A Strong Minimax Theorem for Informationally-Robust Auction Design*

Benjamin Brooks Songzi Du

October 12, 2020

Abstract

We study the design of profit-maximizing mechanisms in environments with interdependent values. A single unit of a good is for sale. There is a known joint distribution of the bidders' ex post values for the good. Two programs are considered:

(i) Maximize over mechanisms the minimum over information structures and equilibria of expected profit;

(ii) Minimize over information structures the maximum over mechanisms and equilibria of expected profit.

These programs are shown to have the same optimal value, which we term the *profit guarantee*.

In addition, we characterize a family of linear programs that relax (i) and produce, for any finite number of actions, a mechanism with a corresponding lower bound on equilibrium profit. An analogous family of linear programs relax (ii) and produce, for any finite number of signals, an information structure with a corresponding upper bound on equilibrium profit. These lower and upper bounds converge to the profit guarantee as the numbers of actions and signals grow large.

Our model can be extended to allow for demand constraints, multiple goods, and ambiguity about the value distribution. We report numerical simulations of approximate solutions to (i) and (ii).

Keywords: Mechanism design, information design, optimal auctions, profit maximization, interdependent values, max-min, Bayes correlated equilibrium, direct mechanism.

JEL Classification: C72, D44, D82, D83.

*Brooks: Department of Economics, University of Chicago, babrooks@uchicago.edu; Du: Department of Economics, University of California San Diego, sodu@ucsd.edu. This research has been supported by NSF #1757222.

1 Introduction

In recent work, we studied an informationally robust auction design problem in which the bidders have a pure common value for the good being sold, drawn from a known prior distribution (Brooks and Du, 2020). The main result of that paper is the construction of a mechanism and a (common-prior) information structure for the bidders such that the mechanism maximizes minimum expected profit across all information structures and equilibria, the information structure minimizes maximum expected profit across all mechanisms and equilibria, and the min-max and max-min profits coincide. This model can roughly be described as a game between a Seller, who chooses the mechanism to maximize profit, and adversarial Nature, who chooses the information structure to minimize profit. It is not a standard zero-sum game, however, because profit depends on which equilibrium is played, and the set of equilibria depends in a complicated manner on the mechanism and the information structure. It is therefore a non-trivial result that min-max profit is equal to max-min profit and that this remains true regardless of which equilibrium is selected.

The purpose of the present paper is to generalize this finding. First and foremost, we drop the common value hypothesis and instead allow for an arbitrary finite-support joint distribution over bidders' values. The only substantive assumption for our results to hold is that the support of the distribution is a product set, although the probability of a given value profile can be arbitrarily small. In addition, the techniques we develop can accommodate a variety of auction design problems, such as additional feasibility constraints on the allocation, the simultaneous auction of multiple goods, and ambiguity about the value distribution.

Throughout our analysis, we restrict attention to finite mechanisms and finite information structures, although the numbers of actions and signals can be arbitrarily large. *Max-2min profit* is defined to be the supremum over mechanisms of the infimum over information structures and Bayes Nash equilibria of expected profit. (The "2" here indicates that we are minimizing over two objects, the information structure and the equilibrium.) Similarly, *min-2max profit* is defined to be the infimum over information structures of the supremum over mechanisms and equilibria of expected profit. Our first main result, Theorem 1, says that max-2min profit is equal to min-2max profit. We refer to this as a *strong minimax theorem*, since the equality of max-min and min-max profit holds regardless of how we select an equilibrium. We refer to the optimal profit level as *the profit guarantee*. Our second main result, Theorem 2, concerns a pair of sequences of linear programs whose solutions are mechanisms and information structures. Theorem 2 shows that these objects have lower and upper bounds on profit, respectively, that converge to the profit guarantee as the number of actions and signals goes to infinity. Thus, the solutions are approximate max-2min mechanisms and min-2max information structures.

To prove these theorems, we study versions of the max-2min and min-2max programs for fixed numbers of actions and signals, respectively. We then relax these programs to obtain upper and lower bounds on min-2max and max-2min profit. These relaxations are the linear programs characterized in Theorem 2. We argue that the relaxations have asymptotically the same optimal value, in the limit as the numbers of actions and signals is large, which is precisely the profit guarantee. The relaxations are obtained by (i) fixing an (arbitrary) order on the actions/signals, (ii) dropping all equilibrium conditions except those pertaining to "local" deviations, and (iii) normalizing the units for signals and actions so that the shadow cost of all local incentive constraints is the same. The resulting relaxations are linear programs. Moreover, it turns out that these programs are "almost" a dual pair, except that in the min-2max relaxation, the local downward constraints bind, as in the standard auction design model of Myerson (1981), whereas in the max-2min relaxation, it is local upward constraints that bind. The remainder of the proof shows that in spite of this difference, the duality gap is small when the number of actions and signals is large.

As mentioned previously, the linear relaxations of the max-2min and min-2max programs produce approximate max-2min mechanisms and min-2max information structures. After proving Theorems 1 and 2, we present numerical simulations of these objects for a wide variety of examples, including: Pure common values; asymmetric perfectly correlated values; independent values; pure common values with asymmetric demands; multiple goods auctioned simultaneously; and ambiguous correlation between values. The simulations suggest features of the saddle point in the continuum action/signal limit. We stop short of providing an analytical characterization of the limit solution. Part of the reason is that there are generally many solutions of the linear relaxations, and additional properties may be needed to isolate those that are theoretically tractable and of practical interest. We comment further on this issue in our discussion. Nonetheless, these simulations can motivate and guide the limit analysis, as they did for our earlier results on pure common values.

After the extensions, we discuss additional theoretical topics, starting with properties of the approximate max-2min mechanisms. We then characterize a lower bound on the rate of convergence of the approximations to the optimal value. We also show that the bounds are linear and continuous in the underlying fundamental distribution of values.

This paper is related to large literatures in mechanism design and information design. As discussed above, the most closely related paper is Brooks and Du (2020),[1] which studies a version of the present problem under the assumption of pure common values. In that paper, we explicitly constructed a saddle point in the limit model with continua of actions and signals. The max-2min mechanism is what we termed a "proportional auction," in which the sum of bidders' allocations and the sum of bidders payments only depend on the sum of their actions, and individual allocations and payments are proportional to actions. We also showed that the limit solution can be approximated by finite mechanisms and information structures. We have no reason to think that this particular form of the saddle point will generalize beyond the pure common value model, and we do not construct continuous solutions for the class of environments considered here. Instead, we argue non-constructively that the profit guarantee exists and can be approximated with finite mechanisms and information structures. Thus, a major part of our present contribution is to develop new tools for simulation, which can then be used to motivate constructions, either finite or continuous, such as those in Brooks and Du (2020).

Our analysis draws on techniques from the theory of mechanism design and information design. With respect to the former, we use direct revelation mechanisms and local

[1] That paper built on earlier analysis of informationally robust auctions under common values by Du (2018) and Bergemann, Brooks, and Morris (2016).

relaxations to bound max-2min profit, as in the revenue equivalence arguments of Myerson (1981). Our bounds on min-2max profit use analogous revelation arguments for information design, specifically *Bayes correlated equilibrium* (BCE) introduced by Bergemann and Morris (2013, 2016).

Various papers have studied informationally robust auction design under the assumption that values are private, including Chung and Ely (2007), Yamashita (2016), Chen and Li (2018). In contrast, our model allows for values to be interdependent. Yamashita and Zhu (2018) also study robust mechanism design with interdependent values, but they focus on conditions under which ex-post incentive compatible mechanisms are also max-min optimal when the Seller-preferred equilibrium is selected. Other related studies of robust mechanism design include Neeman (2003), Brooks (2013), Yamashita (2015), Carroll (2017), Bergemann, Brooks, and Morris (2019), and the literature on algorithmic mechanism design (e.g., Hartline and Roughgarden, 2009).

We conclude this introduction by discussing possible interpretations. Our results can be understood literally as predicting the choices of a Seller who evaluates each mechanism by its worst-case profit across all information structures and equilibria. We do not believe that real-world auction designers have such extreme preferences. At the same time, we suspect that designers in a practical setting may be unable or unwilling to commit to a single information structure and a single equilibrium as the correct description of behavior, as required by the classical Bayesian auction design paradigm. Our view is that the truth is somewhere in between: Designers may know some features of bidders' information without being able to give a complete description. Of course, the ambiguity of bidders' information may be accompanied by distinct concerns about the complexity of the mechanism and/or the accuracy of the equilibrium prediction. It is beyond our present abilities to incorporate all such concerns into the theory of optimal auctions. We can, however, ask what mechanisms are robust to ambiguity about bidders' information in an extreme sense, provided we are still willing to accept the common prior and Bayes Nash equilibrium as an as-if description of behavior.

In our view, the greatest promise of this approach is that it may lead to the discovery of new auction designs, such as the proportional auction, that are compelling both for their optimal worst-case performance as well as their simplicity.[2] The worst-case performance of a mechanism is, in a sense, a measure of how "safe" it is. To be sure, it is just one of many criteria that might be considered in applied auction design. For example, one may also care how the auction performs on particular, benchmark information structures, such as affiliated values. Importantly, there need not be conflict between these criteria: when values are common and the number of bidders is large, the profit guarantee is approximately the entire surplus, so that max-2min mechanisms are near optimal in all information structures (Du, 2018; Brooks and Du, 2020). This will not always be the case, however, and an important task for future work is to evaluate max-2min auctions on particular information structures and under different solution concepts. Such analyses will lead to a more balanced view of

[2]The worst-case analysis naturally leads to a great deal of structure on information and mechanisms, which we view as being relatively "simple," at least compared to the benchmark of full surplus extraction mechanisms in correlated type spaces (Crémer and McLean, 1988; McAfee et al., 1989): As we show, there always exist approximate min-2max information structures with independent signals.

the merits and demerits of the max-2min auctions, and the tradeoff between informational robustness and Bayesian optimality.

The rest of this paper proceeds as follows. Section 2 describes a baseline model where a single unit is being auctioned to bidders with single unit demand, and there is a known joint distribution of the bidders' values. Section 3 states and proves the strong minimax theorem and characterizes the bounding programs. Numerical examples immediately follow the proof. Section 4 describes various extensions of the baseline model, with accompanying simulations. Section 5 discusses additional theoretical topics, such as features of max-2min mechanisms and min-2max information structures, comparative statics in the value distribution, and the rate of convergence to the profit guarantee. Section 6 is a conclusion.

2 Model

One unit of a good is for sale to a finite group of bidders, indexed by $i = 1, \ldots, N$. Each bidder i demands a single unit at a value of $v_i \in \mathbb{R}_+$, with the joint distribution of values being given by $\mu \in \Delta(\mathbb{R}_+^N)$. We assume that μ has a finite support contained in $V = \times_{i=1}^N V_i$, where $V_i \subseteq \mathbb{R}_+$ is a finite set of values for bidder i. We let $\overline{v} = \max_i \max V_i$.

A (finite) *information structure* consists of a finite set of signals S_i for each bidder, with $S = \times_{i=1}^N S_i$, and a joint distribution $\sigma \in \Delta(V \times S)$ such that the marginal of σ on V is μ. An information structure is denoted by $\mathcal{I} = (S, \sigma)$. We let $\mathbf{I}(S)$ denote the set of information structures with signal space S. We let $\mathbf{I}$ denote the set of finite information structures.[3]

A (finite) *mechanism* consists of a finite set of actions A_i for each bidder, with $A = \times_{i=1}^N A_i$; an allocation rule $q : A \to [0, 1]^N$ and $\Sigma q(a) \leqslant 1$ for all $a \in A$;[4] and a transfer rule $t : A \to \mathbb{R}^N$. A mechanism is denoted by $\mathcal{M} = (A, q, t)$. A mechanism is *participation secure* if for all i, there exists an action $0 \in A_i$ such that $t_i(0, a_{-i}) = 0$ for all $a_{-i} \in A_{-i} = \times_{j \neq i} A_j$. We let $\mathbf{M}(A)$ denote the set of participation-secure mechanisms with action space A (which implicitly includes a zero action for each bidder). $\mathbf{M}$ is the set of finite mechanisms.

A mechanism and information structure $(\mathcal{M}, \mathcal{I})$ are a Bayesian game, in which bidder i's strategy is a mapping $b_i : S_i \to \Delta(A_i)$. A strategy profile $b = (b_1, \ldots, b_N)$ is identified with the kernel $b : S \to \Delta(A)$ where $b(s)$ is the product measure $\times_{i=1}^N b_i(s_i)$. Profit from a strategy profile b of a game $(\mathcal{M}, \mathcal{I})$ is

$$\Pi(\mathcal{M}, \mathcal{I}, b) = \sum_{v \in V} \sum_{s \in S} \sum_{a \in A} \Sigma t(a) b(a|s) \sigma(s, v).$$

Bidder i's surplus/utility from a strategy profile b is

$$U_i(b) = \sum_{v \in V} \sum_{s \in S} \sum_{a \in A} (v_i q_i(a) - t_i(a)) b(a|s) \sigma(s, v).$$

[3]The set of finite information structures exists because we can identify finite sets of signals with finite subsets of $\mathbb{N}$. Likewise for the set of finite mechanisms.

[4]Throughout the paper, we adopt the convention that for a vector $x \in \mathbb{R}^N$, Σx denotes the sum $x_1 + \cdots + x_N$.

A *(Bayes Nash) equilibrium* is a strategy profile b such that $U_i(b) \geqslant U_i(b'_i, b_{-i})$ for all i and strategies b'_i. We let $B(\mathcal{M}, \mathcal{I})$ denote the set of equilibria for the game $(\mathcal{M}, \mathcal{I})$, which is always non-empty since the mechanism and information structure are finite.

3 A Strong Minimax Theorem

3.1 Main results

Our first main result is the following:

Theorem 1 (Strong Minimax)**.** *Suppose $\mu(v) > 0$ for all $v \in V$. Then,*

$$\inf_{\mathcal{I} \in \mathbf{I}} \sup_{\mathcal{M} \in \mathbf{M}} \sup_{b \in B(\mathcal{M},\mathcal{I})} \Pi(\mathcal{M}, \mathcal{I}, b) = \sup_{\mathcal{M} \in \mathbf{M}} \inf_{\mathcal{I} \in \mathbf{I}} \inf_{b \in B(\mathcal{M},\mathcal{I})} \Pi(\mathcal{M}, \mathcal{I}, b). \tag{1}$$

We denote the value in (1) by Π^*. The left-hand side of (1) is the min-2max program, as described in the introduction, whose value is min-2max profit, and the right-hand side is the max-2min program, whose value is max-2min profit. Theorem 1 says that min-2max profit is equal to max-2min profit. Equivalently, for any $\epsilon > 0$, the Seller has a mechanism which guarantees a profit of at least $\Pi^* - \epsilon$ across all information structures and equilibria, and Nature has an information structure which guarantees a profit of at most $\Pi^* + \epsilon$ across all mechanisms and equilibria.

Theorem 1 is essentially a minimax theorem for the zero-sum game in which the Seller chooses the mechanism to maximize profit and Nature adversarially chooses information to minimize profit. There are some key differences between this result and standard minimax theorems, as we now explain.

First, the aforementioned zero-sum competition between Seller and Nature is not quite a game, because for a given mechanism and information structure, there may be multiple equilibria with different profit levels (although an equilibrium always exists due to the restriction to finite actions and signals). In the left-hand side of (1), in which the Seller has the second-mover advantage, we have effectively allowed the Seller to also choose the equilibrium, thus giving the most pessimistic value for Nature. Similarly, in the right-hand side of (1), where Nature has the second-mover advantage, we also have Nature choose the equilibrium, thus giving the most pessimistic value for the Seller. The theorem therefore implies that the values of these programs would hold regardless of how we selected an equilibrium. We now state this result formally:

Corollary 1. *Fix a selection $b(\mathcal{M}, \mathcal{I}) \in B(\mathcal{M}, \mathcal{I})$ from the equilibrium correspondence B on $\mathbf{M} \times \mathbf{I}$. If $\mu(v) > 0$ for all $v \in V$, then*

$$\inf_{\mathcal{I} \in \mathbf{I}} \sup_{\mathcal{M} \in \mathbf{M}} \Pi(\mathcal{M}, \mathcal{I}, b(\mathcal{M}, \mathcal{I})) = \sup_{\mathcal{M} \in \mathbf{M}} \inf_{\mathcal{I} \in \mathbf{I}} \Pi(\mathcal{M}, \mathcal{I}, b(\mathcal{M}, \mathcal{I})).$$

Second, in a standard zero-sum game, the problem of computing the value can be reformulated as a linear program. As we will see, the equilibrium constraints prevent us from following the same approach, and both the min-2max and max-2min in (1) are non-linear programs.

Finally, Theorem 1 does not hold when we impose an upper bound on the number of actions/signals. Fix a finite signal space S and define

$$\Pi^{\text{MIN}-2\text{MAX}}(S) = \inf_{\mathcal{I} \in \mathbf{I}(S)} \sup_{\mathcal{M} \in \mathbf{M}} \sup_{b \in B(\mathcal{M}, \mathcal{I})} \Pi(\mathcal{M}, \mathcal{I}, b). \tag{2}$$

Similarly, fix an action space A and define

$$\Pi^{\text{MAX}-2\text{MIN}}(A) = \sup_{\mathcal{M} \in \mathbf{M}(A)} \inf_{\mathcal{I} \in \mathbf{I}} \inf_{b \in B(\mathcal{M}, \mathcal{I})} \Pi(\mathcal{M}, \mathcal{I}, b). \tag{3}$$

Note that the values of these programs only depend on the cardinality of the action and signal space. In our simulations, reported below, we have generally found that $\Pi^{\text{MIN}-2\text{MAX}}(S) > \Pi^{\text{MAX}-2\text{MIN}}(A)$ when S and A are finite. Only as $|S| \to \infty$ and $|A| \to \infty$ does $\Pi^{\text{MIN}-2\text{MAX}}(S) - \Pi^{\text{MAX}-2\text{MIN}}(A)$ tend to zero. This is elaborated in Theorem 2 below, from which Theorem 1 immediately follows.

To develop this result, we need to define two auxiliary linear programs. For a fixed $k \in \mathbb{N}$, let

$$X(k) = \left\{ \frac{l}{k} \,\middle|\, 0 \leqslant l \leqslant k^2, l \in \mathbb{Z} \right\}^N$$

be the space of actions/signals.

Given a function $f : X(k) \to \mathbb{R}^N$, the discrete upward partial derivative $\nabla_i^+ f(x)$ is defined as[5]

$$\nabla_i^+ f(x) = \mathbb{I}_{x_i < k}(k-1)(f_i(x_i + 1/k, x_{-i}) - f_i(x)).$$

We let $\nabla^+ f(x) = (\nabla_1^+ f(x), \ldots, \nabla_N^+ f(x))$. The discrete upward divergence is $\nabla^+ \cdot f(x) = \sum_{i=1}^N \nabla_i^+ f(x)$. Also, let

$$\rho(x) = \left(1 - \frac{1}{k}\right)^{k \Sigma x} \frac{1}{k^{\sum_{i=1}^N \mathbb{I}_{x_i < k}}}. \tag{4}$$

denote the independent censored geometric distribution on $X(k)$ with arrival rate $1/k$.

Consider the programs

$$\begin{aligned} \overline{\Pi}^{\text{MIN}-2\text{MAX}}(k) = \quad & \min_{\gamma: X(k) \to \mathbb{R}_+,\, \sigma: X(k) \times V \to \mathbb{R}_+,\, w: X(k) \to \mathbb{R}_+^N} \sum_{x \in X(k)} \gamma(x) \\ \text{s.t.} \quad & \gamma(x) \geqslant \rho(x) \left[w_i(x) - \nabla_i^+ w(x) \right] \ \forall x; \\ & \sum_{v \in V} \sigma(x, v) = \rho(x) \ \forall x; \\ & \sum_{x \in X(k)} \sigma(x, v) = \mu(v) \ \forall v; \\ & w(x) = \frac{1}{\rho(x)} \sum_{v \in V} v \sigma(x, v) \ \forall x \end{aligned} \tag{5}$$

[5] Given that the increment between elements in $X(k)$ is $1/k$, a seemingly more natural definition of the discrete derivative would have a factor k rather than $k-1$. Of course, these definitions are equivalent in the limit as k tends to ∞, and by defining it with $k-1$, we simplify several calculations in the proof of Theorem 2.

and

$$
\begin{aligned}
\underline{\Pi}^{\text{MAX}-2\text{MIN}}(k) = \quad & \max_{\lambda:V\to\mathbb{R},\, q:X(k)\to\mathbb{R}^N_+,\, t:X(k)\to\mathbb{R}^N} \sum_{v\in V} \mu(v)\lambda(v) \\
\text{s.t. } & \lambda(v) \leqslant \Sigma t(x) + v\cdot\nabla^+ q(x) - \nabla^+\cdot t(x)\ \forall v, x; \\
& \Sigma q(x) \leqslant 1\ \forall x; \\
& t_i(0, x_{-i}) = 0\ \forall i, x_{-i}.
\end{aligned} \tag{6}
$$

Theorem 1 follows immediately from our second main result, which is the following:

Theorem 2 (Bounds)**.** *For all k,*

$$
\overline{\Pi}^{\text{MIN}-2\text{MAX}}(k) \geqslant \Pi^{\text{MIN}-2\text{MAX}}(X(k)) \geqslant \Pi^{\text{MAX}-2\text{MIN}}(X(k)) \geqslant \underline{\Pi}^{\text{MAX}-2\text{MIN}}(k). \tag{7}
$$

Moreover, any optimal solution of problem (5) *yields an information structure such that the maximum profit across mechanisms and equilibria is at most $\overline{\Pi}^{\text{MIN}-2\text{MAX}}(k)$, and any optimal solution of problem* (6) *yields a mechanism such that the minimum profit across information structure and equilibria is at least $\underline{\Pi}^{\text{MAX}-2\text{MIN}}(k)$. Finally, if $\mu(v) > 0$ for all $v \in V$, then*

$$
\lim_{k\to\infty} \overline{\Pi}^{\text{MIN}-2\text{MAX}}(k) = \lim_{k\to\infty} \underline{\Pi}^{\text{MAX}-2\text{MIN}}(k) = \Pi^*,
$$

where Π^* is as given in Theorem 1.

Thus, the finite linear programs (5) and (6) bound (2) and (3), and under the full-support hypothesis, their asymptotic values are equal to the profit guarantee.

3.2 Proof of Theorems 1 and 2

We now prove Theorems 1 and 2. The first step of the proof is showing the inequalities (7). We then show that $\overline{\Pi}^{\text{MIN}-2\text{MAX}}(k) - \underline{\Pi}^{\text{MAX}-2\text{MIN}}(k)$ converges to zero under the full support hypothesis.

3.2.1 Ordering $\Pi^{\text{MIN}-2\text{MAX}}$ and $\Pi^{\text{MAX}-2\text{MIN}}$

An elementary observation is that $\Pi^{\text{MIN}-2\text{MAX}}$ is always greater than $\Pi^{\text{MAX}-2\text{MIN}}$, and these values move closer as the the number of actions and signals increases.

Lemma 1. *For all S and A, $\Pi^{\text{MIN}-2\text{MAX}}(S) \geqslant \Pi^{\text{MAX}-2\text{MIN}}(A)$. Moreover, if $|A_i| \leqslant |A'_i|$ (respectively $|S_i| \leqslant |S'_i|$) for all i, then $\Pi^{\text{MAX}-2\text{MIN}}(A) \leqslant \Pi^{\text{MAX}-2\text{MIN}}(A')$ (respectively $\Pi^{\text{MIN}-2\text{MAX}}(S) \geqslant \Pi^{\text{MIN}-2\text{MAX}}(S')$).*

Proof of Lemma 1. The first part follows from a standard argument in zero-sum games, adapted to the setting where the second mover also chooses the equilibrium. Fix an $\epsilon > 0$, let $\mathcal{M} \in \mathbf{M}(A)$ be a mechanism such that the infimum profit across information structures and equilibria is at least $\Pi^{\text{MAX}-2\text{MIN}}(A) - \epsilon$. Let $\mathcal{I} \in \mathbf{I}(S)$ be an information structure such that the supremum profit across mechanisms and equilibria is at most $\Pi^{\text{MIN}-2\text{MAX}}(S)+\epsilon$. Thus, for any equilibrium $b \in B(\mathcal{M},\mathcal{I})$, $\Pi^{\text{MIN}-2\text{MAX}}(S)+\epsilon \geqslant \Pi(\mathcal{M},\mathcal{I},b) \geqslant \Pi^{\text{MAX}-2\text{MIN}}(A)-\epsilon$, thus showing that $\Pi^{\text{MIN}-2\text{MAX}}(S) \geqslant \Pi^{\text{MAX}-2\text{MIN}}(A)-2\epsilon$. Since ϵ is arbitrary, we conclude that $\Pi^{\text{MIN}-2\text{MAX}}(S) \geqslant \Pi^{\text{MAX}-2\text{MIN}}(A)$.

We now prove the second part. For $\epsilon > 0$, let $\mathcal{M} = (A, q, t)$ be a mechanism such that infimum profit across information structures and equilibria is at least $\Pi^{\text{MAX}-2\text{MIN}}(A) - \epsilon$. Since $|A'_i| > |A_i|$, there exists an onto mapping $f_i : A'_i \to A_i$ for each i. Let $f : A' \to A$ be the product mapping. We define $q'(a') = q(f(a'))$ and $t'(a') = t(f(a'))$, and let $\mathcal{M}' = (A', q', t')$. It is clear that for every $\mathcal{I}$ and Π, there is an equilibrium of $(\mathcal{M},\mathcal{I})$ with profit Π if and only if there is an equilibrium of $(\mathcal{M}',\mathcal{I})$ with profit Π: For given the former, we can construct a profit-equivalent equilibrium of the latter by selecting a single action in $f_i^{-1}(a_i)$ to be played instead of a_i, and given the latter, we can construct a profit equivalent equilibrium in which the action $f_i(a'_i)$ is played instead of a'_i. Thus, $\Pi^{\text{MAX}-2\text{MIN}}(A') \geqslant \Pi^{\text{MAX}-2\text{MIN}}(A)-\epsilon$, and since ϵ is arbitrary, we have $\Pi^{\text{MAX}-2\text{MIN}}(A') \geqslant \Pi^{\text{MAX}-2\text{MIN}}(A)$. The proof for $\Pi^{\text{MIN}-2\text{MAX}}$ is analogous and is omitted. □

3.2.2 Local relaxations

Lemma 2. *For all $k \in \mathbb{N}$, $\overline{\Pi}^{\text{MIN}-2\text{MAX}}(k) \geqslant \Pi^{\text{MIN}-2\text{MAX}}(X(k))$. Moreover, if $(\gamma^*, \sigma^*, w^*)$ is an optimal solution to* (5)*, then $(X(k), \sigma^*)$ is an information structure for which profit in any mechanisms and equilibrium is at most $\overline{\Pi}^{\text{MIN}-2\text{MAX}}(k)$.*

Proof of Lemma 2. Consider the inner maximization program in (2) for a fixed information structure $\mathcal{I} = (X(k), \sigma)$ in $\mathbf{I}(X(k))$, in which we maximize over all participation-secure mechanisms and equilibria. The presence of the participation-security action implies that all bidders must receive non-negative utility in equilibrium. Thus, we can relax the program by dropping the requirement of participation security, and replacing it with the constraint that equilibrium interim bidder surpluses must be non-negative.

By the revelation principle (Myerson, 1981), this relaxed program is equivalent to maximizing profit over incentive compatible and individually rational direct mechanisms. Recall that a *direct mechanism* on the information structure $\mathcal{I}$ is a mechanism with $A = X(k)$. When the action and signal spaces coincide, we let $\overline{b}_i$ denote the *truthful strategies* that place probability one on $a_i = s_i$ for all i. The direct mechanism is *incentive compatible* if $\overline{b}$ is an equilibrium. It is *individually rational* if the truthful strategies give each bidder a non-negative payoff in the interim stage. Thus, the relaxed program is the finite linear program

$$
\begin{aligned}
&\max_{q:S(k)\to\mathbb{R}_+^N,\, t:X(k)\to\mathbb{R}^N} \sum_{x\in X(k)}\sum_{v\in V}\Sigma t(x)\sigma(x,v)\\
\text{s.t. } &\sum_{v\in V}\sum_{x_{-i}\in X_{-i}(k)}(v_i q_i(x_i,x_{-i})-t_i(x_i,x_{-i}))\sigma(x_i,x_{-i},v)\\
&\qquad \geqslant \sum_{v\in V}\sum_{x_{-i}\in X_{-i}(k)}(v_i q_i(x_i',x_{-i})-t_i(x_i',x_{-i}))\sigma(x_i,x_{-i},v)\ \forall i,x_i,x_i';\\
&\sum_{v\in V}\sum_{x_{-i}\in X_{-i}(k)}(v_i q_i(x_i,x_{-i})-t_i(x_i,x_{-i}))\sigma(x_i,x_{-i},v)\geqslant 0\ \forall i,x_i;\\
&\Sigma q(x)\leqslant 1\ \forall x.
\end{aligned}
$$

The value of this program is equal to that of its dual:

$$
\min_{\{\alpha_i:X_i(k)^2\to\mathbb{R}_+,\,\beta_i:X_i(k)\to\mathbb{R}_+\}_{i=1}^N,\,\gamma:X(k)\to\mathbb{R}_+}\ \sum_{x\in X(k)}\gamma(x) \tag{8}
$$

$$
\begin{aligned}
\text{s.t. } \gamma(x)\geqslant &\sum_{x_i'\in X_i(k)}\sum_{v\in V}v_i[\sigma(x,v)\alpha_i(x_i,x_i')-\sigma(x_i',x_{-i},v)\alpha_i(x_i',x_i)]\\
&+\sum_{v\in V}\beta_i(x_i)v_i\sigma(x,v)\ \forall i,x;
\end{aligned} \tag{8a}
$$

$$
\begin{aligned}
\sum_{v\in V}\sigma(x,v)= &\sum_{x_i'\in X_i(k)}\sum_{v\in V}[\sigma(x,v)\alpha_i(x_i,x_i')-\sigma(x_i',x_{-i},v)\alpha_i(x_i',x_i)]\\
&+\sum_{v\in V}\beta_i(x_i)\sigma(x,v)\ \forall i,x,
\end{aligned} \tag{8b}
$$

where $\alpha_i(x_i,x_i')$ is the multiplier on the incentive compatibility constraint for type x_i not misreporting as x_i' and $\beta_i(x_i)$ is the multiplier on the individual rationality constraint for type x_i. Thus, (2) has a value less than or equal to the (non-linear) program of minimizing the value of (8) across all information structures σ and multipliers (α,β,γ).

Note that the value of the inner program (8) will only increase if we hold α and β fixed at particular values. In particular, consider the following values:

$$
\alpha_i(x_i,x_i')=\begin{cases}1 & \text{if } x_i'+\frac{1}{k}=x_i=k;\\ k & \text{if } x_i'+\frac{1}{k}=x_i<k;\\ 0 & \text{otherwise}\end{cases} \tag{9}
$$

and

$$
\beta_i(x_i)=\begin{cases}k & \text{if } x_i=0;\\ 0 & \text{otherwise.}\end{cases} \tag{10}
$$

The constraint (8b) (for which $t_i(x)$ is the multiplier) can be simplified as follows: In a slight abuse of notation, let $\sigma(x)$ denote the marginal of σ on $X(k)$. Then integrating out values and using the particular multipliers, (8b) becomes

$$\sigma(x) = \begin{cases} \frac{k-1}{k}\sigma(x_i - 1/k, x_{-i}) & \text{if } 0 < x_i < k; \\ (k-1)\sigma(k-1/k, x_{-i}) & \text{if } x_i = k. \end{cases} \tag{11}$$

The unique probability distribution satisfying this equation is $\sigma(x) = \rho(x)$ defined by (4). As a result, we can replace (8b) with the constraint

$$\sum_{v \in V} \sigma(x, v) = \rho(x). \tag{12}$$

Thus, the multipliers in (9) and (10) are feasible for problem (8) if σ's marginal distribution on $X(k)$ is ρ.

In addition, substituting the chosen multipliers into (8a), the constraint becomes

$$\gamma(x) \geqslant \begin{cases} k\sum_{v\in V} v_i[\sigma(x,v) - \sigma(x_i + 1/k, x_{-i}, v)] & \text{if } x_i < k - 1/k; \\ \sum_{v\in V} v_i[k\sigma(x,v) - \sigma(x_i + 1/k, x_{-i}, v)] & \text{if } x_i = k - 1/k; \\ \sum_{v\in V} v_i\sigma(x,v) & \text{if } x_i = k. \end{cases} \tag{13}$$

Letting

$$w(x) = \frac{1}{\rho(x)} \sum_{v\in V} v\sigma(x, v) \tag{14}$$

denote the interim expected value of bidder i conditional on the signal profile x, the constraint (13) can be rewritten as

$$\gamma(x) \geqslant \rho(x)\left[w_i(x) - \nabla_i^+ w(x)\right]. \tag{15}$$

Thus, replacing (8b) with (12) and replacing (8a) with (14) and (15) yields a program with weakly higher value than (2). This relaxed program is precisely (5). □

Lemma 3. *For all $k \geqslant 0$, $\underline{\Pi}^{\text{MAX}-2\text{MIN}}(k) \leqslant \Pi^{\text{MAX}-2\text{MIN}}(X(k))$. Moreover, if (λ^*, q^*, t^*) is an optimal solution of problem (6), then $(X(k), q^*, t^*)$ is a mechanism for which profit in any information structure and equilibrium is at least $\underline{\Pi}^{\text{MAX}-2\text{MIN}}(k)$.*

Proof of Lemma 3. Consider the inner minimization program in (3) for a fixed mechanism $\mathcal{M} = (X(k), q, t)$ in $\mathbf{M}(X(k))$. The program of minimizing expected profit over all information structures and equilibria can be reformulated as a finite linear program. Specifically, a *Bayes correlated equilibrium* (BCE) of $\mathcal{M}$ is an information structure with $S = X(k)$ such that the truthful strategies are an equilibrium. The problem of minimizing expected profit over information structures and equilibria is equivalent to minimizing expected profit over

BCE (Bergemann and Morris, 2013, 2016). Explicitly, this program is

$$\begin{aligned}
&\min_{\sigma:X(k)\times V\to\mathbb{R}_+} \sum_{v\in V}\sum_{x\in X(k)} \Sigma t(x)\sigma(x,v)\\
\text{s.t. } &\sum_{v\in V}\sum_{x_{-i}\in X_{-i}(k)} \left[v_i q_i(x_i,x_{-i}) - t_i(x_i,x_{-i})\right]\sigma(x_i,x_{-i},v)\\
&\qquad \geqslant \sum_{v\in V}\sum_{x_{-i}\in X_{-i}(k)} \left[v_i q_i(x_i',x_{-i}) - t_i(x_i',x_{-i})\right]\sigma(x_i,x_{-i},v)\ \forall i, x_i, x_i';\\
&\sum_{x\in X(k)} \sigma(x,v) = \mu(v)\ \forall v.
\end{aligned}$$

The value of this program is equal to that of its dual:

$$\max_{\{\alpha_i:(X_i(k))^2\to\mathbb{R}_+\}_{i=1}^N,\,\lambda:V\to\mathbb{R}} \sum_{v\in V}\lambda(v)\mu(v) \tag{16}$$

$$\begin{aligned}
\text{s.t. } \lambda(v) \leqslant \Sigma t(x) + \sum_{i=1}^N \sum_{x_i'\in X_i(k)} \alpha_i(x_i,x_i')\big[&(v_i q_i(x_i',x_{-i}) - t_i(x_i',x_{-i}))\\
&- (v_i q_i(x_i,x_{-i}) - t_i(x_i,x_{-i}))\big]\ \forall x, v,
\end{aligned}$$

where $\alpha(x_i, x_i')$ is the multiplier on the obedience constraint that a bidder with signal x_i not want to bid x_i', and $\lambda(v)$ is the multiplier on the constraint that the marginal probability of $v \in V$ under σ is $\mu(v)$. Moreover, any feasible solution to the dual is a lower bound on the value of the primal. In particular, consider the following feasible multipliers:

$$\alpha_i(x_i,x_i') = \begin{cases} k-1 & \text{if } x_i' - \frac{1}{k} = x_i;\\ 0 & \text{otherwise.}\end{cases} \tag{17}$$

In this case, the dual constraint becomes

$$\lambda(v) \leqslant \Sigma t(x) + v\cdot\nabla^+ q(x) - \nabla^+\cdot t(x). \tag{18}$$

Thus, the maximum of (16) subject to (18) is a lower bound on the inner minimization program in (3). As a result, the maximum of this lower bound across all participation secure mechanisms, given by the linear program (6), is a lower bound on the value of (3). □

3.2.3 Convergence

Program (5) has the following dual:[6]

$$\max_{\lambda:V\to\mathbb{R},\,\Xi:X(k)\to\mathbb{R},\,q:X(k)\to\mathbb{R}^N_+} \sum_{x\in X(k)} \rho(x)\Xi(x) + \sum_{v\in V} \mu(v)\lambda(v)$$

$$\text{s.t. } \Xi(x) + \lambda(v) \leqslant v \cdot \nabla^- q(x)\ \forall v, x;$$

$$\nabla_i^- q(x) = \begin{cases} kq_i(x) & x_i = 0; \\ k(q_i(x) - q_i(x_i - 1/k, x_{-i})) & 0 < x_i < k; \quad \forall i, x \\ q_i(x) - q_i(x_i - 1/k, x_{-i}) & x_i = k; \end{cases} \tag{19}$$

$$\Sigma q(x) \leqslant 1\ \forall x.$$

Lemma 4. *Suppose $\mu(v) > 0$ for every $v \in V$. Let $\epsilon(k) > \frac{C}{k}$ where C is any constant bigger than $\frac{2\overline{v}/\mu(v)}{v_i''-v_i'}$ for all i, $v \in V$, and $v_i', v_i'' \in V_i$ such that $v_i' \neq v_i''$. Then there exists an optimal solution (λ^*, Ξ^*, q^*) of (19) that satisfies*

$$q_i^*(x_i - 1/k, x_{-i}) \leqslant q_i^*(x) + \epsilon(k) \tag{20}$$

for every i and every x such that $0 < x_i < k$.

We defer the proof of Lemma 4 to the appendix. The argument consists of two main steps. First, we argue that optimal λ can be taken to be bounded as $k \to \infty$. The optimal λ and Ξ in (19) are only defined up to a constant, so it is without loss to normalize them so that the expected value of Ξ is zero and

$$\sum_{v\in V} \mu(v)\lambda(v) = \overline{\Pi}^{\text{MIN}-2\text{MAX}}(k).$$

Now, consider the variation of (19) where we hold fixed some particular λ. The dual of this linear program is similar to (5), except that instead of the constraint on the marginal of σ on v, there is an extra term in the objective, which is now

$$\overline{\Pi}^{\text{MIN}-2\text{MAX}}(k) + \sum_{x\in X(k)} \gamma(x) - \sum_{v\in V} \lambda(v) \sum_{x\in X(k)} \sigma(x, v), \tag{21}$$

which is effectively a linear cost of a value distribution. Now, if $\lambda(v)$ were larger than $\overline{v}$, then a feasible solution to this program is to put probability one on this v. As a result, the second two terms in (21) would be negative, which contradicts the optimal value being $\overline{\Pi}^{\text{MIN}-2\text{MAX}}(k)$. Finally, since λ is bounded above and the value of (19) is non-negative, the full-support hypothesis implies that λ is bounded below, uniformly across k. This is the only place in the proof of Theorems 1 and 2 where we use the hypothesis that μ has full support.

[6] When taking the dual of (5) we find it convenient to replace the constraint (15) by the equivalent constraint (13).

The second step is to show that for k sufficiently large, the optimal value of (19) does not change if we impose (20) for all i and x as additional constraints. This argument is actually made through the dual of the constrained program where we hold fixed an optimal λ. Again, this program looks much like (5), except with the modified objective (21) and no marginal constraint, and there are now extra variables which correspond to multipliers on (20). When k is large, and using the hypothesis that λ is bounded, we use a perturbation argument to show that the optimal multipliers on the constraints (20) must all be zero.

The following lemma is the last piece of the argument that the programs (5) and (6) have the same asymptotic value as k goes to infinity:

Lemma 5. *Suppose $\mu(v) > 0$ for every $v \in V$. Then*

$$\liminf_{k\to\infty} \left(\underline{\Pi}^{\text{MAX}-2\text{MIN}}(k) - \overline{\Pi}^{\text{MIN}-2\text{MAX}}(k)\right) \geqslant 0.$$

Proof of Lemma 5. To prove the lemma, it is convenient to reformulate program (6) so that it closely resembles program (19).

First, the dual of (6) is

$$\min_{\gamma:X(k)\to\mathbb{R}_+,\, \sigma:X(k)\times V\to\mathbb{R}_+} \sum_{x\in X(k)} \gamma(x) \tag{22}$$

$$\text{s.t. } \gamma(x) \geqslant (k-1)\sum_{v\in V} v_i(\sigma(x_i - 1/k, x_{-i}, v)\mathbb{I}_{x_i>0} - \sigma(x,v)\mathbb{I}_{x_i<k}) \;\forall i, x; \tag{22a}$$

$$\sum_v (\sigma(x,v) - (k-1)(\sigma(x_i - 1/k, x_{-i}, v)\mathbb{I}_{x_i>0} - \sigma(x,v)\mathbb{I}_{x_i<k})) = 0 \quad \forall i, x \text{ such that } x_i > 0; \tag{22b}$$

$$\sum_{x\in X(k)} \sigma(x,v) = \mu(v) \;\forall v, \tag{22c}$$

where $\sigma(x,v)$ is the multiplier on the first constraint in (6), and $\gamma(x)$ is the multiplier on the second constraint (feasibility constraint for q).

Second, exactly as we solve the marginal of σ on $X(k)$ in program (8) (see equation (11)), we have picked the multipliers in (17) so that constraint (22b) is equivalent to

$$\sum_{v\in V} \sigma(x,v) = \rho(x) \;\forall x. \tag{23}$$

This follows from an analogous calculation as in the proof of Lemma 2. Finally, after replacing constraint (22b) with (23) in program (22), the dual of (22) becomes:

$$\max_{\lambda:V\to\mathbb{R},\, \Xi:X(k)\to\mathbb{R},\, q:X(k)\to\mathbb{R}^N_+} \sum_{x\in X(k)} \rho(x)\Xi(x) + \sum_{v\in V}\mu(v)\lambda(v) \tag{24}$$

$$\text{s.t. } \Xi(x) + \lambda(v) \leqslant v\cdot\nabla^+ q(x) \;\forall v, x; \tag{24a}$$

$$\Sigma q(x) \leqslant 1 \;\forall x, \tag{24b}$$

where $q_i(x)$ is the multiplier on constraint (22a), $\lambda(v)$ is the multiplier on constraint (22c), and $\Xi(x)$ is the multiplier on constraint (23). By construction, the programs (6) and (24) have the same optimal value of $\underline{\Pi}^{\text{MAX}-2\text{MIN}}(k)$.

Let $\epsilon(k) = \frac{C+\epsilon}{k}$ where C is given in the statement of Lemma 4 and $\epsilon > 0$ is arbitrary. By Lemma 4, let (λ^*, Ξ^*, q^*) be an optimal solution of (19) that satisfies $q_i^*(x_i - 1/k, x_{-i}) \leqslant q_i^*(x) + \epsilon(k)$ for every i and every x such that $0 < x_i < k$.

Define

$$\begin{aligned}
\overline{q}_i(x) &= \begin{cases} \frac{q_i^*(x_i-1/k, x_{-i})}{1+N\epsilon(k)} & \text{if } 0 < x_i < k; \\ 0 & \text{if } x_i = 0 \text{ or } x_i = k; \end{cases} \\
\overline{\lambda}(v) &= \frac{k-1}{k(1+N\epsilon(k))}\lambda^*(v) \quad \forall v \in V; \\
\overline{\Xi}(x) &= \begin{cases} \frac{k-1}{k(1+N\epsilon(k))}\Xi^*(x) & \text{if } x \notin \partial X(k); \\ -(k-1)N\overline{v} - \max_{v\in V}\overline{\lambda}(v) & \text{if } x \in \partial X(k), \end{cases}
\end{aligned} \tag{25}$$

where $\partial X(k) = \{x \in X(k) \mid x_i \geqslant k - 1/k \text{ for some } i\}$.

We claim that $(\overline{\lambda}, \overline{\Xi}, \overline{q})$ is feasible for the program (24): First, the constraint (24a) holds for $x \in \partial X(k)$ because

$$\overline{\Xi}(x) = -(k-1)N\overline{v} - \max_{v\in V}\overline{\lambda}(v) \leqslant v \cdot \nabla^+ \overline{q}(x) - \overline{\lambda}(v)$$

for all v; (24a) also holds for $x \notin \partial X(k)$ because $\nabla^+\overline{q}(x) = \frac{k-1}{k(1+N\epsilon(k))}\nabla^- q^*(x)$, $\overline{\Xi}(x) = \frac{k-1}{k(1+N\epsilon(k))}\Xi^*(x)$, $\overline{\lambda}(v) = \frac{k-1}{k(1+N\epsilon(k))}\lambda^*(v)$, and $\Xi^*(x)+\lambda^*(v) \leqslant v \cdot \nabla^- q^*(x)$. Also, the feasibility constraint (24b) is satisfied, as

$$\sum_{i=1}^N \overline{q}_i(x) = \sum_{i=1}^N \frac{q_i^*(x_i - 1/k, x_{-i})}{1+N\epsilon(k)}\mathbb{I}_{0<x_i<k} \leqslant \sum_{i=1}^N \frac{q_i^*(x) + \epsilon(k)}{1+N\epsilon(k)} \leqslant 1.$$

Finally, the difference in objectives of (19) under (λ^*, Ξ^*, q^*) (which is equal to $\overline{\Pi}^{\text{MIN}-2\text{MAX}}(k)$) and of (24) under $(\overline{\lambda}, \overline{\Xi}, \overline{q})$ is

$$\begin{aligned}
&\sum_{x\in X(k)} \rho(x)(\Xi^*(x) - \overline{\Xi}(x)) + \sum_{v\in V}\mu(v)(\lambda^*(v) - \overline{\lambda}(v)) \\
&= \sum_{x\in X(k)} \rho(x)\left(1 - \frac{k-1}{k(1+N\epsilon(k))}\right)\Xi^*(x) + \sum_{x\in\partial X(k)} \rho(x)\left(\frac{k-1}{k(1+N\epsilon(k))}\Xi^*(x) - \overline{\Xi}(x)\right) \\
&\quad + \sum_{v\in V}\mu(v)\left(1 - \frac{k-1}{k(1+N\epsilon(k))}\right)\lambda^*(v) \\
&= \left(1 - \frac{k-1}{k(1+N\epsilon(k))}\right)\overline{\Pi}^{\text{MIN}-2\text{MAX}}(k) \\
&\quad + \sum_{x\in\partial X(k)}\rho(x)\left(\frac{k-1}{k(1+N\epsilon(k))}\Xi^*(x) + (k-1)N\overline{v} + \max_{v\in V}\frac{k-1}{k(1+N\epsilon(k))}\lambda^*(v)\right) \\
&\leqslant \left(1 - \frac{k-1}{k(1+N\epsilon(k))}\right)\overline{\Pi}^{\text{MIN}-2\text{MAX}}(k) + N(1-1/k)^{k^2-1}\left(\frac{k-1}{k(1+N\epsilon(k))}kN\overline{v} + (k-1)N\overline{v}\right),
\end{aligned}$$

where in the last line we use the fact that $\rho(\partial X(k)) \leqslant N(1-1/k)^{k^2-1}$ and $\Xi^*(x) + \lambda^*(v) \leqslant v \cdot \nabla^- q^*(x) \leqslant kN\overline{v}$. The last line of the display equation vanishes as $k \to \infty$ because $\epsilon(k) \to 0$ and $\overline{\Pi}^{\text{MIN}-2\text{MAX}}(k) \leqslant \overline{v}$ for all k (since $\sigma(x,v) = \rho(x)\mu(v)$ is feasible for the program (5)). This implies the result, since $\underline{\Pi}^{\text{MAX}-2\text{MIN}}(k)$ is equal to the value of the program (24), which is weakly larger than the objective obtained by $(\overline{\lambda}, \overline{\Xi}, \overline{q})$. □

Proof of Theorem 2. From Lemma 2, we know that $\Pi^{\text{MIN}-2\text{MAX}}(X(k)) \leqslant \overline{\Pi}^{\text{MIN}-2\text{MAX}}(k)$ for all k, and Lemma 3 implies $\underline{\Pi}^{\text{MAX}-2\text{MIN}}(k) \leqslant \Pi^{\text{MAX}-2\text{MIN}}(X(k))$ for all k. Lemma 1 implies that $\lim_{k\to\infty} \Pi^{\text{MIN}-2\text{MAX}}(X(k))$ and $\lim_{k\to\infty} \Pi^{\text{MAX}-2\text{MIN}}(X(k))$ exist. Putting these together and applying Lemma 5, we conclude that

$$\begin{aligned}\lim_{k\to\infty} \overline{\Pi}^{\text{MIN}-2\text{MAX}}(k) &= \lim_{k\to\infty} \Pi^{\text{MIN}-2\text{MAX}}(X(k)) \\ &= \lim_{k\to\infty} \Pi^{\text{MAX}-2\text{MIN}}(X(k)) = \lim_{k\to\infty} \underline{\Pi}^{\text{MAX}-2\text{MIN}}(k).\end{aligned}$$

□

Proof of Theorem 1. The left-hand side of (1) is equal to the inf over all finite signal spaces S of $\Pi^{\text{MIN}-2\text{MAX}}(S)$. By Lemma 1, $\Pi^{\text{MIN}-2\text{MAX}}(X(k))$ is a decreasing sequence, so that the left-hand side of (1) is less than or equal to the limit as k goes to infinity of $\Pi^{\text{MIN}-2\text{MAX}}(k)$, which by Theorem 2 is Π^*. A similar argument shows that the right-hand side of (1) is at least Π^* as well. The theorem then follows from Lemma 1, which implies that the left-hand side of (1) is greater than the right-hand side. □

3.3 Discussion

Approximate dual pair The proof of Theorems 1 and 2 essentially shows that the values of the programs (2) and (3) are completely determined by local incentive constraints, and that they are asymptotically equivalent to their respective local relaxations, the programs (5) and (6). At first glance, these programs appear unrelated: in the former we minimize over information structures, and the latter we maximize over mechanisms. In fact, these programs are closely related to one another and are "almost" a dual pair in the following sense. The proof of Lemma 5 shows that transfers can be "solved out" of (6) to obtain the program (24), where we just maximize over the allocation. This program is nearly identical to the dual of (5), program (19), with one exception: In (24), it is local *upward* incentive constraints that are binding, whereas in (19), it is local *downward* constraints that bind.[7] As a result, it is the discrete downward derivative of the allocation that appears in (19), whereas the discrete upward derivative is in (24).

Notice that there are analogous similarities and differences between the program (5) and the dual of (6) (or more specifically its reformulation (24)). Both involve minimization over information structures whose marginal on $X(k)$ is ρ.[8] And in both cases, the variable

[7]This difference in the pattern of binding local incentive constraints is explicit in the proofs of Lemmas 2 and 3.

[8]The marginal constraint on $X(k)$ is explicit in (5). The dual of (24) is briefly described in the proof of Lemma 5, and is the same as the program (22) except that the constraint (22b) is replaced with the marginal constraint on $X(k)$.

$\gamma(x)$ represents a highest "virtual value" of the bidders. The key difference is that in (5), the virtual value, written in terms of the interim value, is

$$w_i(x) - (k-1)(w_i(x_i + 1/k, x_{-i}) - w_i(x))$$

whereas in the dual of (6), the virtual value is

$$w_i(x) - k(w_i(x) - w_i(x_i - /k, x_{-i})).$$

Thus, the key difference between these expressions (when k is large) is whether local upward or local downward constraints are used to compute information rents.

Lemma 5 shows that this difference is inconsequential when k is large, and whether local constraints go up or down, we arrive at the same asymptotic value for the local relaxation. An important unanswered question is whether these programs can be formulated as an exact dual pair in the continuum limit. If so, there are important implications for the structure of the solution, which we discuss in the conclusion.

One-dimensional action/signal space The common signal/action space $X(k)$ in the programs (5) and (6) is one dimensional in the following senses: First, (5) is derived from (2) by dropping all non-local incentive-compatibility constraints; and likewise for (6) and (3). In effect, the proof shows that max-2min mechanisms and min-2max information structures are determined by local incentive compatibility with respect to *some* order, and we have simply labeled actions and signals in $X(k)$ so that the order coincides with the standard order on subsets of $\mathbb{R}$.

Second, Lemma 4 shows that it is without loss to restrict attention to solutions to (6) and (19) in which the allocation does not jump down by more than C/k for a constant C. Thus, in the $k \to \infty$ limit, the subgradient of the allocation q_i is bounded below. In fact, while upward jumps are not ruled out by our argument, simulations reported below suggest that it is without loss for the allocation to be continuous. This is an intuitive feature of max-2min auctions: small changes in messages result in small changes in allocations, which makes the mechanism less susceptible to manipulation through information.

Based on simulations, we conjecture that the order on signals and actions plays an additional role: In every instance where we solved(6) and (19), we have found that there exists an optimal solution in which $q_i(x)$ is non-decreasing in x_i. We have selected such solutions in the simulations we report throughout this paper. The conjecture that non-decreasing q_i is without loss of generality has interesting implications, which we discuss further in Section 5.

Independent signals It is intuitive that the min-2max information structure that solves (2) would have independent signals, as this rules out Crémer and McLean (1988) style constructions, where bidder surplus is extracted by having the bidders make bets about others' types. In the extreme case where the signal distribution satisfies the Crémer and McLean separation conditions and full surplus extraction is possible, bidders' incentive constraints are all slack and only participation constraints bind. In contrast, in the classic revenue equivalence results of Myerson (1981) with independent signals, all local incentive constraints bind.

Remarkably, the proof of Lemma 2 shows that binding local incentive compatibility is actually *equivalent* to independence of the signals. In particular, we derived the independent marginal ρ from the hypothesis that a rich class of local incentive constraints were binding in the local relaxation of (2). Lemma 2 also implicitly uses the binding local incentive compatibility constraint to solve out the interim transfer and represent the profit as an expected virtual value of the winner (Myerson, 1981).

Note that the values of multipliers on the local incentive compatibility constraint in Lemma 2 and also the particular independent signal distribution are merely a normalization. This can be seen in program (8): Suppose that only local downward incentive constraints bind, meaning compatibility constraint, that is, $\alpha_i(x_i, x_i') > 0$ if and only if $x_i' + 1/k = x_i$. Suppose also that individual rationality constraints are all slack, so that $\beta_i(x_i) = 0$. Then letting $\sigma_i(x_i)$ denote the marginal of bidder i's signal, the constraint (8b) reduces to

$$\sigma_i(x_i) = \sigma_i(x_i)\alpha_i(x_i, x_i - 1/k)\mathbb{I}_{x_i>0} - \sigma(x_i + 1/k)\alpha_i(x_i + 1/k, x_i)\mathbb{I}_{x_i<k} \ \forall i, x,$$

where we have integrated out v and x_{-i}. Summing the above equation across $x_i \geqslant x_i'$ yields

$$\sum_{x_i \geqslant x_i'} \sigma(x_i) = \sigma(x_i')\alpha_i(x_i', x_i' - 1/k).$$

In other words, the multiplier on the local incentive compatibility constraint is exactly equal to the inverse hazard rate of the signal distribution.

3.4 Examples

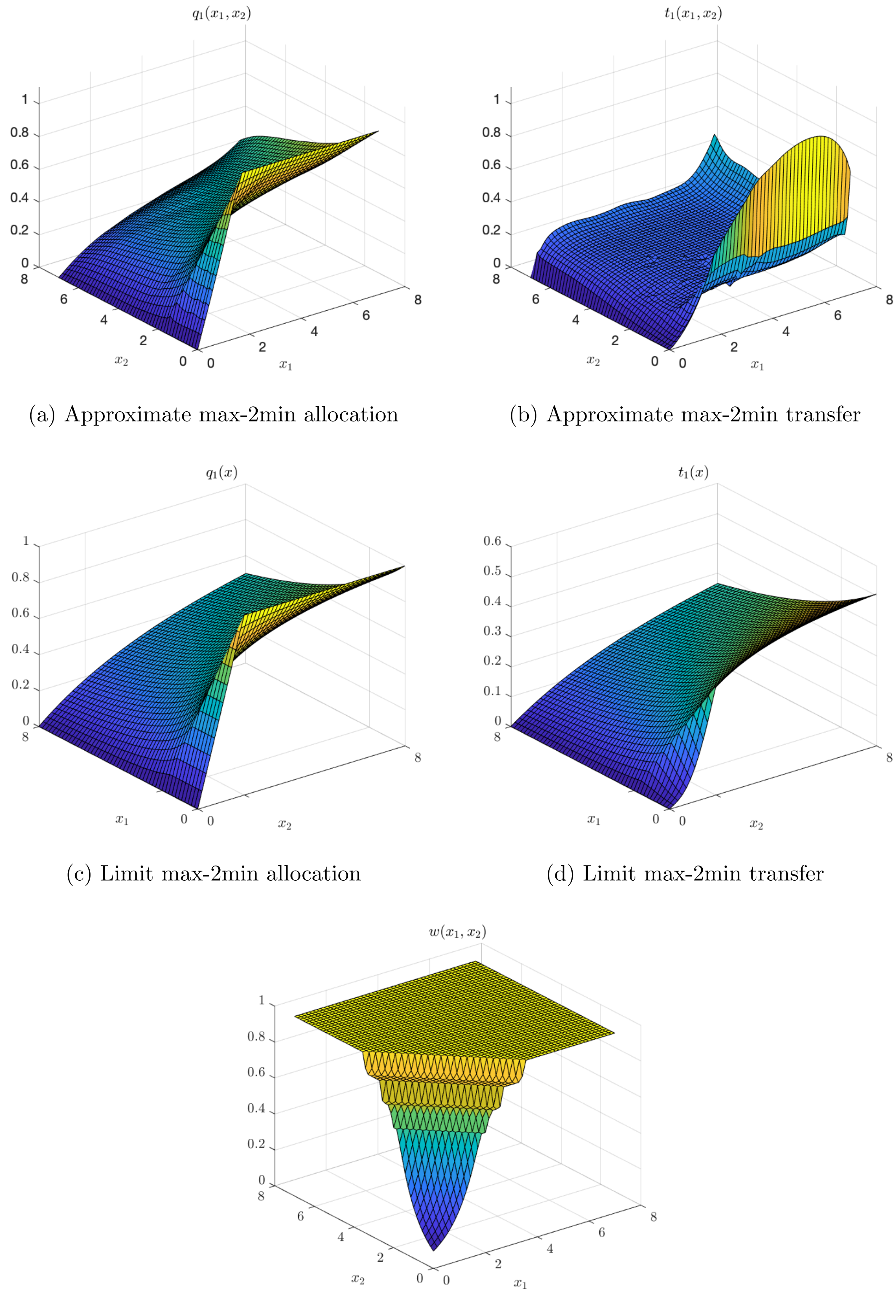
3.4.1 Perfectly correlated values

We now give examples of the approximate min-2max information structures and max-2min mechanisms, which solve (5) and (6), respectively.[9] Our first example is one in which $v_1 = v_2 + c$ for a constant c, i.e., values are perfectly positively correlated. Bidder 2's value v_2 is uniformly distributed on an evenly spaced grid of 10 values between 0 and 1.

Note that this example does not satisfy the full-support hypothesis, so that the asymptotic equivalence of (2) and (3) is not implied by Theorem 1. When values are pure common, however, the equivalence of these programs is a result of the constructive argument in Theorems 3 and 4 of Brooks and Du (2020). And the first part of Theorem 2, that the local relaxations provide profit guarantees, does not depend on μ having full support. In fact, we conjecture that Theorems 1 and 2 remain true in their entirety even without full support.

We now proceed with the discussion of simulations. When $c = 0$, this example reduces to the model of Brooks and Du (2020). In that paper, we presented max-2min mechanisms and min-2max information structures in the limit when the action/signal space is all of $\mathbb{R}_+$. The mechanism has the form of a "proportional auction," in which the aggregate allocation

[9] All computations were performed using Gurobi on a 2019 Macbook Pro with a 2.8 GHz Quad-Core Intel Core i7 processor. Unless otherwise stated, we used the interior point (barrier) algorithm. Each calculation took approximately 5 seconds.



(a) Approximate max-2min allocation

(b) Approximate max-2min transfer

(c) Limit max-2min allocation

(d) Limit max-2min transfer

(e) Approximate min-2max information

Figure 1: Max-2min mechanisms and min-2max information with pure common values.

and aggregate transfer only depend on the aggregate action, and individual allocations and transfers are proportional to actions. For this example, the aggregate allocation has the form $Q(\Sigma x) = \min\{1, \alpha \Sigma x\}$ for a constant α. Thus, each bidder i's allocation on the low rationing region is a simple linear function of their action: $q_i(x) = \alpha x_i$. (This appears to be a general feature of aggregate allocations for mechanisms in which the good is rationed for low aggregate actions.)

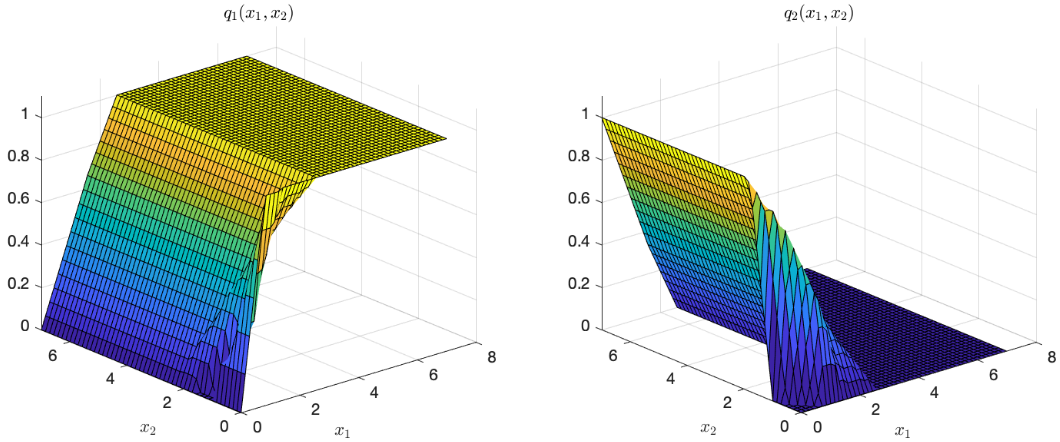
The first row of Figure 1 shows the approximate max-2min mechanism as computed by solving (6) with $k = 7$ (so that each bidder has 50 actions). This mechanism guarantees profit of at least 0.2564, or 51% of the expected value. The proportional auction (for which the message space is all of $\mathbb{R}_+$) is depicted in the second row for comparison. The approximate max-2min allocation bears a close resemblance to the proportional rule, including in the behavior of the aggregate allocation. Indeed, the solution in Brooks and Du (2020) was in part motivated by looking at simulations of this form. The approximate max-2min transfer does not suggest the proportional form. As we will discuss in Section 5, even holding fixed a particular max-2min allocation, there may be many transfer rules which could complete a max-2min mechanism. Numerical simulations of (6) need not produce the most interesting or tractable solution. As a result, for our subsequent examples, we will focus on max-2min allocations, and revisit the question of max-2min transfers in Section 5.

The bottom panel of Figure 1 shows the approximate min-2max information produced by (5). Profit in this information structure is at most 0.2856, so that the gap between $\underline{\Pi}^{\text{MAX}-2\text{MIN}}(k)$ and $\overline{\Pi}^{\text{MIN}-2\text{MAX}}(k)$ is approximately 5.8% of the expected value. Interestingly, the simulated min-2max information very nearly coincides with the theoretical solution with a continuum of signals: The interim expected value $w_1(x) = w_2(x)$ to be an increasing function of the aggregate signal. There is a cutoff, below which the interim expected value grows exponentially, and above which the interim expected value is equal to the ex post value. This structure gives rise to the discontinuities in the value function, evident in Figure 1, which occur when the interim expected value jumps up to the next higher value in the grid with increments of 0.1.

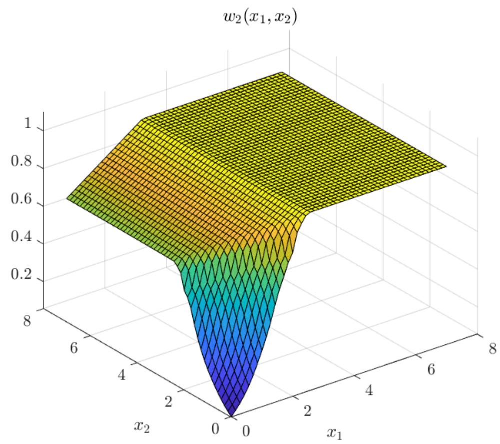
Our next simulation has $c = 0.1$, so that there is common knowledge that bidder 1's value is higher than bidder 2's.[10] In this case, it is socially efficient to always allocate the good to buyer 1. In contrast, the max-2min mechanism takes into account the cost of incentivizing truthtelling, and sometimes allocates to bidder 2 so as to reduce information rents. The simulated allocation is depicted in the top row of Figure 2. Together with the transfer that solves (6), this mechanism guarantees profit at least 0.2977, while the efficient surplus (if the good is always allocated to bidder 1) is now 0.6.[11] Remarkably, it appears that for signals in which the aggregate allocation is 1, the allocation only depends on bidder 1's signal. The approximate min-2max information structure has independent censored geometric signals, with the interim value function for v_2 depicted in the bottom row of Figure 2. (Bidder 1's interim expected value is simply $w_1(x) = w_2(x) + 0.1$.) Profit

[10] Note that this model does not have pure common values and is not characterized by Brooks and Du (2020).

[11] Note that the profit guarantee rises by much less than the increase in the efficient surplus, because in order to realize that gain, it would be necessary to allocate the good to bidder 1, which in turn would necessitate granting bidder 1 a large information rent.



(a) Approximate max-2min allocations



(b) Approximate min-2max interim value

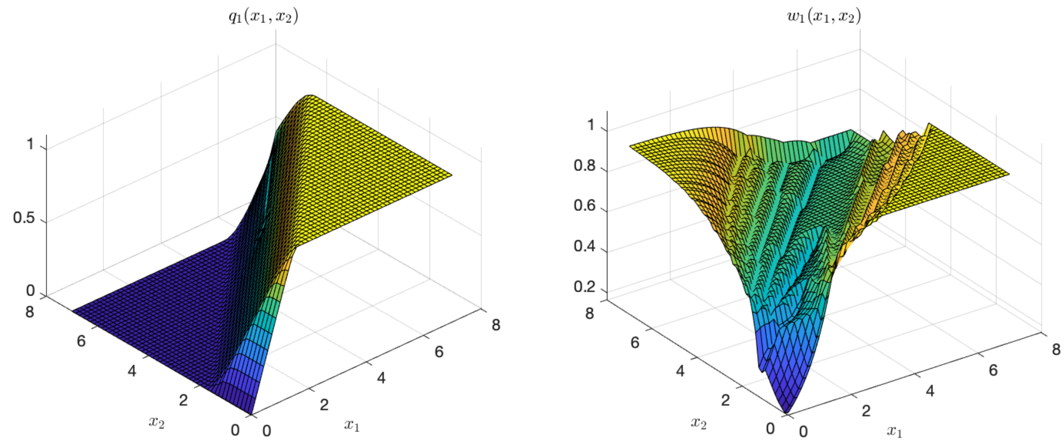
Figure 2: Approximate max-2min mechanism and min-2max information with perfectly correlated asymmetric values.

on this information structure is at most 0.2856. Thus, while this example does not satisfy the full support assumption, we see that the upper and lower bounds on profit are quite close. Note that bidder 1's allocation hits 1, and bidder 2's allocation hits 0, precisely on the region where the interim expected value maxes out, i.e., $w_1 = 1.1$ and $w_2 = 1$.

As c increases, the region where the allocation is interior shrinks. When c is sufficiently large, the optimal mechanism always allocates the good to bidder 1 at a price of 1.

3.4.2 Independent values

Our next example has two bidders whose values are independently distributed on the same ten-point grid in $[0, 1]$. The simulated allocation and interim values for bidder 1 are depicted in Figure 3. The corresponding objects for bidder 2 are symmetric.



(a) Approximate max-2min allocation (b) Approximate min-2max interim value

Figure 3: Approximate max-2min mechanism and min-2max information for independent values.

The allocation that solves (6) is in the left panel. The expected highest value in this discrete example is $0.68\overline{3}$, and the approximate max-2min mechanism guarantees profit of at least 0.2826, or approximately 41% of the efficient surplus.[12] We again see a region where the aggregate action is below a cutoff on which the good is rationed, and each bidder's allocation is linear in their action. A striking result is that on the high region where the good is always allocated, it appears that the allocation only depends on the difference in the bidders' signals, with a bidder's allocation being increasing in their action.

The interim expected value is on the right panel. Maximum profit on this information structure is at most 0.3170. While bidders' ex post values are independent, their interim expectations are highly correlated, with both bidders' interim expected values being higher when the absolute difference in their signals is large. Bidder i's interim value is highest when $x_i - x_j$ is above a threshold. The set of signals where bidder i's interim value is maximized roughly corresponds to the set of actions where their allocation hits 1.

4 Variations

The proof of Theorems 1 and 2 can be extended to a variety of auction design problems which are not formally subsumed in the model of Section 2. We now describe three such variations: Asymmetric demands among the bidders, auctioning multiple goods simultaneously, and ambiguity about the value distribution.

[12] Thus, while the profit guarantee is higher than with perfectly correlated values, it does not rise nearly as much as the expected value. The reason, of course, is that the bidders can obtain higher information rents when their values are independent.

4.1 Asymmetric demands

Even if bidders have the same per-unit value for the good, they may demand different quantities. For example, if the auction is an IPO, then different bidders may have different capacities for the risk associated with owning equity in the firm. These different risk capacities may be related to public information about the bidders, such as the sizes of their portfolios.

Let us suppose that it is public information that bidder i demands at most κ_i units of the good. A mechanism must now satisfy the additional restrictions $q_i(x) \leqslant \kappa_i$ for all i.

We claim that Theorems 1 and 2 can be generalized to models with such asymmetric demands, as we now explain. The program (5) is modified to the following:

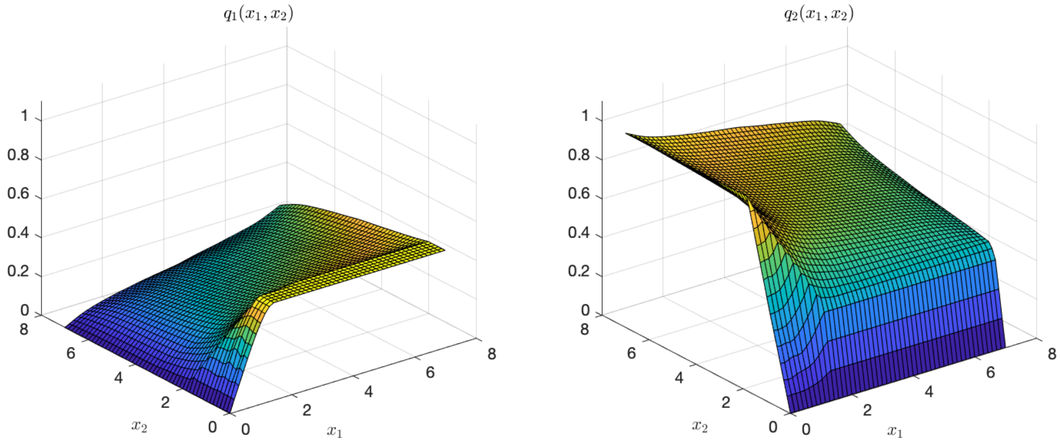
$$\begin{aligned} &\min_{\gamma:X(k)\to\mathbb{R}_+,\eta_i:X(k)\to\mathbb{R}_+,\sigma:X(k)\times V\to\mathbb{R}_+,w:X(k)\to\mathbb{R}^N_+} \sum_{x\in X(k)} \left[\gamma(x) + \sum_i \kappa_i \eta_i(x)\right] \\ \text{s.t. } &\gamma(x) + \eta_i(x) \geqslant \rho(x)\left[w_i(x) - \nabla_i^+ w(x)\right] \ \forall x; \\ &\sum_{v\in V} \sigma(x,v) = \rho(x) \ \forall x; \\ &\sum_{x\in X(k)} \sigma(x,v) = \mu(v) \ \forall v \\ &w(x) = \frac{1}{\rho(x)} \sum_{v\in V} v\sigma(x,v) \ \forall x. \end{aligned} \tag{26}$$

while (6) becomes

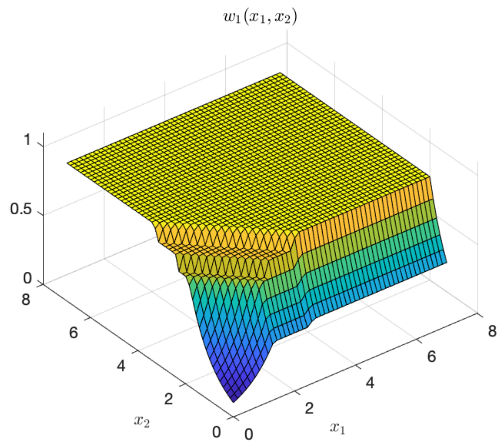
$$\begin{aligned} &\max_{q:X(k)\to A,t:X(k)\to\mathbb{R}^N,\lambda:V\to\mathbb{R}} \sum_{v\in V} \mu(v)\lambda(v) \\ \text{s.t. } &\lambda(v) \leqslant \Sigma t(x) + v\cdot\nabla^+ q(x) - \nabla^+\cdot t(x) \ \forall (v,x); \\ &\Sigma q(x) \leqslant 1 \ \forall x; \\ &q_i(x) \leqslant \kappa_i \ \forall i, x; \\ &t_i(0, x_{-i}) = 0 \ \forall i, x_{-i}. \end{aligned}$$

The proof that these programs bound (2) and (3) follows closely the proofs of Lemmas 2 and 3, with the additional demand constraints. The variables η_i in (26) are the multipliers on the demand constraints in the inner minimization program of (2) and are introduced when we take a dual as in the proof of Lemma 2. The proof of Lemma 3 is essentially unchanged.

In addition, the arguments for bounded λ in Lemma 4 and the shifting argument of Lemma 5 proceed essentially as before. A subtle feature of the proof of Lemma 4 is that in the perturbation which "drives out" the multipliers on the constraints (20), we change the value distribution but do not change γ. When we add demand constraints, the perturbation proceeds as before, and now both γ and η_i are unchanged. For Lemma 5, a key step is the transformation of an optimal solution (λ^*, Ξ^*, q^*) to (19) into a feasible solution $(\overline{q}, \overline{\lambda}, \overline{\Xi})$ for the dual of (6), defined in (25) which has approximately the same value. Critically, we



(a) Approximate max-2min allocations



(b) Approximate min-2max interim value

Figure 4: Approximate max-2min mechanism and min-2max information with pure common values but bidder 1 demands only 0.5 units.

have defined $\overline{q}$ to be either 0 or so that

$$\overline{q}_i(x) \leqslant q_i^*(x_i - 1/k, x_{-i}) \leqslant \kappa_i,$$

so that $\overline{q}$ also satisfies individual demand constraints. The rest of the proof goes through as before.

Figure 4 depicts the approximate max-2min mechanism and min-2max information for the pure common value example of Section 3.4, except that bidder 1 is only willing to buy 0.5 units of the good. In addition to the usual low linear rationing region, we see that there is a rectangular region where bidder 1's action is relatively high and bidder 2's action is relatively low on which bidder 1's allocation maxes out at 0.5. On this region, the good is still rationed, and bidder 2's allocation only depends on their signal. In the information structure, the interim expected value only depends on bidder 2's signal on the region where bidder 1's allocation has maxed out.

Note that the particular form for the feasibility constraint does not play a significant role in the argument, and we could easily generalize to other constraints on the allocation, e.g., a cap on the share of the good allocated to a subset of the bidders. The critical feature is that the constraint is "downward closed," meaning that if all bidders' allocations decrease, the constraint will still be satisfied.

4.2 Simultaneous auction of multiple goods

Suppose that there are L goods for sale, indexed by $l = 1, \ldots, L$, and each bidder demands a single unit of each of the goods. Let $v_{l,i}$ denote bidder i's value for good l. The primitive is a prior distribution over buyers' values for all goods. An information structure now consists of sets of signals and a joint distribution over signals and bidders' values for all goods. A mechanism now specifies sets of actions and allocations for each bidder and good.[13]

The program (5) is generalized as follows:

$$\begin{aligned}
&\min_{\gamma_l:X(k)\to\mathbb{R}_+,\,\sigma:X(k)\times V^L\to\mathbb{R}_+,\,w_l:X(k)\to\mathbb{R}^N_+} \sum_{l=1}^{L}\sum_{x\in X(k)} \gamma_l(x)\\
\text{s.t. }\ &\gamma_l(x) \geqslant \rho(x)\left[w_{l,i}(x) - \nabla_i^+ w_l(x)\right]\ \forall l, i, x;\\
&\sum_{v\in V}\sigma(x,v) = \rho(x)\ \forall x;\\
&\sum_{x\in X(k)}\sigma(x,v) = \mu(v)\ \forall v;\\
&w_l(x) = \frac{1}{\rho(x)}\sum_{v\in V} v_l\sigma(x,v)\ \forall l, x,
\end{aligned}$$

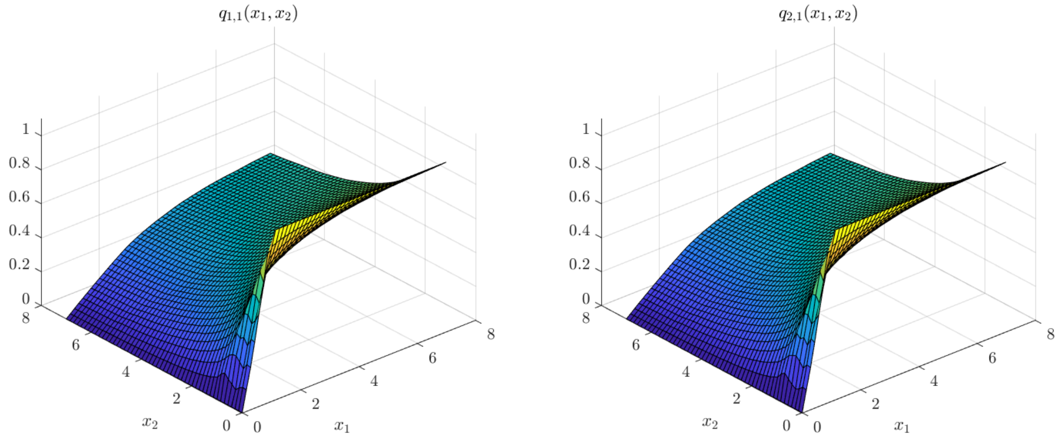
where now $w(x)$ is a matrix that specifies an interim expected value $w_{l,i}(x)$ for each good l and bidder i. The program (6) becomes

$$\begin{aligned}
&\max_{q_l:X(k)\to\mathbb{R}^N_+,\,t_l:X(k)\to\mathbb{R}^N,\,\lambda:V^L\to\mathbb{R}} \sum_{v\in V^L}\mu(v)\lambda(v)\\
\text{s.t. }\ &\lambda(v) \leqslant \sum_{l=1}^{L}\left(\Sigma t_l(x) + v_l\cdot\nabla^+ q_l(x) - \nabla^+\cdot t_l(x)\right)\ \forall (v,x);\\
&\Sigma q_l(x) \leqslant 1\ \forall l, x;\\
&t_{l,i}(0, x_{-i}) = 0\ \forall l, i, x_{-i}.
\end{aligned}$$

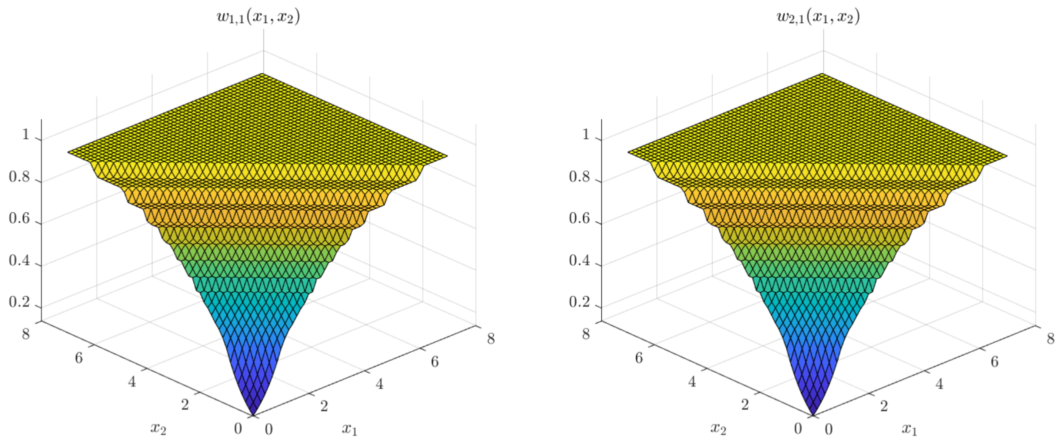
As with asymmetric demands, the proofs of Lemmas 2 and 3 proceed as before, by dropping non-local obedience and incentive constraints and fixing the multipliers on local incentives. Lemma 4 is also generalized, by showing that the constraints

$$q_{l,i}(x_i - 1/k, x_{-i}) \leqslant q_{l,i}(x) + \epsilon(k) \tag{27}$$

[13] The demand constraint example of the previous section could also be modeled with multiple goods, some of which are only assigned positive value by some bidders. Nonetheless, there is independent value to the extension with demand constraints, as it illustrates how more general feasibility constraints could be added to the model, including constraints which cannot be easily mapped into a multiple goods model.



(a) Approximate max-2min allocations



(b) Approximate min-2max interim value

Figure 5: Approximate max-2min mechanism and min-2max information with pure common values and two goods.

are redundant if $\epsilon(k) = C/k$ for C sufficiently large. The proof is as before, via a perturbation of the value distribution that drives out multipliers on (27). The proof of Lemma 5 is also generalized: when k is large, there is an optimal solution (λ^*, Ξ^*, q^*) that satisfies the no-downward-jump constraints (27) for each l and i. We can define a new solution $\overline{q}_{l,i}$ exactly as in (25), which has approximately the same value.

We illustrate this extension with a two-bidder two-good example. First assume bidders have pure common values for each good, so $v_{l,1} = v_{l,2}$ almost surely for each $l = 1, 2$. The common values are independently distributed across goods. The lower and upper bounds on Π^* are 0.5942 and 0.6542, respectively. The optimal mechanism and information are depicted in Figure 5. (Bidder 1's allocations are depicted, with bidder 2's allocations being symmetric.) The simulation clearly indicates that the allocation and interim expected values for the two goods are *exactly the same*. Thus, the two-good pure common value model reduces to a single-good pure common value model, in which the value for the single good is the sum of the values of the two goods. This is in spite of the fact that the underlying values are independent across the two goods. Why should this be the case? Clearly the Seller can treat the two goods as one, and only offer them bundled together, in which case all that matters are bidders' beliefs about the value of the bundle. At the same time, it is possible for Nature to only give bidders information about the value of the bundle, as in the simulation. In this case, the symmetry of the underlying value distribution implies that $w_{1,i}(x) = w_{2,i}(x)$ for all x, i.e., bidders always assign the same interim expected value to both goods. As a result, the Seller can do no better than the optimal profit guarantee when the goods are bundled.

Indeed, we conjecture that the limit analysis of Brooks and Du (2020) can be generalized to formally show that proportional auctions for the grand bundle are max-2min mechanisms when there are multiple goods with pure common values and the distribution of the goods' values is exchangeable. More broadly, let us say that a value distribution is *exchangeable across goods* if all $\mu(v) = \mu(v')$ for all v and v', where v'_i is a permutation of v_i for all i. We conjecture that if values are exchangeable across goods, then the multi-good problem reduces to a single-good problem in which bidders only learn about their value for the grand bundle, and the Seller only offers the grand bundle for sale.

If, however, values are not exchangeable across goods, then the multiple-good problem need not reduce to an auction for the grand bundle, as the following example shows. Let us now suppose that bidder 2's values $v_{l,2}$ are distributed as before, uniform on each good l and independent across goods; bidder 1 has the same value for good 2 as bidder 1 but assigns more value to good 1 than bidder 2: $v_{2,1} = v_{2,2}$ and $v_{1,1} = v_{1,2} + 1$. The approximate max-2min allocations are depicted in Figure 6. As we can see, bidders receive each good with different probabilities. As we would expect, good 1 is mostly allocated to bidder 1, since their value for that good is much higher. Interestingly, bidder 1 also tends to get more shares of good 2 than bidder 2, even though the two bidders have the same value, because of the endogenous bundling of the two goods in the max-2min mechanism.

4.3 Ambiguous correlation between values

Our last two extensions regards the constraint on the value distribution. We have assumed heretofore that the Seller knows the value distribution exactly, while at the same time taking a worst case over bidders' information and the equilibrium strategies. There is a clear tension here. Our last two extensions incorporate ambiguity with regard to the value distribution.

First, suppose that the Seller knows that each bidder i's value is distributed according to $\mu_i \in \Delta(V)$, but the Seller does not know the joint distribution of values.[14] Thus, an

[14] Carroll (2017) studies a robust multiple-good monopoly problem where the designer knows marginal distributions but takes a worst case over the joint distribution. In contrast, we model the sale of a single unit to multiple bidders, where there is ambiguity about the joint distribution of bidders' values. Moreover, Carroll (2017) assumes that the agent knows their ex post value, and just varies the correlation structure, whereas our model incorporates ambiguity about bidders' information.

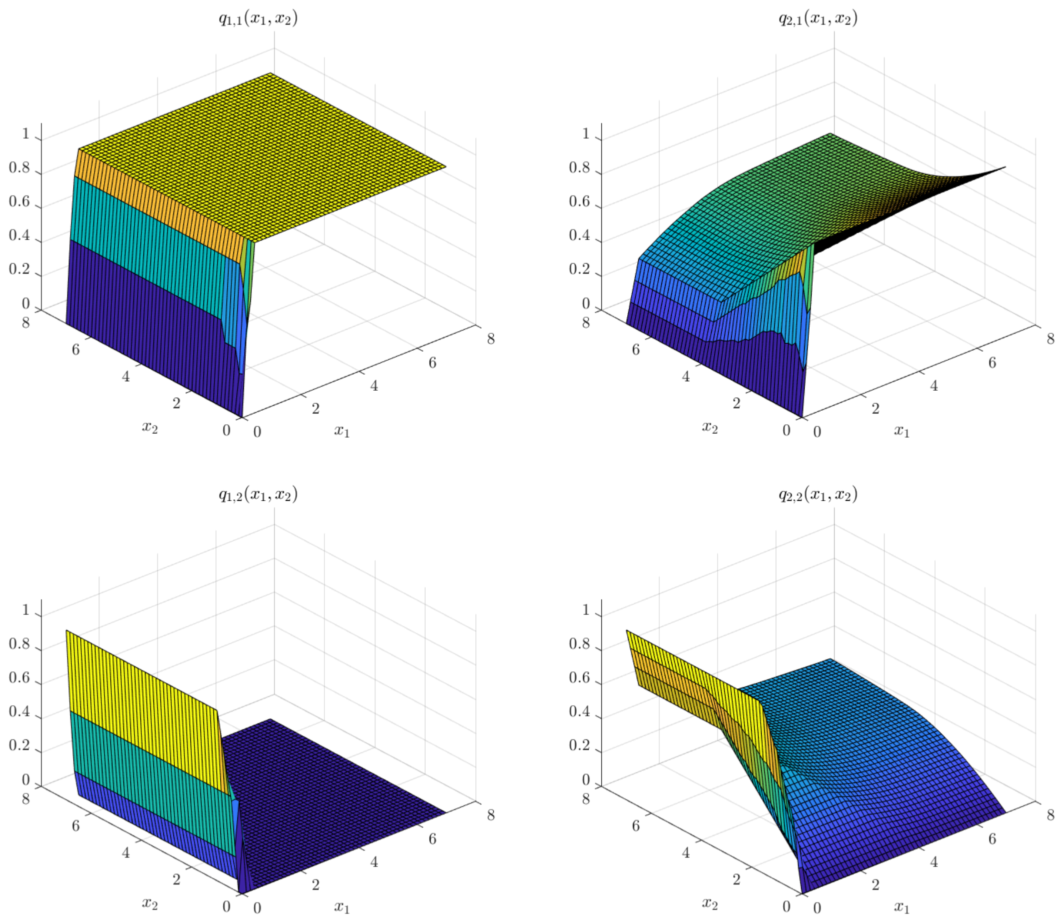


Figure 6: Max-2min allocation with two goods with non-exchangeable values.

information structure (S,σ) need only satisfy

$$\sum_{s,v_{-i}} \sigma(s,v_i,v_{-i}) = \mu_i(v_i)\ \forall i, v_i.$$

These constraints replace the marginal constraint in the program (5). The analogue of program (6) is now:

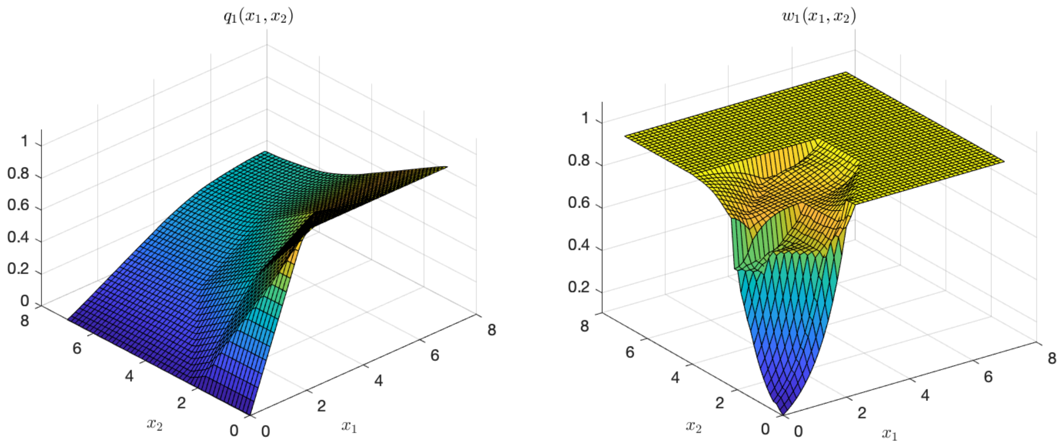
$$\max_{q:X(k)\to\mathbb{R}_+^N,t:X(k)\to\mathbb{R}^N,\lambda:V\to\mathbb{R}} \sum_{i=1}^{N}\sum_{v\in V} \mu_i(v_i)\lambda_i(v_i)$$

$$\text{s.t.}\ \sum_{i=1}^{N}\lambda_i(v_i) \leqslant \Sigma t(x) + v\cdot\nabla^+ q(x) - \nabla^+\cdot t(x)\ \forall (v,x);$$

$$\Sigma q(x) \leqslant 1\ \forall x;$$

$$t_i(0,x_{-i}) = 0\ \forall i, x_{-i}.$$

All of the previous steps in our argument go through as before, where we replace $\lambda(v)$ with $\sum_i \lambda_i(v_i)$, with one exception. In the proof of Lemma 4, we invoked a full-support



(a) Approximate max-2min allocation

(b) Approximate min-2max interim value

Figure 7: Approximate max-2min mechanism and min-2max information with uniformly distributed values and unknown correlation.

hypothesis on μ to prove that optimal λ are bounded. But now, λ will be bounded if each component λ_i is bounded, and the same argument for boundedness of λ_i goes through as long as μ_i has full support on V_i. In particular, we can always normalize the λ_i so that

$$\sum_{v_i \in V_i} \mu_i(v_i)\lambda_i(v_i) = \frac{\overline{\Pi}^{\text{MIN}-2\text{MAX}}(k)}{N}$$

for all i. Under this normalization, λ_i must be bounded above, since otherwise, in the version of (19) with fixed λ_i, Nature could place all of the mass on values with $\lambda_i(v_i)$ going to infinity, which would contradict the hypothesis that the value of the program is $\overline{\Pi}^{\text{MIN}-2\text{MAX}}(k)$. And since the optimal value of (19) is bounded below, λ_i must be bounded below as well. Given this result, the proof of Lemma 5 goes through unchanged.

We illustrate with two bidders whose values are uniformly distributed on the ten point grid in $[0, 1]$. The optimal allocation and interim value are illustrated in Figure 7. The mechanism guarantees profit at least 0.2537 and maximum profit on the information structure is at most 0.2834. Note that both of these numbers are lower than the corresponding figures for pure common uniform values (0.2564 and 0.2856) and independent uniform values (0.2826 and 0.3170), as they should be, since perfect correlation and independence are both feasible joint distributions for the present problem.

4.4 A penalty-based model of ambiguous value distribution

Going a step further, we can dispense with marginal constraints on values altogether, and instead represent the Seller's ambiguity about the value distribution with "cost" for the value distribution. This is in the spirit of the multiplier preferences Hansen and Sargent (2001) and the variational preferences of Maccheroni, Marinacci, and Rustichini (2006), where a penalty function on beliefs is used to discipline a worst-case analysis.

We now argue that Theorems 1 and 2 can be generalized to such a model of ambiguity aversion. In fact, we have already analyzed this model in the proof of Lemma 4, where the penalty function arose endogenously as part of the solution of the program (6). Fix a penalty $\lambda : V \to \mathbb{R}$. The programs (2) and (3) remain the same, except that the domain of information structures is now the set of finite information structures with arbitrary marginal on V, and the objective is now

$$\Pi(\mathcal{M}, \mathcal{I}, \beta) - \sum_{v\in V} \lambda(v) \sum_{s\in S} \sigma(s, v).$$

The program (5) is modified by dropping the marginal constraints on V and changing the objective to

$$\sum_{x\in X(k)} \gamma(x) - \sum_{v\in V} \lambda(v) \sum_{x\in X(k)} \sigma(x, v).$$

The only change to (6) is that λ is fixed and is no longer a variable over which we optimize.

Lemmas 2 and 3 go through as before, with the penalty replacing the marginal constraint on values. In the proof of Lemma 4, boundedness of the optimal λ holds by assumption, and no full support condition on the (endogenous) distribution of values is needed. The only significant change is in the proof of Lemma 5, since the new solution as defined in (25) is no longer feasible (since we cannot change the exogenously given λ). We can instead define a solution for the analogue of (24):

$$\overline{q}_i(x) = \begin{cases} \frac{q_i^*(x_i - 1/k, x_{-i})}{1+N\epsilon(k)} & \text{if } 0 < x_i < k; \\ 0 & \text{if } x_i = 0 \text{ or } x_i = k; \end{cases}$$

$$\overline{\Xi}(x) = \begin{cases} \frac{k-1}{k(1+N\epsilon(k))}\Xi^*(x) & \text{if } x \notin \partial X(k); \\ -(k-1)N\overline{v} - \max_{v\in V} \lambda(v) & \text{if } x \in \partial X(k), \end{cases}$$
$$- \left(1 - \frac{k-1}{k(1+N\epsilon(k))}\right) \max_{v\in V} \lambda(v)$$

It is straightforward to verify that this solution is feasible for (24) with fixed λ and has approximately the same value as (25) when k is large.

5 Further results

We now discuss further theoretical results on the baseline model of Sections 2 and 3.

5.1 Properties of approximate max-2min mechanisms

5.1.1 Approximate max-2min allocations

A key step in our argument (Lemma 4) shows that the dual to (5) has optimal solutions which downward jumps vanish as k becomes large. The following result shows that this is also true for (6):

Proposition 1. *Suppose that $\mu(v) > 0$ for all $v \in V$. Then there exists a constant $C > 0$ such that for every k, there exists an optimal solution to* (6) (λ^*, q^*, t^*) *such that $\nabla_i^+ q^*(x) \geqslant -C$ for all i and x.*

Proof of Proposition 1. As we showed in the proof of Lemma 5, (6) has the same optimal value as the program (24), where we substitute $\Xi(x)$ for the aggregate excess growth $\nabla^+ \cdot t - \Sigma t$, and add the expectation of Ξ under ρ to the objective. Now, exactly the same sequence of steps as in the proof of Lemma 4 (normalizing by $k-1$ instead of k) can be used to show that the optimal λ is bounded and that there exists an optimal solution (λ^*, q^*, t^*) to (24) that satisfies the constraints

$$q_i^*(x) \leqslant q_i^*(x_i + 1/k, x_{-i}) + \epsilon(k) \ \forall i, x_i < k,$$

where $\epsilon(k) = C/(k-1)$ for a constant $C > 0$.

It only remains to show that for any optimal solution (λ^*, Ξ^*, q^*) for (24), there exists a λ' and t' such that (λ', q^*, t') is an optimal solution to (6). To see why this is the case, consider the linear program (6) where we hold fixed $q = q^*$. This program's dual is

$$\begin{aligned} &\min_{\sigma: X(k)\times V \to \mathbb{R}_+} \sum_{x\in X(k)} \sum_{v \in V} \sigma(x,v) v \cdot \nabla^+ q(x) \\ \text{s.t. } &\sum_{v\in V} \sigma(x,v) - (k-1)(\sigma(x_i - 1/k, x_{-i}, v)\mathbb{I}_{x_i>0} - \sigma(x,v)\mathbb{I}_{x_i<k}) = 0 \ \forall i, x \text{ s.t. } x_i > 0; \\ &\sum_{x\in X(k)} \sigma(x,v) = \mu(v) \ \forall v. \end{aligned}$$

As before, we can solve out the first constraint to determine that the marginal of σ on x must be ρ, so that the program reduces to the transportation problem

$$\begin{aligned} &\min_{\sigma: X(k)\times V \to \mathbb{R}_+} \sum_{x,v} \sigma(x,v) v \cdot \nabla^+ q(x) \\ \text{s.t. } &\sum_{v} \sigma(x,v) = \rho(x) \ \forall x; \\ &\sum_{x\in X(k)} \sigma(x,v) = \mu(v) \ \forall v. \end{aligned}$$

The dual of this program is precisely (24), where we again hold fixed $q = q^*$. Thus, we conclude that (6) and (24), where we hold fixed $q = q^*$, have the same value. Since q^* is part of an optimal solution to (24), which has the same optimal value as (6), we conclude that it is also part of an optimal solution to (6). □

5.1.2 Approximate max-2min transfers

The proof of Proposition 1 shows that the programs (6) and (24) have the same value, even when we hold fixed a particular allocation. Thus, there is in some sense an equivalence between $\Xi(x)$ and the *aggregate excess growth* $\nabla^+ \cdot t(x) - \Sigma t(x)$. The following two results characterize this relationship.

Proposition 2. *Fix $\Xi : X(k) \to \mathbb{R}$. There exists a t that solves*

$$\nabla^+ \cdot t(x) - \Sigma t(x) = \Xi(x) \ \forall x; \tag{28}$$

$$t_i(0, x_{-i}) = 0 \ \forall i, x_{-i} \tag{29}$$

if and only if

$$\sum_{x \in X(k)} \rho(x)\Xi(x) = 0. \tag{30}$$

Proof of Proposition 2. Given Ξ, Fredholm's alternative says that there is a t that solves (28) and (29) if and only if there does not exist a ρ' such that

$$\sum_{x \in X(k)} \rho'(x)\Xi(x) \neq 0 \tag{31}$$

$$\rho'(x) = \begin{cases} \frac{k-1}{k}\rho'(x_i - 1/k, x_{-i}) & \text{if } 0 < x_i < k; \\ (k-1)\rho'(k - 1/k, x_{-i}) & \text{if } x_i = k. \end{cases}$$

It is easy to see that the choice of $\rho'(0)$ pins down the rest of ρ', and in fact

$$\rho'(x) = \rho(x)\frac{\rho'(0)}{\rho(0)}.$$

As a result, (31) holds if and only if $\sum_{x \in X(k)} \rho(x)\Xi(x) \neq 0$. Thus, (28) and (29) has a solution if and only if (30) holds. □

Proposition 2 gives a simple way to go back and forth between solutions to (6) and (24). First, given a feasible solution (λ, q, t) to (6), we know that the expected aggregate excess growth is zero. Thus, if we define Ξ according to (28), then the solution (λ, Ξ, q) is feasible for (24) and has the same value. On the other hand, given a feasible solution (λ, Ξ, q) to (24), the solution $(\lambda + C, \Xi - C, q)$ is also feasible for any constant $C \in \mathbb{R}$. We can therefore without loss of generality restrict attention to feasible solutions to (24) for which the expectation of Ξ under ρ is zero. For any such solution, we can find a transfer t with the given excess growth, so that (λ, q, t) is feasible for (6) and has the same value. This discussion is formalized in the following corollary:

Corollary 2. *The triple (λ^*, q^*, t^*) is an optimal solution to (6) if and only if (λ^*, q^*, Ξ^*) is an optimal solution to (24), where $\Xi^* = \nabla^+ \cdot t^* - \Sigma t^*$.*

The aggregate excess growth played a prominent role in our earlier analysis of common values in Brooks and Du (2020). In the limit model with a continuum of actions, we derived an optimal aggregate excess growth function, which turned out to only depend on the aggregate action Σx (where the limit min-2max signals are independent draws from the standard exponential distribution, which is the limit in distribution of ρ as k goes to infinity). Using the fact that the expectation of this Ξ under the independent standard exponential distribution is zero, we explicitly constructed transfers that attained the optimal excess growth and are feasible for the analogue of (6).[15] Proposition 2 shows, non-constructively and in a finite version of the model, that zero expectation is necessary and sufficient for this to be possible.

In fact, in the pure common value model, Brooks and Du (2020) constructed two distinct transfer rules with the same aggregate excess growth. More broadly, we suspect that there may be many solutions to the aggregate excess growth equation, and note that any convex combination of solutions is also a solution. This multiplicity of max-2min transfer rules, not all of which are of practical interest, presents a challenge to the study of max-2min mechanisms, and additional properties may be needed to isolate the most useful transfer rules. For example, in the common value model, the transfer rule in the proportional auction is characterized by the property that the aggregate transfer depends only on the aggregate action.

5.2 Rate of convergence

We now characterize the rate of convergence of the relaxed profit bounds to Π^*:

Proposition 3. *Suppose that $\mu(v) > 0$ for all $v \in V$. For all $k \geqslant 1$,*

$$\left|\overline{\Pi}^{\text{MIN}-2\text{MAX}}(k) - \underline{\Pi}^{\text{MAX}-2\text{MIN}}(k)\right| \leqslant \frac{\overline{v}}{k} + o(1/k).$$

Hence, $\underline{\Pi}^{\text{MAX}-2\text{MIN}}(k)$ and $\overline{\Pi}^{\text{MIN}-2\text{MAX}}(k)$ converge to Π^* at a rate of $1/k$.

Proof of Proposition 3. Theorem 2 shows that $\overline{\Pi}^{\text{MIN}-2\text{MAX}}(k) \geqslant \underline{\Pi}^{\text{MAX}-2\text{MIN}}(k)$. The proof of Lemma 5 shows that

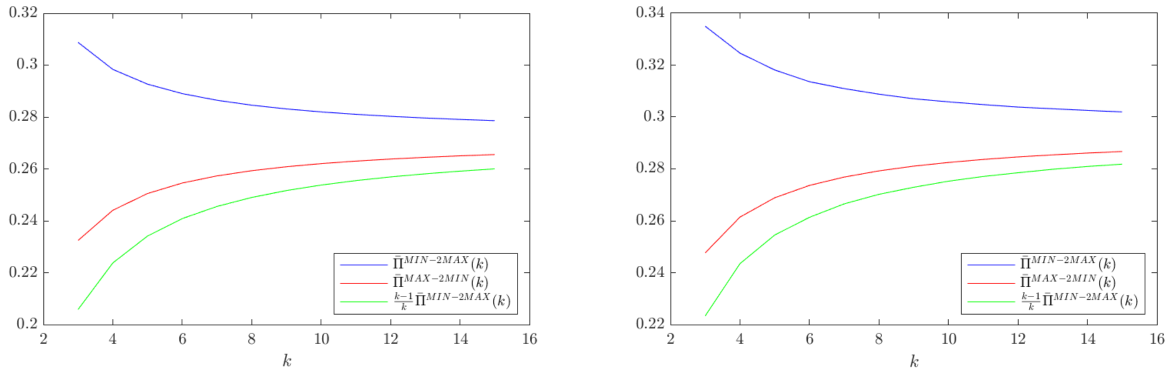
$$\underline{\Pi}^{\text{MAX}-2\text{MIN}}(k) \geqslant \frac{k-1}{k+C}\overline{\Pi}^{\text{MIN}-2\text{MAX}}(k) - \left(1-\frac{1}{k}\right)^{k^2-1}\left(\frac{k-1}{k+C}k + k - 1\right)N\overline{v}.$$

for a constant C. Hence,

$$\overline{\Pi}^{\text{MIN}-2\text{MAX}}(k) - \underline{\Pi}^{\text{MAX}-2\text{MIN}}(k) \leqslant \frac{1+C}{k+C}\overline{\Pi}^{\text{MIN}-2\text{MAX}}(k) + \left(1-\frac{1}{k}\right)^{k^2-1}\left(\frac{k-1}{k+C}k + k - 1\right)N\overline{v}.$$

Since $\overline{\Pi}^{\text{MIN}-2\text{MAX}}(k) \leqslant \overline{v}$, this immediately gives the first result. The second result follows immediately from the fact that $\Pi^* \in [\underline{\Pi}^{\text{MAX}-2\text{MIN}}(k), \overline{\Pi}^{\text{MIN}-2\text{MAX}}(k)]$ for all k. □

[15] In the continuum model, an additional complication arises: If we take the action space to be all of $\mathbb{R}_+$, there may be solutions to the analogue of the aggregate excess growth equation (28), such as setting $\nabla_i t - t_i = \Xi/N$ for all i, that diverge at infinity and lead to a mechanism that has no equilibria on any type space. For that reason, we further restricted attention to bounded solutions. In the finite action model, any solution must be bounded. Moreover, we suspect that the pathological limit solutions cannot arise as limits of finite solutions, because of the need to satisfy the aggregate excess growth equation at the highest signal, where the ∇t_i term does not appear.



(a) Pure common values (b) Independent values

Figure 8: Rate of convergence.

We illustrate this rate of convergence with two examples, depicted in Figure 8. Both examples are for $N = 2$ and values in a grid $V = \{0, 0.1, \ldots, 1\}$. In the left panel, bidders have a pure common value that is uniform $\{0, 0.1, \ldots, 1\}$, i.e., μ is concentrated on the diagonal. In the right-hand panel, values are independent uniform on V. In each case, the blue and red lines are $\overline{\Pi}^{\text{MIN}-\text{2MAX}}(k)$ and $\underline{\Pi}^{\text{MAX}-\text{2MIN}}(k)$, respectively. Note that only the independent example satisfies the full-support hypothesis of Proposition 3. Nonetheless, in both cases the green line is $\overline{\Pi}^{\text{MIN}-\text{2MAX}}(k)(k-1)/k$ and is always below the red line, consistent with the theoretical rate of convergence.

We note that if (γ^*, σ^*) is a solution to (5), then $\overline{\Pi}^{\text{MIN}-\text{2MAX}}(k)$ is only an upper bound on maximum profit of the information structure $(X(k), \sigma^*)$ across all mechanisms and equilibria; because the bound uses only local incentive constraints, it is logically possible that maximum expected profit is strictly lower and converges faster to Π^*. A corresponding statement applies to solutions to (6) and the bound $\underline{\Pi}^{\text{MAX}-\text{2MIN}}(k)$.

5.3 Robustness to fundamentals

An important feature of the approximate max-2min mechanisms, in addition to their optimal worst-case performance, is that we can bound their performance even when the model from which they were derived is misspecified.

To develop this result, we need the following lemma, which asserts that any *feasible* solution to the programs (5) or (6) has a corresponding bound on equilibrium expected profit.

Lemma 6. *Fix $k \geqslant 1$. Suppose (γ, σ, w) is a feasible solution to (5), and let $\mathcal{I} = (X(k), \sigma)$. Then*

$$\sup_{\mathcal{M}' \in \mathbf{M}} \sup_{b \in B(\mathcal{M}', \mathcal{I})} \Pi(\mathcal{M}', \mathcal{I}, \beta) \leqslant \sum_{x \in X(k)} \gamma(x).$$

Suppose (λ, q, t) *is a feasible solution to* (6)*, and let* $\mathcal{M} = (X(k), q, t)$*. Then*

$$\inf_{\mathcal{I}' \in \mathbf{I}} \inf_{b \in B(\mathcal{M}, \mathcal{I}')} \Pi(\mathcal{M}, \mathcal{I}', \beta) \geqslant \sum_{v \in V} \lambda(v)\mu(v).$$

Proof of Lemma 6. This is an immediate implication of the proofs of Lemmas 2 and 3. In particular, the program (5) is obtained by taking the dual of the inner maximization over mechanisms and equilibria from (2), so that any feasible solution to that dual provides an upper bound on the value of the primal, meaning that it provides an upper bound on profit under any mechanism and equilibrium. Similarly, we obtained (6) by taking the dual of the inner minimization program, and any feasible solution to the dual provides a lower bound on profit in the primal program. □

We can now formalize the bounds for a misspecified prior, which generalizes Proposition 7 of Brooks and Du (2020):

Proposition 4. *Suppose that* (λ, q, t) *is a feasible solution to* (6)*, and extend the domain of* λ *to all of* $\mathbb{R}_+^N$ *according to*

$$\lambda(v) = \min_{x \in X(k)} \left[\Sigma t(x) + v \cdot \nabla^+ q(x) - \nabla^+ \cdot t(x)\right]$$

Then for all $\mu' \in \Delta(\mathbb{R}_+^N)$ *with finite support, revenue in any information structure and equilibrium is at least*

$$\sum_{v \in \mathbb{R}_+^N} \lambda(v)\mu'(v). \tag{32}$$

In particular, the bound (32) *holds for an optimal solution* (λ^*, q^*, t^*)*.*

Proof of Proposition 4. Suppose that the prior is μ'. Then clearly (λ, q, t) is a feasible solution for the program (6), where we replace μ with μ'. From Lemma 6, we conclude that (32) is a lower bound on equilibrium profit. □

5.4 The monotonicity conjecture and its implications

We have conducted tens of thousands of simulations of the programs (5) and (6). In every case, we have found that both the dual to (5) and (6) have optimal solutions in which the allocation is such that $q_i^*(x)$ is non-decreasing in x_i for all i. We refer to such an allocation (and the corresponding solution) as *monotone*. We conjecture that a monotone solution always exists for the programs (5) and (6) for all k.[16] This conjecture has interesting implications.

First, suppose that (λ^*, Ξ^*, q^*) is an optimal solution to (24) and is monotone. Clearly, λ^* must satisfy

$$\lambda^*(v) = \min_{x \in X(k)} v \cdot \nabla^+ q^*(x) - \Xi^*(x).$$

[16]Curiously, the allocation proportional auction constructed in Brooks and Du (2020) need not be monotone. Thus, while we conjecture that there always exists a monotone solution, we know for a fact that there are max-2min auctions that are non-monotone.

Monotonicity of q^* implies that $\nabla^+ q^* \geqslant 0$, so that λ^* is a minimum of non-decreasing functions. As a result, λ^* must also be non-decreasing. This implies a simple comparative static result. A set $X \subseteq \mathbb{R}_+^N$ is *upward closed* if $v \in X$ and $v' \geqslant v$ implies that $v' \in X$. We say that μ *first-order stochastically dominates* μ' if for every upward closed set, $\sum_{v \in X} \mu(v) \geqslant \sum_{v \in X} \mu'(v)$. It is a standard result that if μ first-order stochastically dominates μ', then for every non-decreasing function $f : V \to \mathbb{R}$,

$$\sum_{v \in V} f(v)\mu(v) \geqslant \sum_{v \in V} f(v)\mu'(v)$$

(e.g., Sriboonchita et al., 2009, Theorem 3.3). Hence, the lower bound from Proposition 4 is increasing in μ in the first-order stochastic dominance order. If this is true for all k, then Π^* must also be increasing in μ in the first-order stochastic dominance order.

Second, monotonicity implies additional structure on approximate min-2max information structures. Suppose that (19) has an optimal solution (λ^*, Ξ^*, q^*) that is monotone. Further suppose that $q_i^*(x) > 0$ for some x. Then for all $x_i' > x_i$, $q_i^*(x_i', x_{-i}) > 0$ as well. Complementary slackness then implies that for all $x_i' \geqslant x_i$,

$$\gamma(x_i', x_{-i}) = \rho(x_i', x_{-i})[w_i(x_i', x_{-i}) - \nabla_i^+ w(x_i', x_{-i})].$$

This property can be reformulated as follows: Let $\psi_i(x) = w_i(x) - \nabla_i^+ w(x)$ denote bidder i's *virtual value*. For a set $Y \subseteq \mathbb{R}$, a function $f : Y \to \mathbb{R}$ is *strictly single crossing* if $f_i(y) > 0$ implies $f_i(y') \geqslant 0$ for all $y' > y$. We say that virtual values have *strict single crossing differences* if the function

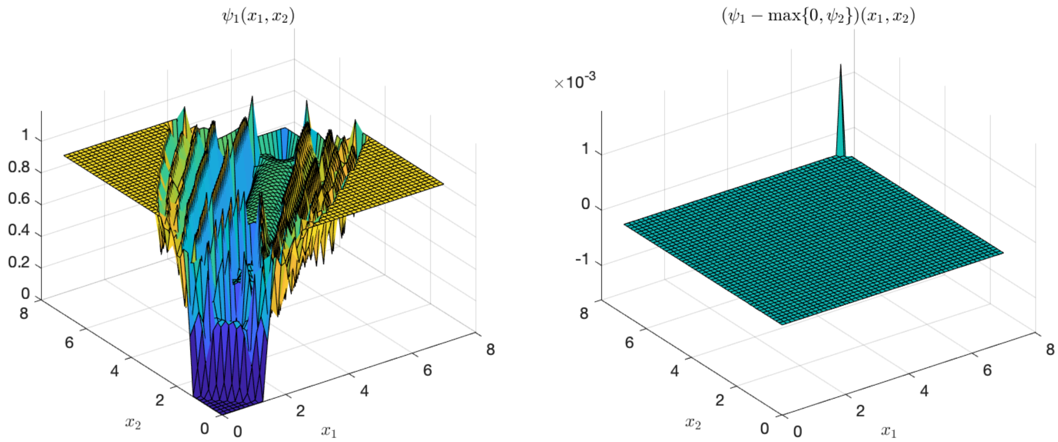
$$\psi_i(x_i, x_{-i}) - \max\{0, \max_{j \neq i} \psi_j(x_i, x_{-i})\} \tag{33}$$

is strictly single crossing as a function of x_i, for all i and x_{-i}. If the monotonicity conjecture is true, then for all k, there exists an approximate min-2max information structure that solves (5) and has strict single crossing differences.

In fact, the simulations suggest that there is even more structure on approximate min-2max virtual values, beyond strict single crossing differences. Figure 9 depicts the virtual values for the independent uniform example of Section 3.4. Note that in the case of independent values, the virtual value function is symmetric across bidders. Also note the large regions where bidders' virtual values are exactly equal. In fact, for the independent values example, the difference between the virtual values is generally less than 10^{-4}, except when bidders have the highest possible signals.[17]

This suggests that the strict single crossing differences conjecture can be significantly strengthened: For $Y \subseteq \mathbb{R}$, we say that a function $f : Y \to \mathbb{R}$ is *weakly single crossing* if $f_i(y) \geqslant 0$ implies $f_i(y') \geqslant 0$ for all $y' > y$. An information structure has virtual values with *weakly single crossing differences* if the virtual value difference (33), viewed as a function of x_i, is weakly single crossing for all i and x_{-i}. This property is satisfied in the simulations,

[17] The barrier algorithm of Gurobi, with which these objects were computed, only asymptotically approaches complementary slackness, so it is expected that there is some noise in the calculation of virtual value differences. This is especially true for high signals, where ρ is small and hence the impact of virtual values on the objective is negligible.



(a) Virtual value

(b) Virtual value difference

Figure 9: Virtual values and virtual value differences.

and we conjecture that it is true more broadly. While the difference may seem subtle, the simulations show that it is frequently not the case that there is a unique highest virtual value. Indeed, the most common situation seems to be that all bidders tie for highest virtual value. Weak single crossing differences implies that if a bidder ties, then they continue to have a highest virtual value if their signal increases. This is consistent with an intuition that min-2max information structures should make the Seller is indifferent between many allocations. For example, in Brooks and Du (2020), the min-2max information structure has the property that the Seller is *always* indifferent as to which bidder is allocated the good.

5.5 Other objectives

The focus of our analysis until this point has been expected profit. One may ask: Does our model have interesting implications for other welfare objectives, in particular total surplus? As the following proposition shows, the answer to this particular question is essentially no. Given a mechanism $\mathcal{M} = (A, q, t)$, information structure $\mathcal{I} = (S, \sigma)$, and strategy profile b, let

$$TS(\mathcal{M}, \mathcal{I}, b) = \sum_{s \in S} \sum_{v \in V} \sum_{a \in A} v_i q_i(a) b(a|s) \sigma(s, v)$$

denote the resulting expected total surplus. Let

$$\underline{TS} = \max_{i=1,\ldots,N} \sum_{v \in V} v_i \mu(v)$$

denote the highest ex ante expected value among the bidders.

Proposition 5.

$$\sup_{\mathcal{M} \in \mathbf{M}} \inf_{\mathcal{I} \in \mathbf{I}} \inf_{b \in BNE(\mathcal{M}, \mathcal{I})} TS(\mathcal{M}, \mathcal{I}, b) = \inf_{\mathcal{I} \in \mathbf{I}} \sup_{\mathcal{M} \in \mathbf{M}} \sup_{b \in BNE(\mathcal{M}, \mathcal{I})} TS(\mathcal{M}, \mathcal{I}, b) = \underline{TS}.$$

Proof of Proposition 5. Let i be the index of a bidder with the highest ex ante expected value. Then clearly a feasible mechanism is $q_i(a) = 1$ and $q_j(a) = 0$ for all a and $j \neq i$, and $t_j(a) = 0$ for all a and j. This mechanism is guaranteed to generate $\underline{TS}$ regardless of the information structure and equilibrium. On the other hand, if Nature chooses the degenerate information structure in which each bidder has a single signal, then total surplus in any mechanism and equilibrium must be less than $\underline{TS}$. $\square$

Thus, in order to obtain non-degenerate results for social welfare, one either needs to modify the objective (such as by using min-max regret) or by imposing restrictions on the set of information structures so that the domain of minimization does not include a least informative information structure.

6 Discussion and conclusion

This paper has proven a strong minimax theorem for informationally robust auction design with interdependent values. The result says that the joint information design/mechanism design game has the same expected profit Π^*, regardless of whether the Seller or Nature moves first, and regardless of how we select an equilibrium. The theorem also implies that for every $\epsilon > 0$, there exists a finite mechanism for which expected profit is at least $\Pi^* - \epsilon$ in all information structures and equilibria, and there exists a finite information structure for which expected profit is at most $\Pi^* + \epsilon$ in all information structures and equilibria. These approximate max-2min mechanisms and min-2max information structures are a promising subject for further study in theoretical and applied mechanism design. There are at least four important directions for future research, which we now discuss.

6.1 Detailed analysis of particular specifications

We have reported simulations of approximate max-2min mechanisms and min-2max for various specifications, e.g., independent values. Each of these models could be analyzed more fully to understand the particular form of max-2min auctions and min-2max information structures, as in Brooks and Du (2020). For example, is it the case that max-2min mechanisms only depend on the difference in signals in the independent uniform case? Exactly when does the max-2min multi-good auction collapse to a single-good auction for the grand bundle?

6.2 The continuum limit

We have restricted attention to finite mechanisms and information structures. This means that our results immediately apply to the finite models which can be solved numerically, and it also allows us to avoid the delicate theoretical issue of equilibrium existence. At the same time, it may be that the profit guarantee is not exactly attained with finitely many actions or signals, and sharper and more interpretable results can be obtained in the continuum limit. Such is the case in Brooks and Du (2020), where we constructed a max-2min mechanism and min-2max information structure with continua of actions and signals that exactly attain the profit guarantee. Moreover, the mechanism can simply be described as a proportional auction, and the information structure has the "additive-exponential" form where signals are iid exponential and interim expected values only depend on the sum of the signals. In the limit, the gap between the min-2max and max-2min programs vanishes, and the max-2min mechanism is a profit maximizing direct mechanism for the min-2max information structure, and the min-2max information structure is a profit-minimizing Bayes correlated equilibrium for the max-2min mechanism. We termed this a *strong max-min solution.*

Can the strong minimax theorem be extended to continua of actions/signals for general environments, so that the profit guarantee is exactly attained? Is there a limit formulation in which the min-2max and max-2min programs are exactly a dual pair? Does a strong max-min solution, in the sense of Brooks and Du (2020), always exist?

6.3 More general environments

Our assumptions are sufficient for the strong minimax theorem, but not necessary. What is the most general environment in which the strong minimax theorem holds? For example, can it be generalized to richer constraints on allocations or richer preferences, such as those with complementarities between goods?

6.4 Restrictions on information

Perhaps most importantly, we have allowed for no lower bound on bidders' information. Does the strong minimax theorem continue to hold if, for example, bidders know their own values? More broadly, we feel that while the present model does address robustness to model misspecification, the bounds obtained with with no restrictions on information are too conservative. We predict that this theory will become more useful as we find ways to model intermediate ambiguity, where there are some restrictions on information and equilibrium, but we do not commit to a single information structure and equilibrium as in the standard model.

References

Bergemann, D., B. Brooks, and S. Morris (2016): "Informationally Robust Optimal Auction Design," Tech. rep., Princeton University and the University of Chicago and Yale University, working paper.

——— (2019): "Revenue guarantee equivalence," *The American Economic Review*, 109, 1911–29.

Bergemann, D. and S. Morris (2013): "Robust Predictions in Games with Incomplete Information," *Econometrica*, 81, 1251–1308.

——— (2016): "Bayes Correlated Equilibrium and the Comparison of Information Structures in Games," *Theoretical Economics*, 11, 487–522.

Brooks, B. (2013): "Surveying and Selling: Belief and Surplus Extraction in Auctions," Working paper.

Brooks, B. and S. Du (2020): "Optimal Auctions with Common Values: An Informationally Robust Approach," Tech. rep., The University of Chicago and University of California-San Diego, working paper.

Carroll, G. (2017): "Robustness and separation in multidimensional screening," *Econometrica*, 85, 453–488.

Chen, Y.-C. and J. Li (2018): "Revisiting the Foundations of Dominant-Strategy Mechanisms," *Journal of Economic Theory*, 178, 294–317.

Chung, K.-S. and J. C. Ely (2007): "Foundations of Dominant-Strategy Mechanisms," *The Review of Economic Studies*, 74, 447–476.

Crémer, J. and R. P. McLean (1988): "Full Extraction of the Surplus in Bayesian and Dominant Strategy Auctions," *Econometrica*, 1247–1257.

Du, S. (2018): "Robust mechanisms under common valuation," *Econometrica*, 86, 1569–1588.

Hansen, L. and T. J. Sargent (2001): "Robust control and model uncertainty," *American Economic Review*, 91, 60–66.

Hartline, J. D. and T. Roughgarden (2009): "Simple Versus Optimal Mechanisms," in *Proceedings of the 10th ACM conference on Electronic commerce*, ACM, 225–234.

Maccheroni, F., M. Marinacci, and A. Rustichini (2006): "Ambiguity aversion, robustness, and the variational representation of preferences," *Econometrica*, 74, 1447–1498.

McAfee, R. P., J. McMillan, and P. J. Reny (1989): "Extracting the Surplus in the Common-Value Auction," *Econometrica*, 57, 1451–1459.

Myerson, R. B. (1981): "Optimal Auction Design," *Mathematics of Operations Research*, 6, 58–73.

Neeman, Z. (2003): "The Effectiveness of English Auctions," *Games and Economic Behavior*, 43, 214–238.

Sriboonchita, S., W.-K. Wong, S. Dhompongsa, and H. T. Nguyen (2009): *Stochastic dominance and applications to finance, risk and economics*, CRC Press.

Yamashita, T. (2015): "Implementation in Weakly Undominated Strategies: Optimality of Second-Price Auction and Posted-Price Mechanism," *The Review of Economic Studies*, 82, 1223–1246.

——— (2016): "Revenue Guarantee in Auction with a (Correlated) Common Prior and Additional Information," Working paper.

Yamashita, T. and S. Zhu (2018): "On the foundations of ex post incentive compatible mechanisms," .

A Omitted proofs

A.1 Proof of Lemma 4

Let (λ^*, Ξ^*, q^*) be an optimal solution of (19). Without loss of optimality, we can assume that

$$\sum_{v\in V} \mu(v)\lambda^*(v) = \overline{\Pi}^{\text{MAX-2MIN}}(k)$$

and

$$\sum_{x\in X(k)} \rho(x)\Xi^*(x) = 0.$$

Lemma 7. *Suppose $\mu(v) > 0$ for all $v \in V$. Then $|\lambda^*(v)| \leqslant \max_{v'\in V} \frac{\overline{v}}{\mu(v')}$ for all $v \in V$ and all k.*

Proof of Lemma 7. We first show that for all k and $v \in V$, $\lambda^*(v) \leqslant \overline{v}$. For the sake of contradiction, suppose not, i.e., there exist some k and v' such that $\lambda^*(v') > \overline{v}$. Consider the dual of (19) where we fix λ^*:

$$\begin{aligned}
&\min_{\gamma:X(k)\to\mathbb{R}_+,\,\sigma:X(k)\times V\to\mathbb{R}_+} \sum_{x\in X(k)} \gamma(x) - \sum_{v\in V, x\in X(k)} \lambda^*(v)(\sigma(v,x) - \mu(v)) \\
&\text{s.t. } \gamma(x) \geqslant \begin{cases} \sum_{v\in V} v_i k[\sigma(x,v) - \sigma(x_i + 1/k, x_{-i}, v)] & \text{if } x_i < k - 1/k, \\ \sum_{v\in V} v_i[k\sigma(x,v) - \sigma(x_i + 1/k, x_{-i}, v)] & \text{if } x_i = k - 1/k, \\ \sum_{v\in V} v_i\sigma(x,v) & \text{if } x_i = k, \end{cases} \\
&\qquad \forall i, x; \\
&\sum_{v\in V} \sigma(x,v) = \rho(x)\ \forall x,
\end{aligned} \tag{34}$$

Let $\overline{\sigma}(x,v) = \rho(x)\delta_{v'}(v)$ and $\overline{\gamma}(x) = \rho(x)\max_i v'_i$. It is easy to check that $(\overline{\sigma}, \overline{\gamma})$ is feasible for the above program and obtains an objective strictly less than $\max_i v'_i - \overline{v} + \overline{\Pi}^{\text{MIN}-2\text{MAX}}(k) < \overline{\Pi}^{\text{MIN}-2\text{MAX}}(k)$. This is an contradiction since the optimal value of (34) is $\overline{\Pi}^{\text{MIN}-2\text{MAX}}(k)$.

Since $\lambda^*(v) \leqslant \overline{v}$ and $\sum_v \mu(v)\lambda^*(v) = \overline{\Pi}^{\text{MIN}-2\text{MAX}}(k)$, we must have $\lambda^*(v) \geqslant (\overline{\Pi}^{\text{MIN}-2\text{MAX}}(k) - \overline{v})/\mu(v) \geqslant -\overline{v}/\mu(v)$. □

Suppose $\epsilon(k) > \frac{C}{k}$ where C is given in the statement of Lemma 4. Consider the program (19) where we fix λ^* and impose our desired condition on q:

$$\max_{\Xi:X(k)\to\mathbb{R},\, q:X(k)\to\mathbb{R}_+^N} \sum_{x\in X(k)} \rho(x)\Xi(x) + \sum_{v\in V}\mu(v)\lambda^*(v)$$
$$\text{s.t. } \Xi(x) + \lambda^*(v) \leqslant v\cdot\nabla^- q(x)\ \forall v, x;$$
$$\nabla_i^- q(x) = \begin{cases} kq_i(x) & \text{if } x_i = 0;\\ k(q_i(x) - q_i(x_i - 1/k, x_{-i})) & \text{if } 0 < x_i < k;\\ q_i(x) - q_i(x_i - 1/k, x_{-i}) & \text{if } x_i = k,\end{cases} \quad \forall i, x; \tag{35}$$
$$\sum_{i=1}^N q_i(x) \leqslant 1\ \forall x;$$
$$q_i(x_i - 1/k, x_{-i}) - q_i(x) \leqslant \epsilon(k)\ \forall i, x \text{ such that } 0 < x_i < k.$$

Lemma 4 follows if we show that programs (35) and (19) have the same value. The dual of (35) is:

$$\min_{\substack{\gamma:X(k)\to\mathbb{R}_+,\,\sigma:X(k)\times V\to\mathbb{R}_+,\\ \zeta:X(k)\to\mathbb{R}_+^N}} \sum_x \gamma(x) - \sum_{v,x}\lambda^*(v)(\sigma(v,x) - \mu(v)) + \sum_{i,x}\zeta_i(x)\epsilon(k)$$
$$\text{s.t. } \gamma(x) \geqslant \begin{cases} \sum_{v\in V} v_i k[\sigma(x,v) - \sigma(x_i + 1/k, x_{-i}, v)] & \\ \quad +\zeta_i(x) - \zeta_i(x_i + 1/k, x_{-i}) & \text{if } x_i < k - 1/k;\\ \sum_{v\in V} v_i[k\sigma(x,v) - \sigma(x_i + 1/k, x_{-i}, v)] & \\ \quad +\zeta_i(x) & \text{if } x_i = k - 1/k;\\ \sum_{v\in V} v_i\sigma(x,v) & \text{if } x_i = k,\end{cases} \tag{36}$$
$$\forall i, x;$$
$$\zeta_i(0, x_{-i}) = 0\ \forall i, x_{-i};$$
$$\sum_{v\in V}\sigma(x,v) = \rho(x)\ \forall x,$$

where $\zeta_i(x)$ is the multiplier on the constraint $q_i(x_i - 1/k, x_{-i}) - q_i(x) \leqslant \epsilon(k)$. Let $(\gamma^*, \sigma^*, \zeta^*)$ be an optimal solution to (36). Let $\overline{v}_i = \max V_i$ and $\underline{v}_i = \min V_i$.

Lemma 8. *Suppose that $\mu(v) > 0$ for all $v \in V$. Then $\zeta_i^*(x) = 0$ for all i and $x \in X(k)$.*

Proof of Lemma 8. For the sake of contradiction, let x be a signal profile with the lowest Σx such that $\zeta_i^*(x) > 0$ for some i. Notice that by construction, $0 < x_i < k$. Let $w^*(x)$ be the interim expected values at x under σ^*:

$$w^*(x) = \frac{1}{\rho(x)}\sum_{v\in V} v\,\sigma^*(x,v).$$

Case 1: $w_i^*(x) < \overline{v}_i$.

In this case, there must exist a v such that $v_i < \overline{v}_i$ and $\sigma^*(x, v) > 0$. Fix such a v, and define

$$\overline{\sigma}(x', v') = \begin{cases} (1-\tau)\sigma^*(x', v) & \text{if } x' = x, v' = v; \\ \tau\sigma^*(x', v) + \sigma^*(x', v') & \text{if } x' = x, v' = (\overline{v}_i, v_{-i}); \\ \sigma^*(x', v') & \text{otherwise.} \end{cases}$$

Choose $\tau > 0$ such that

$$0 < \sum_{v' \in V} v'_i \overline{\sigma}(x, v')k - \sum_{v' \in V} v'_i \sigma^*(x, v')k = (\overline{v}_i - v_i)\tau\sigma^*(x, v)k \leqslant \zeta_i^*(x).$$

Set

$$\overline{\zeta}_i(x') = \begin{cases} \zeta_i^*(x) - (\overline{v}_i - v_i)\tau\sigma^*(x, v)k & \text{if } x' = x; \\ \zeta_i^*(x') & \text{otherwise,} \end{cases} \qquad \overline{\gamma}(x') = \gamma^*(x').$$

By construction, $(\overline{\gamma}, \overline{\sigma}, \overline{\zeta})$ is feasible for (36). Notice that

$$\sum_{v' \in V} \lambda^*(v')\overline{\sigma}(x, v') - \sum_{v' \in V} \lambda^*(v')\sigma^*(x, v') = (\lambda^*(\overline{v}_i, v_{-i}) - \lambda^*(v))\tau\sigma^*(x, v).$$

Therefore, the difference between the objectives of $(\gamma^*, \sigma^*, \zeta^*)$ and $(\overline{\gamma}, \overline{\sigma}, \overline{\zeta})$ in (36) is:

$$\left(\sum_{x'} \gamma^*(x') - \sum_{v', x'} \lambda^*(v')\sigma^*(v', x') + \sum_{i, x'} \zeta_i^*(x')\epsilon(k)\right) - \left(\sum_{x'} \overline{\gamma}(x') - \sum_{v', x'} \lambda^*(v')\overline{\sigma}(v', x') + \sum_{i, x'} \overline{\zeta}_i(x')\epsilon(k)\right)$$
$$= \epsilon(k)(\overline{v}_i - v_i)\tau\sigma^*(x, v)k + (\lambda^*(\overline{v}_i, v_{-i}) - \lambda^*(v))\tau\sigma^*(x, v) > 0,$$

since $\epsilon(k)k > (\lambda^*(\overline{v}_i, v_{-i}) - \lambda^*(v))/(\overline{v}_i - v_i)$ by Lemma 7. This is a contradiction.

Case 2: $w_i^*(x) = \overline{v}_i$ and $w_i^*(x_i - 1/k, x_{-i}) > \underline{v}_i$.

In this case, there must exist a v such that $v_i > \underline{v}_i$ and $\sigma^*(x_i - 1/k, x_{-i}, v) > 0$. Fix such a v, and define

$$\overline{\sigma}(x', v') = \begin{cases} (1-\tau)\sigma^*(x', v) & \text{if } x' = (x_i - 1/k, x_{-i}), v' = v; \\ \tau\sigma^*(x', v) + \sigma^*(x', v') & \text{if } x' = (x_i - 1/k, x_{-i}), v' = (\underline{v}_i, v_{-i}); \\ \sigma^*(x', v') & \text{otherwise.} \end{cases}$$

Choose $\tau > 0$ such that

$$0 < \sum_{v' \in V} v'_i \sigma^*(x_i - 1/k, x_{-i}, v')k - \sum_{v' \in V} v'_i \overline{\sigma}(x_i - 1/k, x_{-i}, v')k = (v_i - \underline{v}_i)\tau\sigma^*(x_i - 1/k, x_{-i}, v)k \leqslant \zeta_i^*(x).$$

Set

$$\overline{\zeta}_i(x') = \begin{cases} \zeta_i^*(x) - (v_i - \underline{v}_i)\tau\sigma^*(x_i - 1/k, x_{-i}, v)k & \text{if } x' = x; \\ \zeta_i^*(x') & \text{otherwise,} \end{cases} \qquad \overline{\gamma}(x') = \gamma^*(x').$$

As in Case 1, $(\overline{\gamma}, \overline{\sigma}, \overline{\zeta})$ is feasible for (36) and has a strictly lower objective than $(\gamma^*, \sigma^*, w^*, \zeta^*)$, a contradiction.

Case 3: $w_i^*(x) = \overline{v}_i$ and $w_i^*(x_i - 1/k, x_{-i}) = \underline{v}_i$.

The virtual value at $(x_i - 1/k, x_{-i})$ is $\underline{v}_i - (k-1)(\overline{v}_i - \underline{v}_i) < 0$ when k is sufficiently large. Since $\gamma^*(x_i - 1/k, x_{-i}) \geqslant 0$ and $\zeta_i^*(x_i - 1/k, x_{-i}) = 0$, we must have

$$\gamma^*(x_i - 1/k, x_{-i}) > \sum_{v' \in V} v_i' k[\sigma^*(x_i - 1/k, x_{-i}, v') - \sigma^*(x, v')] + \zeta_i^*(x_i - 1/k, x_{-i}) - \zeta_i^*(x),$$

so we can decrease $\zeta_i^*(x)$ to lower strictly the objective of $(\gamma^*, \sigma^*, \zeta^*)$, a contradiction. $\square$

Lemma 8 implies that program (35) has the same value even if we drop the constraints (20). This implies Lemma 4.