*(a)$ and r^*.

We must have

$$\Phi^*(a) = \frac{\lambda F_0}{(n-1)(1-F_0)} \leq 1 \Leftrightarrow F_0 \leq \frac{n-1}{n-1+\lambda}$$

and

$$r^* = \frac{n-(n-1)F_0}{n(1-F_0)} a \leq b \Leftrightarrow F_0 \leq \frac{n(1-\tilde{a})}{n-(n-1)\tilde{a}}$$

We will show that $F_0 \leq \frac{n-1}{n}$.

We note that $F_0(\tilde{a} = \frac{\lambda n}{(1+\lambda)n-1}) = \frac{n-1}{n-1+\lambda}$ because when we plug in $\tilde{a} = \frac{\lambda n}{(1+\lambda)n-1}, F_0 = \frac{n-1}{n-1+\lambda}$ in the $L(F_0, \tilde{a})$ expression we get zero. We claim that $L(F_0, \tilde{a} = \frac{\lambda n}{(1+\lambda)n-1})$ is increasing in F_0 for $\frac{n-1}{n-1+\lambda} \leq F_0 \leq 1$. If this is true then we are done, because if for some $\tilde{a} < \frac{\lambda n}{(1+\lambda)n-1}$ we have $F_0(\tilde{a}) > \frac{n-1}{n-1+\lambda}$, then $0 = L(F_0(\tilde{a}), \tilde{a}) > L(F_0(\tilde{a}), \frac{\lambda n}{(1+\lambda)n-1}) > L(\frac{n-1}{n-1+\lambda}, \frac{\lambda n}{(1+\lambda)n-1}) = 0$, a contradiction.

To show that $L(F_0, \frac{\lambda n}{(1+\lambda)n-1})$ is increasing in F_0 for $1 \geq F_0 \geq \frac{n-1}{n-1+\lambda}$, we calculate (taking the expression from the $\text{fn}'(F_0)$ earlier)

$$\frac{\partial L(F_0, \tilde{a} = \frac{\lambda n}{(1+\lambda)n-1})}{F_0} =$$

$$= \frac{\lambda F_0^{n-1}}{(n-1)(1-F_0)^2} + \left(1 - \frac{\lambda(n-(n-1)F_0)}{(1+\lambda)n-1}\right)^{n-2} \cdot (n-1)\lambda \left(\frac{1}{n-(n-1)F_0} - \frac{n-1+\lambda}{(1+\lambda)n-1}\right) \geq 0$$

This is true because

$$\frac{1}{n-(n-1)F_0} \geq \frac{1}{n-(n-1)\frac{\lambda n}{(1+\lambda)n-1}} = \frac{n-1+\lambda}{(1+\lambda)n-1}$$

and

$$1 - \frac{\lambda(n-(n-1)F_0)}{(1+\lambda)n-1} > 0$$

so the derivative is positive, as desired. (Note also that the expression $\frac{\partial L(F_0, \tilde{a} = \frac{\lambda n}{(1+\lambda)n-1})}{F_0}$ is increasing in F_0 so we can also just plug in $F_0 = \frac{n-1}{n-1+\lambda}$ in that expression and check that the resulting expression is > 0.

Lastly, we want to show that $F_0 \leq \frac{n(1-\tilde{a})}{n-(n-1)\tilde{a}}$. Note that $\tilde{a} < \frac{\lambda n}{(1+\lambda)n-1}$ implies that $\frac{n-1}{n-1+\lambda} < \frac{n(1-\tilde{a})}{n-(n-1)\tilde{a}}$, so this inequality is immediately implied by $F_0 \leq \frac{n-1}{n-1+\lambda}$ which we just proved.

We conclude that for $k_l \leq \tilde{a} \leq \frac{\lambda n}{(1+\lambda)n-1}$, these parameters give rise to a feasible mechanism that is minimax optimal, as desired.

$\square$

Proposition 11. *The worst-case regret of* SPA(r)*, the second-price auction with fixed reserve r, is*

$$R(\text{SPA}(r), \mathcal{F}_{\text{iid}}) = \begin{cases} -(1-\lambda)b + (b-a)\left(\frac{n-1}{n-1+\lambda}\right)^{n-1} & \textit{if } r = a \\ -(1-\lambda)b + \frac{(b-r)^n (n-1)^{n-1}}{((n-1+\lambda)(b-r)-r)^{n-1}} & \textit{if } a < r \leq \frac{\lambda}{1+\lambda} b \\ \lambda r & \textit{if } \frac{\lambda}{1+\lambda} b \leq r \leq b \end{cases}$$

Therefore,

$$\inf_{r \in [a,b]} R(\text{SPA}(r), \mathcal{F}_{\text{iid}}) = \begin{cases} -(1-\lambda)b + \left(\frac{n}{n+\lambda}\right)^n b & \textit{if } \frac{a}{b} \leq 1 - \left(\frac{n}{n+\lambda}\right)^n \left(\frac{n-1+\lambda}{n-1}\right)^{n-1} \\ -(1-\lambda)b + \left(\frac{n-1}{n-1+\lambda}\right)^{n-1} (b-a) & \textit{if } \frac{a}{b} \geq 1 - \left(\frac{n}{n+\lambda}\right)^n \left(\frac{n-1+\lambda}{n-1}\right)^{n-1} \end{cases}$$

The optimal r in the first case ("low" a/b) is $r = \frac{\lambda}{n+\lambda} b$ and the optimal r in the second case ("high" a/b) is $r = a$.

This result is valid for any $n \geq 1$, if for $n = 1$ we interpret any term with the $n-1$ exponent as 1. In agreement with results from $\mathcal{M}_{\text{SPA-rand}}$ and our intuition, when scale information is important (a/b is high), the regret-minimizing reserve is no reserve $r = a$.

Before we prove the main result (Proposition 11) characterizing the minimax λ-regret SPA with fixed reserve (including no reserve), we first derive the regret expression of SPA(r) for a fixed r.

Proposition 12. *Fix $r \in [a, b]$. The regret of* SPA(r) *against a joint distribution $\boldsymbol{F}$ is*

$$-(1-\lambda)b + r\boldsymbol{F}_n^{(1)}(r^-) - \lambda \int_{v=a}^{v=r} \boldsymbol{F}_n^{(1)}(v)dv + \int_{v=r}^{v=b} (\boldsymbol{F}_n^{(2)}(v) - \lambda \boldsymbol{F}_n^{(1)}(v))dv$$

where $\boldsymbol{F}_n^{(1)}(r^-) = \Pr(v^{(1)} < v)$. If we further assume that $\boldsymbol{F}$ is n i.i.d. with marginal F, then the regret is

$$-(1-\lambda)b + rF(r^-)^n - \lambda \int_{v=a}^{v=r} F(v)^n dv + \int_{v=r}^{v=b} (nF(v)^{n-1} - (n-1+\lambda)F(v)^n)dv$$

Proof. The regret is

$$\begin{aligned}
&\lambda\mathbb{E}[v^{(1)}] - \mathbb{E}[\max(v^{(2)}, r)\mathbf{1}(v^{(1)} \geq r)] \\
&= \lambda\mathbb{E}[v^{(1)}] - \mathbb{E}[v^{(2)}\mathbf{1}(v^{(2)} > r)] - \mathbb{E}[r\mathbf{1}(v^{(2)} \leq r \leq v^{(1)})].
\end{aligned}$$

The first term is

$$\mathbb{E}[v^{(1)}] = \int_{v \geq 0} \Pr(v^{(1)} > v)dv = b - \int_{v=a}^{v=b} \boldsymbol{F}_n^{(1)}(v)dv.$$

The second term's calculation is analogous to that of Lemma 3. We have

$$\begin{aligned}
\mathbb{E}[v^{(2)}\mathbf{1}(v^{(2)} > r)] &= \int_{v' \in (r,b]} v' d\boldsymbol{F}_n^{(2)}(v') = \int_{v' \in (r,b]} \left(a + \int_{v \in [a,v')} dv\right) d\boldsymbol{F}_n^{(2)}(v') \\
&= a(\boldsymbol{F}_n^{(2)}(b) - \boldsymbol{F}_n^{(2)}(r)) + \int_{v \in [a,r]} \int_{v' \in (r,b]} d\boldsymbol{F}_n^{(2)}(v')dv + \int_{v \in (r,b]} \int_{v' \in (v,b]} d\boldsymbol{F}_n^{(2)}(v')dv \\
&= a(1 - \boldsymbol{F}_n^{(2)}(r)) + \int_{v \in [a,r]} (1 - \boldsymbol{F}_n^{(2)}(r))dv + \int_{v \in (r,b]} (1 - \boldsymbol{F}_n^{(2)}(v))dv \\
&= b - r\boldsymbol{F}_n^{(2)}(r) - \int_{v=r}^{v=b} \boldsymbol{F}_n^{(2)}(v)dv.
\end{aligned}$$

The third term is

$$\mathbb{E}[r\mathbf{1}(v^{(2)} \leq r \leq v^{(1)})] = r\Pr(v^{(2)} \leq r \leq v^{(1)}) = r(\Pr(v^{(2)} \leq r) - \Pr(v^{(1)} < r)) = r(\boldsymbol{F}_n^{(2)}(r) - \boldsymbol{F}_n^{(1)}(r^-))$$

Together, we have, for $r \in [a, b]$,

$$R_n(\mathrm{SPA}(r), \boldsymbol{F}_n) = -(1-\lambda)b + r\boldsymbol{F}_n^{(1)}(r^-) - \lambda \int_{v=a}^{v=r} \boldsymbol{F}_n^{(1)}(v)dv + \int_{v=r}^{v=b} (\boldsymbol{F}_n^{(2)}(v) - \lambda\boldsymbol{F}_n^{(1)}(v))dv$$

as desired.

When $\boldsymbol{F}$ is n i.i.d. F, we have $\boldsymbol{F}_n^{(1)}(r^-) = \Pr(\max(\boldsymbol{v}) < r) = \prod_{i=1}^n \Pr(v_i < r) = F(r^-)^n$ where the second-to-last equality uses the fact that the v_i's are independent. We also have $\boldsymbol{F}_n^{(1)}(v) = F(v)^n$ and $\boldsymbol{F}_n^{(2)}(v) = nF(v)^{n-1} - (n-1)F(v)^n$. $\square$

Now we are ready to prove the main result.

Proof of Proposition 11. From Proposition 12, the λ-regret of SPA(r) is

$$R_n(\mathrm{SPA}(r), F) = -(1-\lambda)b + rF_{-}(r)^n - \lambda \int_{v=a}^{v=r} F(v)^n dv + \int_{v=r}^{v=b} nF(v)^{n-1} - (n-1+\lambda)F(v)^n dv.$$

We first assume that $n \geq 2$ and $r \in (a, b]$. Let $c = F_{-}(r)$. (Note here that we require $r > a$ in order for us to have the freedom to set the value of $c = \Pr(v < r)$, the mass *strictly* below r. If $r = a$, i.e. there is no reserve, then $c = 0$ by definition. This is why we consider the case $r = a$, i.e. no reserve, separately.) Note that the integrand $nF(v)^{n-1} - (n-1+\lambda)F(v)^n$ is increasing for $F(v) \leq \frac{n-1}{n-1+\lambda}$ and is decreasing for $F(v) \geq \frac{n-1}{n-1+\lambda}$. To minimize $\int_{v \in [a,r]^n} F(v)^n dv$ we must have $F(v) = 0$ for $v \in [a, r-\epsilon]$ for arbitrarily small $\epsilon > 0$, and to maximize $\int_{v \in (r,b]} nF(v)^{n-1} - nF(v)^n dv$, the only constraint we have is $F(v) \geq c$ so for $v \in (r, b]$ we set $F(v) = \frac{n-1}{n-1+\lambda}$ if $c \leq \frac{n-1}{n-1+\lambda}$ and $F(v) = c$ otherwise. Note that the sup over first case of $c \leq \frac{n-1}{n-1+\lambda}$ is simply the second case with $c = \frac{n-1}{n-1+\lambda}$. Because we take the sup over F, we can let $\epsilon \downarrow 0$ and get that the worst-case regret is

$$-(1-\lambda)b + \sup_{c \in [\frac{n-1}{n-1+\lambda}, 1]} rc^n + (b-r)(nc^{n-1} - (n-1+\lambda)c^n)$$

Now, the derivative of this expression of c is $nc^{n-2}\left[rc + (b-r)(n-1-(n-1+\lambda)c)\right]$. The expression in $[\cdots]$ is linear in c. At $c = \frac{n-1}{n-1+\lambda}$, the derivative expression is $nr\left(\frac{n-1}{n-1+\lambda}\right)^{n-1} \geq 0$. At $c = 1$, the expression is $n((1+\lambda)r - \lambda b)$. So if $r \geq \frac{\lambda}{1+\lambda}b$, the first derivative is always ≥ 0, so the maximum is achieved at $c = 1$ and the value is λr. If $r \leq \frac{\lambda}{1+\lambda}b$, the maximum is achieved at $c^* = \frac{(n-1)(b-r)}{(n-1+\lambda)(b-r)-r} \in \left[\frac{n-1}{n-1+\lambda}, 1\right]$ and the value is $-(1-\lambda)b + \frac{(b-r)^n(n-1)^{n-1}}{((n-1+\lambda)(b-r)-r)^{n-1}}$.

Now we consider the case $r = a$. In this case, by definition $c = 0$ and we have

$$R_n(\mathrm{SPA}(a), F) = -(1-\lambda)b + \int_{v \in (a,b]} nF(v)^{n-1} - (n-1+\lambda)F(v)^n dv$$

so

$$R_n(\mathrm{SPA}(a), \mathcal{F}) = -(1-\lambda)b + (b-a) \sup_{z \in [0,1]} nz^{n-1} - (n-1+\lambda)z^n$$

The maximum occurs at $z = \frac{n-1}{n-1+\lambda}$ which gives

$$R_n(\mathrm{SPA}(a), \mathcal{F}) = -(1-\lambda)b + (b-a)\left(\frac{n-1}{n-1+\lambda}\right)^{n-1}$$

Now we deal with the case $n = 1$. The regret expression reduces to

$$R_1(\mathrm{SPA}(r), F) = -(1-\lambda)b + rF_{-}(r) - \lambda \int_{v\in[a,r]} F(v)dv + \int_{v\in(r,b]} (1-\lambda F(v))dv.$$

For $r > a$, we have

$$R_1(\mathrm{SPA}(r), \mathcal{F}) = -(1-\lambda)b + \sup_{c\in[0,1]} rc + (b-r)(1-\lambda c) = \max(\lambda b - r, \lambda r) = \begin{cases} \lambda b - r & \text{if } a < r \leq \frac{\lambda}{1+\lambda} b \\ \lambda r & \text{if } r \geq \frac{\lambda}{1+\lambda} b \end{cases}$$

because the expression under sup is linear in c so it achieves the extrema at one of the end points, either at $c = 0$ or $c = 1$.

For $r = a$ we have

$$R_1(\mathrm{SPA}(a), F) = -(1-\lambda)b + \int_{v\in(a,b]} (1-\lambda F(v))dv.$$

This is maximized when $F(v) = 0$ for all $v \in (a, b]$ and we get

$$R_1(\mathrm{SPA}(a), \mathcal{F}) = -(1-\lambda)b + (b-a) = \lambda b - a$$

We therefore have

$$R_1(\mathrm{SPA}(r), \mathcal{F}) = \begin{cases} \lambda b - a & \text{if } r = a \\ \lambda b - r & \text{if } a < r \leq \frac{\lambda}{1+\lambda} b \\ \lambda r & \text{if } r \geq \frac{\lambda}{1+\lambda} b \end{cases}$$

Note that the second regime and the first regime are continuous whenever the second regime is applicable, but we will keep them separate for clarity (because the first regime $r = a$ is always applicable, whereas the second regime $r \in (a, \frac{\lambda}{1+\lambda} b]$ is applicable only when $\frac{a}{b} < \frac{\lambda}{1+\lambda}$.

Now we want to choose the optimal r to minimize the worst-case regret. First consider the case $n \geq 2$. We note that

$$(b-a)\left(\frac{n-1}{n-1+\lambda}\right)^{n-1} \leq \frac{(b-a)^n(n-1)^{n-1}}{((n-1+\lambda)(b-a)-a)^{n-1}}$$

with equality if and only if $a = 0$. Therefore, if $\frac{a}{b} < \frac{\lambda}{1+\lambda}$, that is, the regime $r \in (a, \frac{\lambda}{1+\lambda}b]$ is permissible, then the worst-case regret under $r = a$ is lower than under $r = a^+$, slightly above a. In contrast, the worst-case regret is continuous at $r = \frac{\lambda}{1+\lambda}b$. Given that the regret in the third regime λr is linear in r, the worst r (lowest regret) occurs at $r = \frac{\lambda}{1+\lambda}b$ with regret $\frac{\lambda^2}{1+\lambda}b$.

First consider the case $\frac{a}{b} < \frac{\lambda}{1+\lambda}$, so all 3 regimes of r are permissible.

In the $r \in [\frac{\lambda}{1+\lambda}b, b]$ regime, the regret is λr, so the lowest regret occurs at $r = \frac{\lambda}{1+\lambda}b$ and has value $\frac{\lambda^2}{1+\lambda}b$.

In the $r \in (a, \frac{\lambda}{1+\lambda}b]$ regime, the regret is

$$-(1-\lambda)b + (n-1)^{n-1}\exp\left\{n\log(b-r) - (n-1)\log((n-1+\lambda)b - (n+\lambda)r)\right\}$$

The derivative of the expression in $\{\cdots\}$ is

$$-\frac{n}{b-r} + \frac{(n-1)(n+\lambda)}{(n-1+\lambda)b-(n+\lambda)r} = \frac{(n+\lambda)r - \lambda b}{(b-r)((n-1+\lambda)b-(n+\lambda)r)}$$

Therefore, in this second regime, the worst-case regret is decreasing for $\frac{a}{b} \leq \frac{\lambda}{n+\lambda}$ and increasing for $\frac{a}{b} \geq \frac{\lambda}{n+\lambda}$. So if $\frac{a}{b} \leq \frac{\lambda}{n+\lambda}$, the r that minimizes worst-case regret is $r = \frac{\lambda}{n+\lambda}b$, which gives the regret

$$-(1-\lambda)b + \left(\frac{n}{n+\lambda}\right)^n b$$

Therefore, the overall worst-case regret, including $r = a$ also, has regret

$$\min\left(-(1-\lambda)b + \left(\frac{n}{n+\lambda}\right)^n b, -(1-\lambda)b + \left(\frac{n-1}{n-1+\lambda}\right)^{n-1}(b-a)\right)$$

corresponding to $r = \frac{\lambda}{n+\lambda}b$ and $r = a$ respectively.

We can show that

$$\left(\frac{n-1}{n-1+\lambda}\right)^{n-1} \geq \left(\frac{n}{n+\lambda}\right)^{n}$$

So , for $\frac{a}{b} \leq 1-\left(\frac{n}{n+\lambda}\right)^{n}\left(\frac{n-1+\lambda}{n-1}\right)^{n-1}$, $r=\frac{\lambda}{n+\lambda}b$ gives the lowest worst-case regret, and for $\frac{a}{b} \geq 1-\left(\frac{n}{n+\lambda}\right)^{n}\left(\frac{n-1+\lambda}{n-1}\right)^{n-1}$, $r=a$ gives the lowest worst-case regret. We can show that

$$0 \leq 1-\left(\frac{n}{n+\lambda}\right)^{n}\left(\frac{n-1+\lambda}{n-1}\right)^{n-1} \leq \frac{\lambda}{n+\lambda}$$

so this threshold is always interior.

For $\frac{\lambda}{n+\lambda} \leq \frac{a}{b} \leq \frac{\lambda}{1+\lambda}$, the worst-case regret is increasing in r for the second regime, so the worst-case in the second regime is when $r=a^{+}$, but we already know that the worst-case regret is lower under $r=a$ than under $r=a^{+}$, so the best r is $r=a$ with regret

$$-(1-\lambda)b+\left(\frac{n-1}{n-1+\lambda}\right)^{n-1}(b-a)$$

For $\frac{a}{b} \geq \frac{\lambda}{1+\lambda}$, the second regime is not possible, and the third regime's worst case is again $r=a^{+}$ which has highest regret than $r=a$, so again the best r is $r=a$ with regret

$$-(1-\lambda)b+\left(\frac{n-1}{n-1+\lambda}\right)^{n-1}(b-a)$$

, so the r that minimizes worst-case regret is $r=\frac{\lambda}{n+\lambda}b$, which gives the regret

$$-(1-\lambda)b+\left(\frac{n}{n+\lambda}\right)^{n}b$$

If $\frac{a}{b}<\frac{\lambda}{n+\lambda}$, then the worst-case regret is minimized at $r=\frac{\lambda}{n+\lambda}b$ and the worst-case regret value is

$$-(1-\lambda)b+\frac{n^{n}}{(n+\lambda)^{n}}b$$

If $\frac{a}{b} \geq \frac{\lambda}{n+\lambda}$, then we always have $r>a \geq \frac{\lambda}{n+\lambda}b$, so the worst-case regret is minimized at $r=a^{+}$, but we have already shown that $r=a^{+}$ always has higher regret (worse) than $r=a$.

Therefore, for $\frac{a}{b} < \frac{\lambda}{n+\lambda}$, the worst-case regret is

$$\min\left(-(1-\lambda)b + \frac{(b-a)(n-1)^{n-1}}{(n-1+\lambda)^{n-1}}, \left(\frac{n^n}{(n+\lambda)^n} - 1 + \lambda\right)b, \frac{\lambda^2}{1+\lambda}b\right)$$

The first, second, and third terms correspond to $r = a$, $r \in (a, \frac{\lambda}{1+\lambda}b]$ and $r \in [\frac{\lambda}{1+\lambda}b, b]$ respectively.

The third one is higher than the second one because the third one, as we have already shown, is the second one with $r = \frac{\lambda}{1+\lambda}b$ which by our proof has higher regret than that at $r = \frac{\lambda}{1+\lambda}b$. So the worst-case regret becomes

$$\min\left(\frac{(n-1)^{n-1}}{(n-1+\lambda)^{n-1}}(b-a) - (1-\lambda)b, \frac{n^n}{(n+\lambda)^n}b - (1-\lambda)b\right)$$

We note that

$$\left(\frac{n-1}{n-1+\lambda}\right)^{n-1} \geq \left(\frac{n}{n+\lambda}\right)^n$$

This is true because $\left(\frac{x}{x+\lambda}\right)^x = \exp\left\{x\log(x) - x\log(x+\lambda)\right\}$ is a decreasing function of x: the derivative of the expression in the $\{\cdots\}$ is

$$\left(x \cdot \frac{1}{x} + \log(x)\right) - \left(x \cdot \frac{1}{x+\lambda} + \log(x+\lambda)\right) = 1 - \frac{x}{x+\lambda} + \log\left(\frac{x}{x+\lambda}\right) \leq 0$$

because $1 + \log(u) \leq u$ for all u. (Let $u' = \log(u)$; this because the well-known $1 + u' \leq \exp(u')$.)

We however have

$$\left(\frac{n-1}{n-1+\lambda}\right)^{n-1} \leq \left(\frac{n}{n+\lambda}\right)^{n-1}$$

Therefore, we have

$$\left(\frac{n-1}{n-1+\lambda}\right)^{n-1}(b-a) - (1-\lambda)b \geq \left(\frac{n}{\lambda+n}\right)^n b - (1-\lambda)b$$

when $a = 0$ but

$$\left(\frac{n-1}{n-1+\lambda}\right)^{n-1}(b-a) - (1-\lambda)b \leq \left(\frac{n}{\lambda+n}\right)^{n} b - (1-\lambda)b$$

when $a = \frac{\lambda}{n+\lambda b}$. The threshold to define which one is better is therefore always in the middle, at $1 - \left(\frac{n}{n+\lambda}\right)^n \left(\frac{n-1+\lambda}{n-1}\right)^{n-1}$.

We now consider $\frac{\lambda}{n+\lambda} \leq \frac{a}{b} \leq \frac{\lambda}{1+\lambda}$. We have shown that the worst-case regret in the second regime occurs at $r = a^+$ which is always higher regret than $r = a$, so we only need to consider the first and third regime: the worst-case regret is

$$\min\left(-(1-\lambda)b + (b-a)\left(\frac{n-1}{n-1+\lambda}\right)^{n-1}, \frac{\lambda^2}{1+\lambda}b\right) = \min\left((b-a)\left(\frac{n-1}{n-1+\lambda}\right)^{n-1}, \frac{1}{1+\lambda}b\right) - (1-\lambda)b$$

We know that $\left(\frac{n-1}{n-1+\lambda}\right)^{n-1} \leq \frac{1}{1+\lambda}$ because $n - 1 \geq 1$ and $\left(\frac{x}{x+\lambda}\right)^x$ is a decreasing function of x. We also know that $b - a \leq b$. Therefore, the first expression in the min (first regime) is always lower than the second expression (third regime). So the worst case regret in this case is just

$$-(1-\lambda)b + (b-a)\left(\frac{n-1}{n-1+\lambda}\right)^{n-1}$$

which is achieved at $r = a$.

Lastly, we consider the case $\frac{a}{b} \geq \frac{\lambda}{1+\lambda}$. Then the second regime is never applicable, and the worst-case regret is

$$\min\left(-(1-\lambda)b + (b-a)\left(\frac{n-1}{n-1+\lambda}\right)^{n-1}, \frac{\lambda^2}{1+\lambda}b\right) = -(1-\lambda)b + (b-a)\left(\frac{n-1}{n-1+\lambda}\right)^{n-1}$$

where we know the first term in the min is less than the second term by what we just proved. This is also achieved when $r = a$.

We therefore conclude that for $n \geq 2$ the worst-case regret (and the corresponding optimal reserve r^*) as a function of a and b is as follows.

$$\inf_{r\in[a,b]} R_n(\mathrm{SPA}(r), \mathcal{F}) = \begin{cases} \left(\frac{n-1}{n-1+\lambda}\right)^{n-1}(b-a) - (1-\lambda)b & \text{if } \frac{a}{b} \leq 1 - \left(\frac{n}{n+\lambda}\right)^n \left(\frac{n-1+\lambda}{n-1}\right)^{n-1} \text{ or } \frac{a}{b} \geq \frac{\lambda}{n+\lambda} \\ \left(\frac{n}{n+\lambda}\right)^n b - (1-\lambda)b & \text{if } 1 - \left(\frac{n}{n+\lambda}\right)^n \left(\frac{n-1+\lambda}{n-1}\right)^{n-1} \leq \frac{a}{b} < \frac{\lambda}{n+\lambda} \end{cases}$$

In the first case, $r^* = a$. In the second case, $r^* = \frac{\lambda}{n+\lambda} b$.

Now we calculate the optimal r^* and the best worst-case regret for the case $n = 1$. In the case $\frac{a}{b} < \frac{\lambda}{1+\lambda}$, then all 3 regimes are possible. The lowest worst-case regret in the second regime is $\frac{\lambda^2}{1+\lambda} b$ when $r = \frac{\lambda}{1+\lambda} b$, which is the same as the lowest worst-case regret in the third regime. Therefore,

$$\inf_{r \in [a,b]} R_1(\mathrm{SPA}(r), \mathcal{F}) = \min \left(\lambda b - a, \frac{\lambda^2}{1+\lambda} b \right) = \frac{\lambda^2}{1+\lambda} b$$

where the last part is true because $\frac{a}{b} < \frac{\lambda}{1+\lambda}$ implies $\lambda b - a > \frac{\lambda^2}{1+\lambda} b$. Here, $r^* = \frac{\lambda}{1+\lambda} b$.

Now consider the case $\frac{a}{b} \geq \frac{\lambda}{1+\lambda}$, then the second regime is inapplicable, and the third case holds for any $r \in (a, b]$, and the lowest worst-case regret in this regime is λa at $r = a^+$, so

$$\inf_{r \in [a,b]} R_1(\mathrm{SPA}(r), \mathcal{F}) = \min (\lambda b - a, \lambda a) = \lambda b - a$$

where the last part is true because $\frac{a}{b} \geq \frac{\lambda}{1+\lambda} b$ implies $\lambda b - a \leq \lambda a$. Here, $r^* = a$.

□

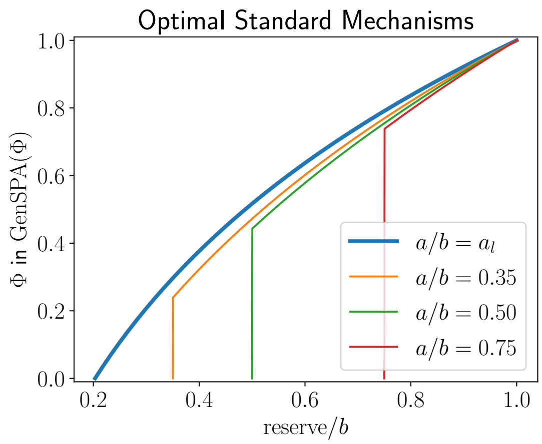
F Structure and Performance of Optimal Mechanisms within Subclasses

F.1 Structure of Optimal Mechanisms within Subclasses

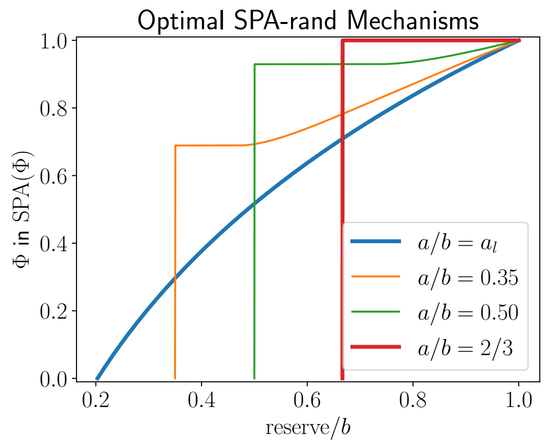
We now use the analytical results previously derived in Section 4 to show the structure and evaluate the performance of these mechanisms that are optimal in the subclasses $\mathcal{M}_{\text{std}}$, $\mathcal{M}_{\text{SPA-rand}}$, $\mathcal{M}_{\text{SPA-det}}$ and $\mathcal{M}_{\text{SPA-a}}$.

Figure 7 shows the structure of the optimal mechanisms in $\mathcal{M}_{\text{std}}$, the class of standard mechanisms, and $\mathcal{M}_{\text{SPA-rand}}$, the class of SPAs with random reserve. Note that for $a/b \leq k_l$, the optimal mechanisms in both classes are the same as the optimal mechanism in $\mathcal{M}_{\text{all}}$, namely SPA(Φ). Henceforth we will focus on the case $a/b \geq k_l$. (Note also that the distribution Φs in Figures 7a and 7b are *not* comparable, even for the same Φ, because the mechanism structures are different.)

Theorem 3 states that the optimal mechanism in $\mathcal{M}_{\text{std}}$ is GenSPA(Φ), a *generous* SPA with reserve distribution Φ, meaning that the mechanism allocates like SPA(Φ), except in the case when



(a) Φ of GenSPA(Φ) in $\mathcal{M}_{\text{std}}$



(b) Φ of SPA(Φ) in $\mathcal{M}_{\text{SPA-rand}}$

Figure 7: Structure of the minimax regret optimal mechanisms in the classes of standard mechanisms (generous SPAs) and SPAs with random reserve with $n = 2$ bidders.

all but one bidder have the lowest value a, then it allocates to the highest (and only non-lowest) value with probability one. This distribution Φ has a point mass at a and a density on $(a, b]$. Figure 7a shows the plot of Φ for $a/b \in \{k_l, 0.35, 0.50, 0.75\}$. We see that as a/b gets larger, the point mass at a becomes bigger, and the density component of Φ for different a/b are very close to each other.

Theorem 4 states that the optimal Φ of SPA(Φ) in the class $\mathcal{M}_{\text{SPA-rand}}$ has three regimes determined by the lower threshold k_l and the upper threshold $\lambda n/((1+\lambda)n-1)$. Here, we consider the minimax regret problem $\lambda = 1$ and $n = 2$ bidders, so the thresholds are $k_l \approx 0.2032$ and $2/3$. For $a/b \le k_l$, Φ has a density in $[k_l b, b]$, while for $a/b \ge 2/3$, Φ is a point mass at a, namely, no reserve is optimal. Figure 7b shows the optimal Φ in the intermediate regime $k_l \le a/b \le 2/3$. In this regime, Φ has a point mass at a and a density in $(a, b]$. We can see that the Φ in this regime interpolates between the smooth density (blue) and a single point mass (red).

G Extensions to other Distribution Classes

G.1 Formal Definitions of Distribution Classes

The main text already gives verbal definitions of distribution classes in Figure 8 that we consider, but for completeness we also give formal definitions here as well. The precise formulations of these definitions are taken from Anunrojwong et al. [2022] but they are standard; they are included here for completeness.

Definition 8. *A distribution $\boldsymbol{F}$ over n real valuations is* exchangeable *if it has the same joint distribution under any ordering of those valuations. Formally, a vector of random variables $\boldsymbol{v} = (v_1, \ldots, v_n)$ is exchangeable if for any permutation $\sigma : \{1, 2, \ldots, n\} \to \{1, 2, \ldots, n\}$, the random variables $(v_{\sigma(1)}, v_{\sigma(2)}, \ldots, v_{\sigma(n)})$ and $(v_1, \ldots, v_n)$ have the same joint distributions. A joint distribution $\boldsymbol{F}$ is exchangeable if the corresponding random variable is exchangeable.*

To give a formal definition of affiliation, we need a few auxillary definitions.[11] First, a subset A of $\mathbb{R}^n$ is called *increasing* if its indicator function $\mathbf{1}_A$ is nondecreasing. Second, a subset S of $\mathbb{R}^n$ is a *sublattice* if its indicator function $\mathbf{1}_S$ is affiliated, i.e., for any $\boldsymbol{v}, \boldsymbol{v}' \in S$ we have $\boldsymbol{v} \wedge \boldsymbol{v}' \in S$ and $\boldsymbol{v} \vee \boldsymbol{v}' \in S$ where $\boldsymbol{v} \wedge \boldsymbol{v}' = (\min(v_1, v_1'), \ldots, \min(v_n, v_n'))$ and $\boldsymbol{v} \vee \boldsymbol{v}' = (\max(v_1, v_1'), \ldots, \max(v_n, v_n'))$ denote the component-wise minimum and maximum of $\boldsymbol{v}$ and $\boldsymbol{v}'$, respectively. For a distribution $\boldsymbol{F}$ and sets A and S, we write $\Pr_{\boldsymbol{F}}(A|S)$ as the probability that a random variable drawn from the conditional distribution $\boldsymbol{F}$ restricted to S is in A.

We can now give the definition of affiliation.

Definition 9. *A distribution $\boldsymbol{F}$ is affiliated if for all increasing sets A and B and every sublattice S, $\Pr_{\boldsymbol{F}}(A \cap B|S) \geq \Pr_{\boldsymbol{F}}(A|S) \Pr_{\boldsymbol{F}}(B|S)$.*

Lastly, we give the definition of mixtures of i.i.d.

Definition 10. *$\mathcal{F}_{mix}$ is the class of all distributions $\boldsymbol{F}$ such that there exists a probability measure μ on the set of distributions satisfying $\boldsymbol{F}(\boldsymbol{v}) = \int \prod_{i=1}^n G(v_i) d\mu(G)$ for every $\boldsymbol{v}$.*

In this definition, G is the i.i.d. distribution that n agent valuations are drawn i.i.d. from, and $\mu(G)$ is the weight of G in the mixture that comprises $\boldsymbol{F}$.

G.2 Minimax λ-Regret across distribution classes

So far, we have only considered the class of i.i.d. distributions $\mathcal{F}_{\text{iid}}$. However, our results can also be extended to other distribution classes $\mathcal{F}$ that capture different forms of dependence between bidder valuations. Even when bidder valuations are not i.i.d., they are often "positively dependent" in

[11] We follow the treatment in the appendix of Milgrom and Weber [1982].

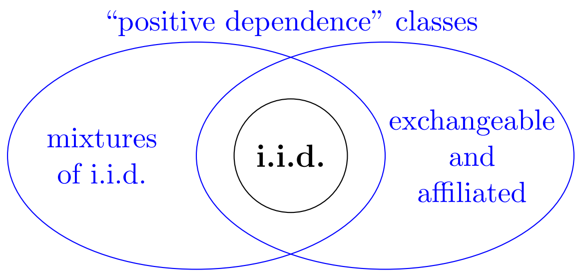


Figure 8: All these distribution classes have the same power in robust settings.

some sense. We will analyze two classes that capture this property, as shown in Figure 8. We will give verbal definitions of those classes here; the precise formal definitions are in Appendix G.1.

The first class is the class $\mathcal{F}_{\text{aff}}$ of exchangeable and affiliated distributions. A distribution is *exchangeable* if it has the same joint distribution under any permutation of the bidders. In other words, the bidders are "symmetric" in the sense that their identities are irrelevant. The notion of affiliation is introduced by Milgrom and Weber [1982] and is commonly used to model positive dependence in Bayesian mechanism design. A formal definition of *affiliation* for arbitrary distributions is quite involved[12] and is deferred to Appendix G.1, but roughly, "affiliation means that a high value of one bidder's estimate makes high values of the others' estimates more likely" [Milgrom and Weber, 1982].

The second class is the class $\mathcal{F}_{\text{mix}}$ of all mixtures of i.i.d. distributions over n bidders. Mixture of i.i.d. is commonly used to model positive dependence through latent variables. In other words, there is an unknown random state of the world, and bidder valuations are i.i.d. conditional on that state. We can also show that if bidders' values are mixtures of i.i.d. then any pair of bidder valuations have non-negative covariance, which further supports the notion that $\mathcal{F}_{\text{mix}}$ captures positive dependence.

We can show that both classes $\mathcal{F}_{\text{aff}}$ and $\mathcal{F}_{\text{mix}}$ have the *exact same power* as $\mathcal{F}_{\text{iid}}$ in the sense that all the saddle points for $\mathcal{M} \in \{\mathcal{M}_{\text{all}}, \mathcal{M}_{\text{SPA-rand}}, \mathcal{M}_{\text{SPA-det}}, \mathcal{M}_{\text{SPA-a}}\}$ in Theorems 1 and 4 and Proposition 11 still hold[13] for $\mathcal{F} \in \{\mathcal{F}_{\text{aff}}, \mathcal{F}_{\text{mix}}\}$. The proof sketch goes as follows. Because Seller's saddle fixes the distribution and optimizes over the mechanism class $\mathcal{M}$ which stays the same here, it still holds. Nature's saddle fixes the mechanism and optimizes over the distributions.

[12] In a special case when the distribution $\boldsymbol{F}$ has a density f, affiliation is equivalent to log-supermodularity, that is, for any $\boldsymbol{v}', \boldsymbol{v}''$, $f(\boldsymbol{v}' \wedge \boldsymbol{v}'')f(\boldsymbol{v}' \vee \boldsymbol{v}'') \geq f(\boldsymbol{v}')f(\boldsymbol{v}'')$, where $\wedge$ and $\vee$ are coordinate-wise min and max, respectively.

[13] Note that we do not have results for $\mathcal{M} = \mathcal{M}_{\text{std}}$ here because we only have expressions for GenSPA(Φ) with i.i.d. distributions (Proposition 9). For further discussions, see Section 4.1.

We therefore need to show that the original worst-case distribution $\boldsymbol{F}^* \in \mathcal{F}_{\text{iid}}$ is still optimal over a larger class $\mathcal{F}$. The regret expression is given by (Regret-$\boldsymbol{F}$) in Proposition 4 which is in terms of $\boldsymbol{F}_n^{(1)}(v)$ and $\boldsymbol{F}_n^{(2)}(v)$, the CDF of the 1st-highest and 2nd-highest among the n values. In the i.i.d. case, we proceed by writing both expressions in terms of the marginal CDF $F(v)$ and optimize over $F(v)$ piecewise. The key idea here is to reduce two functions to one. In our case with $\mathcal{F} \in \{\mathcal{F}_{\text{aff}}, \mathcal{F}_{\text{mix}}\}$, we let $\boldsymbol{F}_{n-1}^{(1)}(v)$ be the CDF of the 1st-highest among $n-1$ values, which is well defined because $\boldsymbol{F}$ is exchangeable. We have the identity $\boldsymbol{F}_n^{(2)}(v) = n\boldsymbol{F}_{n-1}^{(1)}(v) - (n-1)\boldsymbol{F}_n^{(1)}(v)$ [David and Nagaraja, 2003]. Now we use the key inequality from Lemma 1 of Anunrojwong et al. [2022]: $\boldsymbol{F}_{n-1}^{(1)}(v) \leq \boldsymbol{F}_n^{(1)}(v)^{(n-1)/n}$ which holds for $\mathcal{F} \in \{\mathcal{F}_{\text{aff}}, \mathcal{F}_{\text{mix}}\}$. Once we verify that the coefficient of $\boldsymbol{F}_{n-1}^{(1)}(v)$ is nonnegative, the bound for $\boldsymbol{F}_{n-1}^{(1)}(v)$ goes through, and the resulting regret bound is exactly the same as the i.i.d. expression with $F(v)$ replaced by $\boldsymbol{F}_n^{(1)}(v)^{1/n}$ and the rest of the saddle point proof therefore goes through in exactly the same way.

On the other hand, we can restrict the i.i.d. class even further to the i.i.d. regular class $\mathcal{F}_{\text{iid-reg}}$ and require that the valuation distribution F of each bidder is *regular*, meaning that if F has density f, then the virtual value function $v - (1 - F(v))/f(v)$ is a (weakly) increasing function of v.[14] Regular distributions are important because in Myerson's Bayesian mechanism design theory, if the distribution is i.i.d. and regular, the optimal mechanism is a SPA with a reserve, whereas if the distribution is irregular, an ironing procedure (which is a form of pooling) is needed. Here, we note that every worst-case distribution in the saddle points shown in Theorems 2, 6, 7 and Proposition 11 are all regular. In particular, they are all isorevenue distributions on their supports (virtual value is zero) except the case of $\mathcal{M}_{\text{all}}$ and $a/b \geq k_h$ where the distribution has a constant positive virtual value. Therefore, they are all saddle points for the same $\mathcal{M}$ and $\mathcal{F}_{\text{iid-reg}}$ as well.

In a sense, therefore, distributional classes "do not matter" in that any "reasonable" specification has the same power, but mechanism classes do matter and the power of mechanism classes are quantified by the gaps between classes (cf. Section 4.3). We formalize the results in this section in Proposition 13.

Proposition 13. *For each $\mathcal{M} \in \{\mathcal{M}_{\text{all}}, \mathcal{M}_{\text{SPA-rand}}, \mathcal{M}_{\text{SPA-det}}, \mathcal{M}_{\text{SPA-a}}\}$, we have $R_{\lambda,n}(\mathcal{M}, \mathcal{F}) = R_{\lambda,n}(\mathcal{M}, \mathcal{F}_{\text{iid}})$ for $\mathcal{F} \in \{\mathcal{F}_{\text{iid-reg}}, \mathcal{F}_{\text{iid}}, \mathcal{F}_{\text{aff}}, \mathcal{F}_{\text{mix}}\}$. We also have $R_{\lambda,n}(\mathcal{M}_{\text{std}}, \mathcal{F}_{\text{iid-reg}}) = R_{\lambda,n}(\mathcal{M}_{\text{std}}, \mathcal{F}_{\text{iid}})$.*

[14] Equivalently, the revenue function is concave in the quantity ("quantile") space. Here, we consider our distribution to be regular with a point mass at the end. We can also find a regular distribution in a standard sense that is arbitrarily close to our distribution, e.g., by replacing the point mass with a density peak of width $\epsilon \downarrow 0$.

Lastly, we can show that the first-best benchmark (considered in this paper) and the second-best benchmark give the exact same minimax λ-regret *against arbitrary distributions* $\mathcal{F}_{\text{all}}$. See Appendix I for details.

G.3 Proof of the Main Result (Proposition 13)

Proof of Proposition 13. The argument in the main text is sufficient to prove the equivalence between $\mathcal{F}_{\text{iid-reg}}$ and $\mathcal{F}_{\text{iid}}$. Henceforth, when we talk about positively dependent classes we will mean $\{\mathcal{F}_{\text{iid}}, \mathcal{F}_{\text{aff}}, \mathcal{F}_{\text{mix}}\}$.

For positively dependent $\mathcal{F}$, we will prove that the original saddle point $(m^*, \boldsymbol{F}^*)$ of $(\mathcal{M}, \mathcal{F}_{\text{iid}})$ is also a saddle point of the new problem with $(\mathcal{M}, \mathcal{F})$. Note that because $\mathcal{M}$ is the same, the Seller's Saddle is automatically satisfied. So we only need to verify Nature's Saddle.

Here, we only need two properties of $\mathcal{F}$. Firstly, that $\mathcal{F}_{\text{iid}} \subseteq \mathcal{F} \subseteq \mathcal{F}_{\text{exc}}$ (so $\boldsymbol{F}^*$ from i.i.d. is still a valid distribution in $\mathcal{F}$). At the same time, exchangeability implies that the notion of $\boldsymbol{F}_{n-1}$ for the restriction to $n-1$ agents is well-defined and independent of agent identities. Secondly, for any $\boldsymbol{F} \equiv \boldsymbol{F}_n \in \mathcal{F}$, $\boldsymbol{F}_n^{(1)}(v)^{1/n} \geq \boldsymbol{F}_{n-1}^{(1)}(v)^{1/(n-1)}$ for every v. This is proved in Lemma 1 of Anunrojwong et al. [2022] in the case of $\mathcal{F}_{\text{aff}}$. For the case of $\mathcal{F}_{\text{mix}}$, this inequality follows from Jensen's inequality or power mean inequality. Namely, if $\boldsymbol{F}(\boldsymbol{v}) = \int \prod_{i=1}^n G(v_i) d\mu(G)$, then the inequality is $\left(\int G(v)^n d\mu(G)\right)^{1/n} \geq \left(\int G(v)^{n-1} d\mu(G)\right)^{1/(n-1)}$ which is true.

We now work out the proof strategy outlined above in cases.

Case $\mathcal{M} \in \{\mathcal{M}_{\text{all}}, \mathcal{M}_{\text{SPA-rand}}\}$. For both $\mathcal{M}_{\text{all}}$ and $\mathcal{M}_{\text{SPA-rand}}$ against $\mathcal{F}_{\text{iid}}$, the optimal mechanism in the saddle point is a (g_u^*, g_d^*) mechanism with continuous functions g_u^*, g_d^*.

We use the identity $\boldsymbol{F}_n^{(2)}(v) = n\boldsymbol{F}_n^{(1)}(v) - (n-1)\boldsymbol{F}_{n-1}^{(1)}(v)$ (see, e.g., David and Nagaraja [2003]). The regret expression from Proposition 4 then gives $\boldsymbol{F}_n^{(2)}(v) = n\boldsymbol{F}_n^{(1)}(v) - (n-1)\boldsymbol{F}_{n-1}^{(1)}(v)$:

$$\begin{aligned} R(g^*, \boldsymbol{F}) = a(\lambda - g_u^*(a) - (n-1)g_d^*(a)) &+ \int_{v=a}^{v=b} (\lambda - g_u^*(v) + g_d^*(v) - v g_u'^*(v) + (v - na) g_d'^*(v)) dv \\ &+ \int_{v=a}^{v=b} \left(-\lambda - (n-1)(g_u^*(v) - g_d^*(v)) + v(g_u'^*(v) + (n-1)g_d'^*(v))\right) \boldsymbol{F}_n^{(1)}(v) \\ &+ \int_{v=a}^{v=b} n(g_u^*(v) - g_d^*(v) - (v-a)g_d'^*(v)) \boldsymbol{F}_{n-1}^{(1)}(v) dv \end{aligned}$$

We know that $g_u^*(v) - g_d^*(v) - (v-a)g_d'^*(v) \geq 0$ due to (SOC). Therefore, we can use the bound

$\boldsymbol{F}^{(1)}_{n-1}(v) \geq \boldsymbol{F}^{(1)}_n(v)^{(n-1)/n}$. Writing $\tilde{F}(v) = \boldsymbol{F}_n(v)^{1/n}$, we therefore have

$$\begin{aligned} R(g^*, \boldsymbol{F}) \leq a(\lambda - g^*_u(a) - (n-1)g^*_d(a)) + \int_{v=a}^{v=b} (\lambda - g^*_u(v) + g^*_d(v) - v g'^*_u(v) + (v - na) g'^*_d(v)) dv \\ + \int_{v=a}^{v=b} \Big[\left(-\lambda - (n-1)(g^*_u(v) - g^*_d(v)) + v(g'^*_u(v) + (n-1)g'^*_d(v))\right) \tilde{F}(v)^n \\ + n(g^*_u(v) - g^*_d(v) - (v-a)g'^*_d(v))\tilde{F}(v)^{n-1}\Big] dv = \sup_{\tilde{F}(v) \in [0,1]} R(g^*, \tilde{F}^n) \end{aligned}$$

Note that $\tilde{F}(v) = \boldsymbol{F}_n(v)^{1/n}$ is (weakly) increasing in v, and $\tilde{F}(b) = 1$, so $\tilde{F}$ is a valid CDF of a distribution. $R(g^*, \tilde{F}^n)$ is the regret of the $g^* = (g^*_u, g^*_d)$ mechanism against $\tilde{F}^n$, or the distribution with n i.i.d. with marginal $\tilde{F}$. By definition of g^*, fixing g^*, the (i.i.d.) distribution $(F^*)^n$ maximizes the regret, so $R(g^*, \tilde{F}^n) \leq R(g^*, (F^*)^n) = R(g^*, \boldsymbol{F}^*)$. Therefore, we have proved Nature's saddle.

Case $\mathcal{M} \in \{\mathcal{M}_{\text{SPA-det}}, \mathcal{M}_{\text{SPA-a}}\}$. We have the following expression for SPA(r) against a joint distribution $\boldsymbol{F} \in \{\mathcal{F}_{\text{iid}}, \mathcal{F}_{\text{aff}}, \mathcal{F}_{\text{mix}}\}$:

$$-(1-\lambda)b + r\boldsymbol{F}^{(1)}_n(r^-) - \lambda \int_{v \in [a,r]} \boldsymbol{F}^{(1)}_n(v) dv + \int_{v \in (r,b]} (\boldsymbol{F}^{(2)}_n(v) - \lambda \boldsymbol{F}^{(1)}_n(v)) dv$$

With $r > a$, we have that the regret is upper bounded by

$$\begin{aligned} &\leq -(1-\lambda)b + r\boldsymbol{F}^{(1)}_n(r) + \int_{v \in (r,b]} (\boldsymbol{F}^{(2)}_n(v) - \lambda \boldsymbol{F}^{(1)}_n(v)) dv \\ &= r\boldsymbol{F}^{(1)}_n(r) + \int_{v \in (r,b]} (n\boldsymbol{F}^{(1)}_{n-1}(v) - (n-1+\lambda)\boldsymbol{F}^{(1)}_n(v)) dv \\ &\leq r\boldsymbol{F}^{(1)}_n(r) + \int_{v \in (r,b]} (n\boldsymbol{F}^{(1)}_n(v)^{(n-1)/n} - (n-1+\lambda)\boldsymbol{F}^{(1)}_n(v)) dv \\ &\leq \sup_{c \in [0,1]} \left[rc^n + \sup_{z \in [c,1]} \left\{ nz^{n-1} - (n-1+\lambda)z^n \right\} \right] \end{aligned}$$

The third line uses the equation $\boldsymbol{F}^{(2)}_n(v) = n\boldsymbol{F}^{(1)}_{n-1}(v) - (n-1)\boldsymbol{F}^{(1)}_n(v)$. The fourth line uses the inequality $\boldsymbol{F}^{(1)}_{n-1}(v) \leq \boldsymbol{F}^{(1)}_n(v)^{(n-1)/n}$. The fifth line is a pointwise maximization over $c = \boldsymbol{F}^{(1)}_n(r)^{1/n}$ and z denoting $\boldsymbol{F}^{(1)}_n(v)^{1/n}$ for $v > r$ which is $\geq c$. The last expression is the same expression that we got for $\mathcal{F}_{\text{iid}}$, so it has the same minimax regret.

For $r = a$, the regret is

$$
\begin{aligned}
& -(1-\lambda)b + \int_{v\in(a,b]} (\boldsymbol{F}_n^{(2)}(v) - \lambda \boldsymbol{F}_n^{(1)}(v))dv \\
&= -(1-\lambda)b + \int_{v\in(a,b]} (n\boldsymbol{F}_{n-1}^{(1)}(v) - (n-1+\lambda)\boldsymbol{F}_n^{(1)}(v))dv \\
&\le -(1-\lambda)b + \int_{v\in(a,b]} (n\boldsymbol{F}_n^{(1)}(v)^{(n-1)/n} - (n-1+\lambda)\boldsymbol{F}_n^{(1)}(v))dv \\
&\le -(1-\lambda)b + (b-a)\sup_{z\in[0,1]} nz^{n-1} - (n-1+\lambda)z^n
\end{aligned}
$$

The last expression is the same expression that we got for $\mathcal{F}_{\text{iid}}$, so it has the same minimax regret.

□

H Quantifying The Gap Between First-Best and Second-Best

In this work, we consider the benchmark which is the maximum revenue when the valuations are known to the seller. We can call this the *first best* and denote it by $\text{FirstBest}(\boldsymbol{F}) := \mathbb{E}_{\boldsymbol{v}\sim\boldsymbol{F}} \max(\boldsymbol{v})$. Another quantity of potential interest is the maximum revenue when the *distribution*, but not the valuations, are known to the seller. We call this the *second best* and denote it by $\text{SecondBest}(\boldsymbol{F};\mathcal{M}) = \sup_{(x,p)\in\mathcal{M}} \mathbb{E}_{\boldsymbol{v}\sim\boldsymbol{F}} \sum_{i=1}^n p_i(\boldsymbol{v})$.

Our results (Theorem 5, 6, 7 and Proposition 11) explicitly show that the minimax λ-regret problem has a *pure* saddle point, that is, Nature's saddle point is a single distribution $\boldsymbol{F}$ rather than a mixture of distributions.[15] In other words, the zero-sum game between Seller and Nature has a pure Nash equilibrium. It follows immediately as a corollary that the ordering of the min and the max can be swapped:

$$
\begin{aligned}
\inf_{m\in\mathcal{M}} \sup_{\boldsymbol{F}\in\mathcal{F}} \text{Regret}(m,\boldsymbol{F}) &= \sup_{\boldsymbol{F}\in\mathcal{F}} \inf_{m\in\mathcal{M}} \mathbb{E}_{\boldsymbol{v}\sim\boldsymbol{F}} \left[\max(\boldsymbol{v}) - \sum_{i=1}^n p_i(\boldsymbol{v})\right] &&= \sup_{\boldsymbol{F}\in\mathcal{F}} [\text{FirstBest}(\boldsymbol{F}) - \text{SecondBest}(\boldsymbol{F};\mathcal{M})] \\
\sup_{m\in\mathcal{M}} \inf_{\boldsymbol{F}\in\mathcal{F}} \text{Ratio}(m,\boldsymbol{F}) &= \inf_{\boldsymbol{F}\in\mathcal{F}} \sup_{m\in\mathcal{M}} \frac{\mathbb{E}_{\boldsymbol{v}\sim\boldsymbol{F}} \sum_{i=1}^n p_i(\boldsymbol{v})}{\mathbb{E}_{\boldsymbol{v}\sim\boldsymbol{F}} \max(\boldsymbol{v})} &&= \inf_{\boldsymbol{F}\in\mathcal{F}} \frac{\text{SecondBest}(\boldsymbol{F};\mathcal{M})}{\text{FirstBest}(\boldsymbol{F})}.
\end{aligned}
$$

Therefore, the minimax regret and maximin ratio values for the mechanism class $\mathcal{M}$ against the

[15]There is a caveat: for SPA(r) with deterministic reserve $r > a$, the worst-case distribution has a point mass at $r - \epsilon$ which does not achieve exact optimality but can get arbitrarily close to optimal as $\epsilon \downarrow 0$. The same argument holds true with ϵ-slacks, which disappear in the limit.

valuation distribution $\mathcal{F}$ are precisely the worst-case absolute and relative gap between the first-best and the second-best revenues. The ratios in Table 2 quantify the value of information, i.e., the relative revenue loss the seller needs to incur from not knowing the buyers' value and eliciting them truthfully. (As Section G.2 will show, the same pure saddle points in previous theorems also hold for distribution classes $\mathcal{F}$ other than $\mathcal{F}_{\text{iid}}$ also, and thus so are the results in this subsection.)

I Equivalence of First-Best and Second-Best Benchmarks Against Arbitrary Distributions

We can show that against arbitrary distributions, the first-best benchmark (considered in this paper) and the second-best benchmark give the exact same minimax λ-regret, even though on the surface it seems like the first-best should be bigger than the second-best.

Proposition 14 (Equivalence of First-Best and Second-Best Benchmarks Against Arbitrary Distributions)**.** *For $\mathcal{M} \in \{\mathcal{M}_{\text{all}}, \mathcal{M}_{\text{std}}, \mathcal{M}_{\text{SPA-rand}}, \mathcal{M}_{\text{SPA-det}}\}$ and $\mathcal{F} = \mathcal{F}_{\text{all}}$, we have*

$$\begin{aligned}
&\inf_{(x,p)\in\mathcal{M}} \sup_{\boldsymbol{F}\in\mathcal{F}} \left\{ \lambda \text{FirstBest}(\boldsymbol{F}) - \mathbb{E}_{\boldsymbol{v}\sim\boldsymbol{F}}\left[\sum_{i=1}^{n} p_i(\boldsymbol{v})\right] \right\} \\
&= \inf_{(x,p)\in\mathcal{M}} \sup_{\boldsymbol{F}\in\mathcal{F}} \left\{ \lambda \text{SecondBest}(\boldsymbol{F};\mathcal{M}) - \mathbb{E}_{\boldsymbol{v}\sim\boldsymbol{F}}\left[\sum_{i=1}^{n} p_i(\boldsymbol{v})\right] \right\}
\end{aligned}$$

Note that for $n = 1$, we have $\mathcal{F}_{\text{all}} = \mathcal{F}_{\text{exc}} = \mathcal{F}_{\text{aff}} = \mathcal{F}_{\text{mix}} = \mathcal{F}_{\text{iid}}$, so the equivalence between first best and second best always holds in the one-bidder no-competition case. The intuition behind this equality is that when $\mathcal{F}$ is unrestricted, the worst-case $\boldsymbol{F}$ against second best is essentially a collection of one-point distributions, where the firs-best and second-best benchmarks are the same.

Proof of Proposition 14. Let

$$\begin{aligned}
R_{\lambda,n,FB}(\mathcal{M},\mathcal{F}) &= \inf_{(x,p)\in\mathcal{M}} \sup_{\boldsymbol{F}\in\mathcal{F}} \lambda \text{FirstBest}(\boldsymbol{F}) - \mathbb{E}_{\boldsymbol{v}\sim\boldsymbol{F}}\left[\sum_{i=1}^{n} p_i(\boldsymbol{v})\right] \\
R_{\lambda,n,SB}(\mathcal{M},\mathcal{F}) &= \inf_{(x,p)\in\mathcal{M}} \sup_{\boldsymbol{F}\in\mathcal{F}} \lambda \text{SecondBest}(\boldsymbol{F};\mathcal{M}) - \mathbb{E}_{\boldsymbol{v}\sim\boldsymbol{F}}\left[\sum_{i=1}^{n} p_i(\boldsymbol{v})\right]
\end{aligned}$$

We want to show that for $\mathcal{M} \in \{\mathcal{M}_{\text{all}}, \mathcal{M}_{\text{std}}, \mathcal{M}_{\text{SPA-rand}}, \mathcal{M}_{\text{SPA-det}}\}$, $R_{\lambda,n,FB}(\mathcal{M},\mathcal{F}_{\text{all}}) = R_{\lambda,n,SB}(\mathcal{M},\mathcal{F}_{\text{all}})$.

Clearly, $\text{FirstBest}(\boldsymbol{F}) \geq \text{SecondBest}(\boldsymbol{F};\mathcal{M})$, so $R_{\lambda,n,FB}(\mathcal{M},\mathcal{F}_{\text{all}}) \geq R_{\lambda,n,SB}(\mathcal{M},\mathcal{F}_{\text{all}})$. We will now prove the opposite direction $R_{\lambda,n,FB}(\mathcal{M},\mathcal{F}_{\text{all}}) \leq R_{\lambda,n,SB}(\mathcal{M},\mathcal{F}_{\text{all}})$. For each $\boldsymbol{v} \in [a,b]^n$, let $\delta_{\boldsymbol{v}}$ be a distribution that puts all weight on a single point $\boldsymbol{v}$. Note that for any of our $\mathcal{M}$, $\text{FirstBest}(\delta_{\boldsymbol{v}}) = \text{SecondBest}(\delta_{\boldsymbol{v}};\mathcal{M}) = \max(\boldsymbol{v})$. This is because the mechanism $\text{SPA}(\max(\boldsymbol{v})) \in \mathcal{M}$ achieves the first best revenue $\max(\boldsymbol{v})$ against $\delta_{\boldsymbol{v}}$. Therefore,

$$
\begin{aligned}
&R_{\lambda,n,FB}(\mathcal{M},\mathcal{F}_{\text{all}}) \\
&= \inf_{(x,p)\in\mathcal{M}} \sup_{\boldsymbol{F}\in\mathcal{F}} \mathbb{E}_{\boldsymbol{v}\sim\boldsymbol{F}}\left[\lambda\max(\boldsymbol{v}) - \sum_{i=1}^{n} p_i(\boldsymbol{v})\right] \\
&\leq \inf_{(x,p)\in\mathcal{M}} \sup_{\boldsymbol{v}\in[a,b]^n} \left[\lambda\max(\boldsymbol{v}) - \sum_{i=1}^{n} p_i(\boldsymbol{v})\right] \\
&= \inf_{(x,p)\in\mathcal{M}} \sup_{\delta_{\boldsymbol{v}}\in\mathcal{F}_{\text{all}}} \lambda\text{SecondBest}(\delta_{\boldsymbol{v}};\mathcal{M}) - \mathbb{E}_{\tilde{v}\sim\delta_{\boldsymbol{v}}}\left[\sum_{i=1}^{n} p_i(\boldsymbol{v})\right] \\
&\leq \inf_{(x,p)\in\mathcal{M}} \sup_{\boldsymbol{F}\in\mathcal{F}_{\text{all}}} \lambda\text{SecondBest}(\boldsymbol{F};\mathcal{M}) - \mathbb{E}_{\tilde{\boldsymbol{v}}\sim\boldsymbol{F}}\left[\sum_{i=1}^{n} p_i(\tilde{\boldsymbol{v}})\right] \\
&= R_{\lambda,n,SB}(\mathcal{M},\mathcal{F}_{\text{all}})
\end{aligned}
$$

where the last inequality holds because $\delta_{\boldsymbol{v}} \in \mathcal{F}_{\text{all}}$ for every $\boldsymbol{v}$, so maximizing over δ_v is weakly less than maximizing over $\mathcal{F}_{\text{all}}$.

□

When $\mathcal{F} = \mathcal{F}_{\text{iid}}$ is restricted, the one-point-distribution argument no longer works because it is not i.i.d. In fact, we expect that the equality would no longer hold, and that the second-best problem is significantly more challenging. While the proofs in this paper rely on finding a pure saddle point, this technique cannot work against the second-best benchmark because the saddle point must be mixed. This is because, as discussed in Appendix H, having a pure saddle point means we can swap min and max, but (assuming $\lambda = 1$, standard regret), for any fixed $\boldsymbol{F}$, the inf over all mechanisms of the regret against the second-best benchmark is zero by definition, and the problem becomes degenerate; clearly, strong duality does not hold here. To tackle the second-best benchmark against i.i.d. distributions, we need to develop new techniques to identify and handle mixed saddle points, and this is an exciting direction for future work.