Proper Robustness and the Efficiency of Monopoly Screening*

Ashwin Kambhampati†

April 12, 2025

Abstract

This paper proposes a refinement of the maxmin criterion for robust mechanism design and characterizes its implications in a canonical monopoly screening environment. In the model, a seller produces quality-differentiated goods to sell to a buyer of unknown payoff type. A mechanism is *properly robust* if it lexicographically maximizes expected profit with respect to a lexicographic probability system capturing first-order and higher-order uncertainty aversion about the buyer's type. It is shown that a mechanism is properly robust if and only if it is efficient and revenue maximizing. That is, asymmetric information does not lead to economic inefficiency.

Keywords: robust mechanism design, screening, equilibrium refinements, lexicographic probabilities.

JEL codes: D81, D82.

---

*First circulated: October 11, 2024. The views expressed here are solely my own and do not in any way represent the views of the U.S. Naval Academy, U.S. Navy, or the Department of Defense.

†Department of Economics, United States Naval Academy; `kambhamp@usna.edu`.

1 Introduction

Many markets are characterized by asymmetric information. For instance, a consumer may know its willingness-to-pay for various cellphone features better than the company selling the phone. A central finding in the Bayesian analysis of such markets is that asymmetric information leads to economic inefficiency (Akerlof (1970), Spence (1973), and Rothschild and Stiglitz (1976)). For instance, the cellphone company may degrade the quality of its lower-tier cellphones in order to make them less attractive to high-value consumers, thereby allowing the company to increase the price of its premium cellphones without affecting their demand (Mussa and Rosen (1978)).

Unfortunately, in the canonical Bayesian screening model, optimal allocations and prices are sensitive to the assumed prior distribution over buyer valuations (types). A common way to address this critique is to assume that the seller chooses a selling mechanism that performs well regardless of the environment she faces. That is, she is assumed to choose a "maxmin" (or, "robust") mechanism that maximizes her minimum payoff across all distributions over buyer types.[1]

The maxmin criterion, however, is too weak in many applications. Consider, for instance, the decision problem in Figure 1a; the decision-maker chooses a row and each column corresponds to a state (L, C, or R). Observe that U and D are maxmin optimal strategies; each ensures a payoff of 1, whereas M only ensures a payoff of 0. However, U weakly dominates D (U yields at least the same payoff as D in all states and, in some state, yields a strictly higher payoff). So, D is maxmin optimal, but violates the principle of admissibility.[2]

More broadly, even if dominated strategies are ruled out by assumption, the decision-maker is permitted to abandon her supposedly conservative attitude

[1]See Carroll (2019) for a survey of the rapidly growing literature on robust mechanism design.

[2]The principle of admissibility has a long heritage in decision and game theory (see, e.g., Arrow (1951) and Luce and Raiffa (1957)). Börgers (2017) argues that admissibility should be a minimal requirement for "robustness" in maxmin mechanism design problems.

	L	C	R
U	1	3	4
M	0	6	6
D	1	2	3

(a) *Inadmissibility.*

	L	C	R
U	1	3	4
M	0	6	6
D	1	2	5

(b) *Lexicographic inconsistency.*

Figure 1: *Properties of maxmin optimal strategies.*

towards uncertainty outside of worst-case scenarios. Consider, for instance, the decision problem in Figure 1b. U and D, again, are the only maxmin optimal strategies. But, the payoffs have been modified so that both are (weakly) undominated. Nevertheless, if the decision-maker strictly prefers U and D to M, shouldn't she strictly prefer U to D? U ensures a payoff of 3 outside of state L, whereas D only ensures a payoff of 2. Such considerations are not captured by the maxmin criterion.[3]

This paper proposes a refinement of the maxmin criterion for robust mechanism design that addresses several of its criticisms, and sharply characterizes its predictions in a canonical screening environment. It is shown that, if the seller uses this refinement, then asymmetric information does not lead to *any* loss in social surplus. The contribution is thus two-fold — a modeling framework for robust mechanism design and a substantial economic prediction.

The model is briefly outlined here, so that the refinement can be more precisely defined. As in Mussa and Rosen (1978), a profit-maximizing seller can produce costly quality-differentiated goods to sell to a buyer of unknown type. The seller knows the set of possible types of the buyer, but not its distribution. She commits to a (possibly stochastic) direct mechanism — a quality and price pair targeted to each buyer type. The buyer's payoff function satisfies strictly increasing differences and there is a unique efficient (surplus maximizing) allocation for each type.

The seller's attitude towards distributional uncertainty is characterized by the non-Archimedean extension of subjective expected utility proposed by

[3]This type of reasoning can be traced back to Dresher (1961) (Chapter 3, Section 18), who suggests that a decision-maker should maximize the (minimum) gain resulting from the "mistakes" of a strategic adversary.

Blume, Brandenburger, and Dekel (1991a). Specifically, the seller possesses a *lexicographic probability system (LPS)* — an ordered collection of prior distributions. The first prior distribution can be interpreted as the seller's primary belief, the second as her secondary belief, and so on. The seller ranks mechanisms lexicographically using the ordered vector of expected utilities corresponding to her LPS. For example, in Figure 1b, the decision-maker's first belief could be the point mass on L, her second belief could be the point mass on C, and her third could be the point mass on R. Under such an LPS, U yields the vector of expected utilities $(1, 3, 4)$, whereas D yields $(1, 2, 5)$. So, U is strictly preferred to D (because $1 \geq 1$ and $3 > 2$).

The approach taken in the paper is to impose restrictions on the seller's LPS and study the implications of these restrictions on the structure of optimal mechanisms. An LPS is *adversarial* if the seller's first-order belief coheres with the maxmin criterion; given her mechanism, the seller believes her payoff is primarily determined by a worst-case distribution over types. For example, if the decision-maker chooses U in either Figure 1a or Figure 1b, then the decision-maker's primary belief in any adversarial LPS must be the point mass on L. Strengthening this property, an LPS is *strongly adversarial* if the seller's higher-order beliefs also reflect her conservative attitude towards uncertainty; any buyer type yielding the seller a strictly lower payoff than another first appears with positive probability in a lower-order prior distribution. For example, if the decision-maker chooses U in either Figure 1a or Figure 1b, then C must appear with positive probability before R. Finally, an LPS has *full support* if each buyer type occurs with positive probability in some probability distribution contained in the LPS. For example, given any strategy of the decision-maker in Figures 1a and 1b, L, C, and R must appear with positive probability in *some* belief contained in her LPS.

The analysis begins with two preliminary results. The first preliminary result, Theorem 1, completely characterizes *robust mechanisms*, each of which is a best-response to some adversarial LPS. Any robust mechanism prescribes an efficient allocation for the lowest type. But, there are few other restrictions on the form of robust mechanisms. For instance, robust mechanisms can be inadmissible. The second preliminary result, Theorem 2, completely characterizes *perfectly robust mechanisms*, each of which is a best-response to some adversarial LPS with full support. It is shown that the set of perfectly robust mechanisms is equivalent to the set of robust and admissible mechanisms. In settings with two types, the unique perfectly robust mechanism is the efficient and revenue maximizing mechanism. But, with more than two types, there is little discipline on the allocations targeted to "intermediate" types. Moreover, the higher-order beliefs that generate inefficiency for these types are inconsistent with the seller's lower-order beliefs.

The main result, Theorem 3, completely characterizes *properly robust mechanisms*, each of which is a best-response to some *strongly* adversarial LPS with full support. It is shown that a mechanism is properly robust if and only if it is efficient and revenue maximizing. The basic economic insight is that, if the seller is concerned about her payoff guarantee, then she is most concerned with maximizing profit from the lowest type. Hence, there is no distortion at the bottom. If she is next most concerned about the second lowest type, then this type should also receive an undistorted allocation because the lowest type's allocation has already been fixed. Inductive reasoning leads all types to receive an efficient allocation.

Section 4.1 formalizes the intuition behind the main result by exploiting an equivalence between lexicographic optimality and Bayesian optimality with respect to the limits of the "trembles" justifying proper equilibria (Myerson (1978)) in normal-form games. A "proper" sequence of trembles placing more weight on lower types than higher types is exhibited. In the limit of the sequence, inverse hazard rates vanish. Hence, quality distortions vanish in any corresponding sequence of Bayesian optimal mechanisms. Section 4.2 provides an alternative foundation for the efficient and revenue maximizing mechanism using a variation of the "leximin" criterion (Sen (1970), Rawls (1971), Mackenzie (2024)), which has been proposed as an alternative to the maxmin criterion in the literature on social choice theory. It is shown, however, that leximin optimality does not in general coincide with perfect or proper robustness; a leximin optimal strategy need not be a best-response to *any* full support LPS.

1.1 Related literature

The paper's contribution is now briefly related to the literatures on Bayesian screening, non-Bayesian mechanism design, and equilibrium refinements.

Relative to the literature on Bayesian screening, the model nests (discrete-type versions of) the product design model of Mussa and Rosen (1978), the nonlinear pricing model of Maskin and Riley (1984), and the single-agent case of the optimal auction model of Myerson (1981). A key finding in this literature is that there is "no distortion at top", i.e., the highest type receives an undistorted allocation, whereas the allocations of buyers with lower valuations can be distorted downwards. The results in this paper demonstrate that robustness considerations "reverse" this intuition; distorting the allocation of the lowest type is suboptimal under any robustly optimal mechanism and this property cascades upwards in any properly robust mechanism. Applied to optimal auction settings in which it is efficient to allocate an indivisible good to all buyer types, the results show that a posted price pinned down by the binding participation constraint of the lowest type is properly robust.

Other authors have studied non-Bayesian variants of canonical screening problems. Closest to the setting of the paper, Bergemann and Schlag (2011) and Carrasco, Farinha Luz, Kos, Messner, Monteiro, and Moreira (2018) study the problem of selling a single indivisible good to a buyer of unknown type.[4][5] Bergemann and Schlag (2011) assume that the seller has knowledge that the true distribution lies in some neighborhood of a baseline distribution, whereas Carrasco, Farinha Luz, Kos, Messner, Monteiro, and Moreira (2018) assume that the seller knows the first N moments of the distribution.[6] The main result

[4] Bergemann and Schlag (2008) and Bergemann and Schlag (2011) also study the form of mechanisms that minimize worst-case regret (Savage (1951)). A foundational issue with the regret minimization criterion is that it is dependent on irrelevant alternatives (Chernoff (1954)). Proper robustness does not suffer this drawback.

[5] Madarász and Prat (2017) study local preference uncertainty that may cause non-local incentive compatibility constraints to bind in an optimal mechanism. Che (2022) extends the setting of Carrasco, Farinha Luz, Kos, Messner, Monteiro, and Moreira (2018) to the case of multiple buyers.

[6] Carroll (2017) considers a multi-dimensional screening setting in which the marginal distributions over values are known, but not the joint distribution. His main result is that

in Bergemann and Schlag (2011) is that the seller chooses a posted price to maximize her Bayesian payoff against a worst-case distribution. Carrasco, Farinha Luz, Kos, Messner, Monteiro, and Moreira (2018) show that the seller's worst-case payoff is a linear combination of the known moments of the distribution over types. Theorem 1 coheres with Bergemann and Schlag (2011)'s result; if any distribution is possible, i.e., the seller entertains an arbitrarily large neighborhood around the known distribution, then the seller maximizes her payoff against the point mass on the lowest buyer valuation type. The purpose of reprising (a version of) this result in this paper, however, is to demonstrate that if the seller's uncertainty set is large, then maxmin predictions are weak. Bergemann and Schlag (2011) and Carrasco, Farinha Luz, Kos, Messner, Monteiro, and Moreira (2018) escape this issue by imposing assumptions on the seller's knowledge. This paper offers a complementary modeling approach.

Within the broader literature on robust mechanism design, the point of view taken in this paper is inspired by Börgers (2017). Börgers (2017) critiques the foundations for dominant strategy mechanisms provided by Chung and Ely (2007) in a setting featuring multiple agents on the grounds that such mechanisms are weakly dominated. Others have built upon this implicitly lexicographic approach; Dworczak and Pavan (2022) assume that an information designer primarily optimizes against a worst-case distribution. Among the set of worst-case optimal mechanisms, the designer chooses a mechanism that maximizes her payoff under a baseline distribution. In the language of this paper, Dworczak and Pavan (2022)'s approach corresponds to assuming the designer chooses a mechanism that is optimal with respect to an adversarial LPS containing exactly two prior distributions. This paper shows that fruitful insights arise if it is also required that the designer's LPS has full support and is lexicographically consistent with her uncertainty aversion.

Finally, the restrictions placed on the seller's LPS correspond to those used to characterize the strategies that survive tremble-based refinements in

bundling is suboptimal. Che and Zhong (2024) consider alternative ambiguity sets that lead to the optimality of bundling.

normal-form games (Blume, Brandenburger, and Dekel (1991b)). Indeed, the terms "perfect robustness" and "proper robustness" are directly inspired by the notions of (trembling hand) perfect equilibrium (Selten (1975)) and proper equilibrium (Myerson (1978)). A detailed discussion of this relationship is contained in Section 2.4. Predating Blume, Brandenburger, and Dekel (1991a) and Blume, Brandenburger, and Dekel (1991b), Dresher (1961) (Chapter 3, Section 18) proposes an approach to selecting strategies in two-player, zero-sum games that involves maximizing the (guaranteed) gain resulting from a strategic opponent's "mistakes". Van Damme (1983) observes that this procedure identifies proper equilibrium strategies in such games (see Chapter 3, Theorem 3.5.5). Choosing an optimal mechanism with respect to the criterion considered in this paper thus coheres with what would arise following the protocol of Dresher (1961), once the problem is formulated as a zero-sum game between the seller and an opponent who chooses the distribution over buyer types.

2 Model

2.1 Environment

There is a single seller (she) and a single buyer (he). The seller can produce a good of variable quality $q \in \mathcal{Q}$, where $\mathcal{Q} \subseteq \mathbb{R}_+$ is a compact set and $0 \in \mathcal{Q}$, at a cost determined by the continuous function $c : \mathcal{Q} \to \mathbb{R}_+$. The seller's ex post payoff from selling a good of quality $q \in \mathcal{Q}$ at a price of $p \in \mathbb{R}_+$ is given by the function $v : \mathcal{Q} \times \mathbb{R}_+ \to \mathbb{R}$, where

$$v(q, p) = p - c(q).$$

The buyer has private information about her payoff type $\theta \in \Theta$, where $\Theta := \{\theta_1, \ldots, \theta_N\} \subseteq \mathbb{R}_+$ and $\theta_1 < \cdots < \theta_N$. Specifically, the buyer's ex post payoff from purchasing a good of quality $q \in \mathcal{Q}$ at a price of $p \in \mathbb{R}_+$ is

$$u(q, \theta) - p,$$

where $u : \mathcal{Q} \times \Theta \to [0, M]$ for $M > 0$ is continuous and satisfies $u(0, \theta) = 0$ for all $\theta \in \Theta$. Moreover, u is non-decreasing in q and has strictly increasing differences: if $q' > q$ and $\theta' > \theta$, then

$$u(q', \theta') - u(q, \theta') > u(q', \theta) - u(q, \theta).$$

The buyer's payoff from not transacting with the seller is normalized to zero. Finally, the surplus function $s(\cdot, \theta_i) := u(\cdot, \theta_i) - c(\cdot)$ is strictly quasiconcave for each $\theta_i \in \Theta$.[7]

2.2 Mechanisms

The seller has commitment power and can sell different qualities at different prices. By the Revelation Principle, without loss of optimality, the seller can restrict attention to direct mechanisms. A **(direct) mechanism** is an allocation rule and transfer rule,

$$Q : \Theta \to \Delta(\mathcal{Q}) \quad \text{and} \quad P : \Theta \to \mathbb{R}_+,$$

that together are **incentive compatible** and **individually rational**: for all $\theta \in \Theta$,

$$Eu(Q(\theta), \theta) - P(\theta) \geq Eu(Q(\theta'), \theta) - P(\theta') \quad \text{for all } \theta' \in \Theta$$

[7] The assumptions that $\mathcal{Q}$ is compact and that u and v are continuous ensure that an efficient mechanism exists. The assumption that the surplus function is strictly quasiconcave ensures that there is a unique efficient mechanism and it is deterministic. If all such assumptions are dropped, then the proceeding arguments establish that any deterministic mechanism that is efficient and revenue maximizing mechanism is properly robust. Moreover, if the seller restricts herself to use deterministic mechanisms, then the set of properly robust mechanisms coincides with the set of efficient and revenue maximizing mechanisms (which may be empty or contain more than one element).

and

$$Eu(Q(\theta), \theta) - P(\theta) \geq 0,$$

where $\Delta(\mathcal{Q})$ is the space of Borel measures on $\mathcal{Q}$ and $Eu : \Delta(\mathcal{Q}) \times \Theta \to \mathbb{R}$ is the extension of u taking expectations over allocations. Let $\mathcal{M} \subset (\Delta(\mathcal{Q}) \times [0, M])^N$ denote the compact metric space of all mechanisms.[8]

It is worth remarking that it is without loss of optimality to restrict attention to deterministic transfer rules; any stochastic transfer to type θ can be replaced with its mean and yield the seller equivalent expected profit. Moreover, randomizing over mechanisms (the set $\mathcal{M}$) cannot improve the seller's payoff, i.e., there are no "hedging" advantages to randomization.[9] Hence, the restrictions on the mechanism space entail no loss of optimality.[10]

2.3 Lexicographic probability systems

The seller's attitude towards uncertainty coheres with the non-Archimedean variant of subjective expected utility proposed by Blume, Brandenburger, and Dekel (1991a). Specifically, the seller possesses a lexicographic probability system and ranks mechanisms using a corresponding vector of expected utilities.

A **lexicographic probability system (LPS)** is a vector of prior beliefs $\mu = (\mu_1, \dots, \mu_K)$, where K is a strictly positive integer and $\mu_k \in \Delta(\Theta)$ for each $k \in \{1, \dots, K\}$. Fixing an LPS $\mu = (\mu_1, \dots, \mu_K)$, each mechanism $(Q, P) \in \mathcal{M}$

[8] $\Delta(\mathcal{Q})$ is compact in the topology of weak convergence because $\mathcal{Q}$ is compact in the Euclidean topology. Moreover, $[0, M]$ is compact in the Euclidean topology. So, $(\Delta(\mathcal{Q}) \times [0, M])^N$ is compact in the corresponding product topology. The set of functions $(Q, P) \in (\Delta(\mathcal{Q}) \times [0, M])^N$ satisfying the incentive compatibility and individual rationality constraints form a closed subset of $(\Delta(\mathcal{Q}) \times [0, M])^N$. Hence, $\mathcal{M}$ is a closed subset of a compact metric space and, therefore, compact.

[9] Randomizing over allocation rules results in a compound lottery for each type. These lotteries can be reduced to a simple lottery. The reduced lottery yields the seller and the buyer the same payoffs as under the original randomization against any distribution over types. In robust principal-agent problems with technological uncertainty, this argument does not hold because randomization affects the set of uncertainty (see Kambhampati (2023) and Kambhampati, Toikka, and Vohra (2024)).

[10] Strausz (2006) presents an example in which randomization over allocations is strictly optimal in a Bayesian model, even under standard assumptions on u and c. That such stochastic allocation rules are not properly robust is therefore a result of this paper.

gives rise to a K-vector of expected payoffs. The seller's **k-th order payoff** from $(Q, P) \in \mathcal{M}$ is

$$\sum_{\theta\in\Theta} \mu_k(\theta) Ev(Q(\theta), P(\theta)),$$

where $Ev : \Delta(\mathcal{Q}) \times \mathbb{R} \to \mathbb{R}$ is the extension of v taking expectations over allocations. Then, a mechanism $(Q, P) \in \mathcal{M}$ is **lexicographically preferred** to the mechanism $(Q', P') \in \mathcal{M}$ if

$$\left(\sum_{\theta\in\Theta} \mu_k(\theta) Ev(Q(\theta), P(\theta))\right)_{k=1}^{K} \geq_L \left(\sum_{\theta\in\Theta} \mu_k(\theta) Ev(Q'(\theta), P'(\theta))\right)_{k=1}^{K},$$

where $\geq_L$ is the lexicographic order.[11] Finally, a mechanism $(Q, P) \in \mathcal{M}$ is **μ-optimal** if it is lexicographically preferred to any mechanism $(Q', P') \in \mathcal{M}$. If μ is a single-dimensional LPS, i.e., $\mu = (\rho)$ for some $\rho \in \Delta(\Theta)$, and $(Q, P) \in \mathcal{M}$ is μ-optimal, then it is simply written that (Q, P) is ρ-optimal.

Notice that, in the standard monopoly screening problem, the seller possesses a single prior belief about the type of the buyer. Such a belief is a special case of an LPS in which $K = 1$, corresponding to the setting in which the decision-maker is a subjective expected utility maximizer. A mechanism $(Q, P) \in \mathcal{M}$ is thus **Bayesian optimal** if it is μ-optimal with respect to *some* LPS μ.[12]

This paper is concerned with the role of restrictions on the seller's LPS and the implications of these restrictions on the form of lexicographically optimal mechanisms. To define these restrictions, it will be useful to introduce an order on types with respect to an LPS. Given an LPS μ and types $\theta, \theta' \in \Theta$, write $\theta >_\mu \theta'$ if

$$\min\{k \in \{1, \ldots, K\} : \mu_k(\theta) > 0\} < \min\{k \in \{1, \ldots, K\} : \mu_k(\theta') > 0\}.$$

[11] For $a, b \in \mathbb{R}^K$, $a \geq_L b$ if whenever $b_k > a_k$, there exists a $j < k$ such that $a_j > b_j$.

[12] The usual definition of Bayesian optimality is that there exists a prior $\rho \in \Delta(\theta)$ such that $(Q, P) \in \mathcal{M}$ is ρ-optimal. The point here is that, if $(Q, P) \in \mathcal{M}$ is μ-optimal for some $\mu = (\mu_1, \ldots, \mu_K)$, where $K > 1$, then it must be μ_1-optimal. Hence, the set of Bayesian optimal mechanisms in the usual sense is identical to the set of Bayesian optimal mechanisms defined here.

If $\theta >_\mu \theta'$, then θ is said to be "infinitely more likely" than θ'. The weak order is defined in the usual way: if it is not the case that $\theta' >_\mu \theta$, then $\theta \geq_\mu \theta'$.

Three properties will be of interest. First, an LPS is adversarial given her mechanism if the seller's first-order belief is that the type distribution minimizes her expected payoff.

Definition 1
*An LPS $\mu = (\mu_1, \ldots, \mu_K)$ is **adversarial** with respect to $(Q, P) \in \mathcal{M}$ if $\mu_1(\theta) > 0$ implies*

$$Ev(Q(\theta), P(\theta)) \leq Ev(Q(\theta'), P(\theta')) \quad \textit{for all } \theta' \in \Theta.$$

Second, an LPS is strongly adversarial given her mechanism if the seller's higher-order beliefs cohere with her lower-order beliefs; if type θ' is not infinitely more likely than θ, then the seller must receive a higher expected payoff when transacting with type θ' than with type θ.

Definition 2
*An LPS $\mu = (\mu_1, \ldots, \mu_K)$ is **strongly adversarial** with respect to $(Q, P) \in \mathcal{M}$ if, for all $\theta \in \Theta$,*

$$Ev(Q(\theta), P(\theta)) \leq Ev(Q(\theta'), P(\theta')) \quad \textit{for all } \theta' \in \Theta \textit{ with } \theta \geq_\mu \theta'.$$

Finally, an LPS has full support if each buyer type occurs with positive probability in some probability distribution contained in the LPS.

Definition 3
*An LPS $\mu = (\mu_1, \ldots, \mu_K)$ has **full support** if, for each $\theta \in \Theta$, there exists a $k \in \{1, \ldots, K\}$ such that $\mu_k(\theta) > 0$.*

Observe that if an LPS is strongly adversarial, then it is adversarial. To see why, suppose that μ is *not* adversarial. Then, there exist types $\theta, \theta' \in \Theta$ such that $\mu_1(\theta) > 0$, but

$$Ev(Q(\theta), P(\theta)) > Ev(Q(\theta'), P(\theta')).$$

Because $\theta \geq_\mu \theta'$, it follows that μ is not strongly adversarial. On the other hand, an LPS can be strongly adversarial, but not have full support. For instance, given any mechanism $(Q, P) \in \mathcal{M}$ satisfying $Q(\theta) = Q(\theta')$ and $P(\theta) = P(\theta')$ for all $\theta, \theta' \in \Theta$, the seller's payoff is constant in the buyer's type. So, the LPS with $K = 1$ that places probability one on the lowest type is strongly adversarial. But, it does not have full support.

2.4 Robust optimality

The following optimality criteria incorporate increasingly stronger restrictions on the seller's LPS.

Definition 4 (Robust optimality criteria)

1. *A mechanism $(Q, P) \in \mathcal{M}$ is* ***robust*** *if there exists an LPS, μ, that is adversarial with respect to (Q, P) and for which (Q, P) is μ-optimal.*
2. *A mechanism $(Q, P) \in \mathcal{M}$ is* ***perfectly robust*** *if there exists a full support LPS, μ, that is adversarial with respect to (Q, P) and for which (Q, P) is μ-optimal.*
3. *A mechanism $(Q, P) \in \mathcal{M}$ is* ***properly robust*** *if there exists a full support LPS, μ, that is strongly adversarial with respect to (Q, P) and for which (Q, P) is μ-optimal.*

Five remarks are in order regarding the optimality criteria:

1. Because any adversarial LPS with full support is an adversarial LPS, any perfectly robust mechanism is robust. Moreover, any properly robust mechanism is perfectly robust because any strongly adversarial LPS with full support is an adversarial LPS with full support.
2. If a mechanism is robust, then, by definition, it is a part of a saddle point in a fictitious zero-sum game between the seller and an adversarial "Nature" that chooses the distribution of types. An alternative formulation

of the maxmin criterion is sequential; given the seller's mechanism, one computes her infimum expected payoff over all distributions. A maxmin optimal mechanism is then defined as a mechanism that maximizes the seller's infimum payoff. Such a criterion is *more* permissive than the robustness criterion defined here. That is, a robustly optimal mechanism in the sense of this paper is maxmin optimal according to the sequential criterion, but the converse need not hold. Nevertheless, this distinction plays no role in the analysis of the model considered in this paper.

3. The restrictions on the seller's LPS sustaining a perfectly robust mechanism correspond to those that sustain a (trembling-hand) perfect equilibrium strategy in the fictituous game between the seller and Nature. Precisely, Blume, Brandenburger, and Dekel (1991b) show that, if a decision-maker's strategy is a part of a perfect equilibrium in a finite game, then she must play a strategy that is optimal with respect to a full support LPS whose first-order belief coheres with the other's chosen strategy (see Proposition 4). So, any robust mechanism that is *not* perfectly robust cannot be played in a perfect equilibrium of any game between the seller and Nature in which the seller's strategy space is discretized, but sufficiently rich.

4. The restrictions on the seller's LPS sustaining a properly robust mechanism correspond to those that sustain a proper equilibrium strategy in the fictituous game between the seller and Nature. Blume, Brandenburger, and Dekel (1991b) show that if a decision-maker's strategy is a part of a proper equilibrium in such a game, then she must play a strategy that is optimal with respect to a full support LPS whose higher-order beliefs respect the preferences of the other player (see Proposition 5). In the context of zero-sum games, the restriction to preference-respecting LPSs corresponds precisely to the requirement that the decision-maker play a best-response to a strongly adversarial belief. So, any perfectly robust mechanism that is *not* properly robust cannot be played in a proper equilibrium in any sufficiently rich discretized game between the

seller and Nature.

5. Perfection (properness) of a mechanism does not immediately guarantee that it is played in some perfect (proper) equilibrium. The reason is that no restrictions are placed on the perfection or properness of the distribution chosen by Nature in response to the seller's mechanism. This seems a natural relaxation given that Nature is an artificial player used to facilitate understanding of the seller's attitude towards uncertainty. It is left as a conjecture, however, that the strategy profile in which Nature chooses the point mass on θ_1 and the seller chooses an efficient and revenue maximizing mechanism is, in fact, a proper equilibrium under any of the three extensions of the concept defined for infinite games by Simon and Stinchcombe (1995).

3 Analysis

3.1 Preliminaries

Some useful definitions and results from the standard theory of monopoly screening are recorded here before proceeding to the main analysis (see, e.g., Chapter 6 of Vohra (2011)). Because $\mathcal{Q}$ is compact and $s(\cdot, \theta_i)$ is continuous and strictly quasiconcave for each $i = 1, \ldots, N$, there exists a unique efficient quality for each type: for each $i = 1, \ldots, N$,

$$\{q_i^*\} := \operatorname*{argmax}_{q \in \mathcal{Q}} \quad u(q, \theta_i) - c(q).$$

Hence, randomization is strictly inefficient, i.e., the unique solution to

$$\max_{Q \in \Delta(\mathcal{Q})} \quad Eu(Q, \theta_i) - Ec(Q),$$

where $Ec : \Delta(\mathcal{Q}) \to \mathbb{R}_+$ is the extension of c taking expectations over allocations, is the Dirac measure on q_i^*.

An allocation $Q(\theta_i) \in \Delta(\mathcal{Q})$ is **efficient for type** θ_i if $Q(\theta_i)$ is the Dirac measure on the efficient quality q_i^*. An allocation rule $Q : \Theta \to \Delta(\mathcal{Q})$ is **efficient** if it is efficient for all types $\theta \in \Theta$. Because u satisfies strictly increasing differences, the Monotone Selection Theorem (Milgrom and Shannon (1994)) ensures that efficient allocations are non-decreasing in type: $q_1^* \leq q_2^* \leq \cdots \leq q_N^*$. Hence, the unique efficient allocation rule is implementable (Rochet (1987)). Given an efficient allocation rule and any full support distribution over types, the unique expected revenue maximizing prices are characterized by the binding downward adjacent incentive compatibility constraints and the individual rationality constraint for the lowest type. Specifically, the mechanism $(Q, P) \in \mathcal{M}$ is **efficient and revenue maximizing** if and only if Q is efficient and P satisfies

$$P(\theta_1) = Eu(Q(\theta_1), \theta_1) \quad \text{and}$$
$$P(\theta_j) = Eu(Q(\theta_1), \theta_1) + \sum_{i=2}^{j} Eu(Q(\theta_i), \theta_i) - Eu(Q(\theta_{i-1}), \theta_i) \quad \text{for } j \in \{2, ..., N\}. \tag{1}$$

There is a well-known relationship between full-information optimality and efficiency. Let $\delta_i \in \Delta(\Theta)$ denote the measure that places probability one on buyer type θ_i. Then, $(Q, P) \in \mathcal{M}$ is δ_i-optimal if and only if it is efficient for type θ_i and the seller extracts full surplus,

$$P(\theta_i) = u(q_i^*, \theta_i).$$

Any such mechanism is said to be **full-information optimal for type θ_i**.

3.2 Robust mechanisms

The first result establishes that there are few restrictions placed on the seller's chosen mechanism if the only requirement is that she choose a best-response to an adversarial LPS; any mechanism that is "efficient at the bottom" and does not yield excessive rent to other types is robustly optimal.

Theorem 1 (Characterization of robust mechanisms)
A mechanism $(Q, P) \in \mathcal{M}$ is robust if and only if it is full-information optimal

for the lowest type, θ_1, and the seller attains the lowest payoff from the lowest type,

$$Ev(Q(\theta), P(\theta)) \geq Ev(Q(\theta_1), P(\theta_1)) \quad \textit{for all } \theta \in \Theta.$$

Proof. To prove sufficiency, suppose $(Q, P) \in \mathcal{M}$ is a mechanism for which $Ev(Q(\theta), P(\theta)) \geq Ev(Q(\theta_1), P(\theta_1))$ for all $\theta \in \Theta$. Then, the LPS $\mu = (\delta_1)$ is adversarial with respect to (Q, P). If, in addition, (Q, P) is full-information optimal for type θ_1, then it maximizes the seller's first-order payoff. Hence, (Q, P) is μ-optimal and, thus, robustly optimal.

To prove necessity, suppose that under $(Q, P) \in \mathcal{M}$ there exists a type $\theta > \theta_1$ such that $Ev(Q(\theta), P(\theta)) < Ev(Q(\theta_1), P(\theta_1))$. Then, under any LPS μ that is adversarial with respect to (Q, P), the seller's first-order payoff is strictly smaller than $Ev(Q(\theta_1), P(\theta_1))$. So, (Q, P) cannot be μ-optimal; the seller obtains a payoff of $Ev(Q(\theta_1), P(\theta_1))$ from the mechanism $(Q', P') \in \mathcal{M}$ under which $Q'(\theta) = Q(\theta_1)$ and $P'(\theta) = P(\theta_1)$ for all $\theta \in \Theta$.

It remains to establish the necessity of full-information optimality for type θ_1. Suppose $(Q, P) \in \mathcal{M}$ is not full-information optimal for type θ_1 and $Ev(Q(\theta), P(\theta)) \geq Ev(Q(\theta_1), P(\theta_1))$ for all $\theta \in \Theta$. Then, under any LPS μ that is adversarial with respect to (Q, P), the seller's first-order payoff is no larger than $Ev(Q(\theta_1), P(\theta_1))$. But, if (Q', P') sets $Q'(\theta)$ equal to the Dirac measure on q_1^* and $P'(\theta) = u(q_1^*, \theta_1)$ for all $\theta \in \Theta$, then $(Q', P') \in \mathcal{M}$ and $Ev(Q'(\theta), P'(\theta)) > Ev(Q(\theta_1), P(\theta_1))$ for all $\theta \in \Theta$. Hence, (Q', P') yields a strictly higher first-order payoff than (Q, P) against μ. It follows that (Q, P) is not lexicographically preferred to (Q', P') and, thus, cannot be robustly optimal. $\square$

A simple example illustrates the multiplicity of robustly optimal mechanisms.

Example 1. Suppose $\Theta = \{\theta_1, \theta_2\}$, $\mathcal{Q} = [0, \max(\Theta)]$, $u(q, \theta) = \theta q$, and $c(q) = \frac{1}{2}q^2$. Then, the full-information optimal quality for type θ_i is $q_i^* := \theta_i$. Theorem 1 establishes that a mechanism is robustly optimal if and only if $Q(\theta_1)$ is the Dirac measure on θ_1, $P(\theta_1) = u(q_1^*, \theta_1) = \theta_1^2$, the incentive compatibility constraints are satisfied,

$$0 \geq \theta_1 \mathbb{E}_{Q(\theta_2)}[q] - P(\theta_2) \quad \text{and} \quad \theta_2 \mathbb{E}_{Q(\theta_2)}[q] - P(\theta_2) \geq \theta_2\theta_1 - \theta_1^2,$$

and a profitability constraint is satisfied,

$$P(\theta_2) - \frac{1}{2}\mathbb{E}_{Q(\theta_2)}[q^2] = Ev(Q(\theta_2), P(\theta_2)) \geq Ev(Q(\theta_1), P(\theta_1)) = \frac{1}{2}\theta_1^2.$$

Notice that the quality for type θ_2 need not be efficient and can be distorted upwards or downwards: for any Dirac measure $Q(\theta_2)$ on $q_2 \in \mathbb{R}_+$ such that

$$q_2^2 - 2\theta_2 q_2 + 2\theta_2\theta_1 - \theta_1^2 \leq 0,$$

there exists a price $P(\theta_2) \in \mathbb{R}_+$ such that (Q, P) satisfies both the incentive compatibility and profitability constraint (e.g., set $P(\theta_2)$ so that the downward incentive compatibility constraint binds). For instance, if $\theta_1 = 1$ and $\theta_2 = 2$, then any Dirac measure on $q_2 \in [1, 3]$ is a part of a robustly optimal mechanism.

3.3 Perfectly robust mechanisms

One objection to the maxmin criterion is that it permits the use of weakly dominated mechanisms.

Definition 5
A mechanism $(Q', P') \in \mathcal{M}$ ***weakly dominates*** *the mechanism* $(Q, P) \in \mathcal{M}$ *if*

$$Ev(Q'(\theta), P'(\theta)) \geq Ev(Q(\theta), P(\theta)) \quad \text{for all } \theta \in \Theta$$

and the inequality is strict for some $\theta \in \Theta$.

This possibility is illustrated in the setting of Example 1.

Example 2. Return to the setting of Example 1. The unique efficient and revenue maximizing mechanism, $(Q^*, P^*) \in \mathcal{M}$, sets $Q^*(\theta_1)$ equal to the Dirac measure on $q_1^* = \theta_1$, $Q^*(\theta_2)$ equal to the Dirac measure on $q_2^* = \theta_2$, $P^*(\theta_1) = p_1^* := \theta_1^2$, and $P^*(\theta_2) = p_2^* := \theta_2^2 - \theta_1^2$. It is shown that any robust mechanism that is *not* equal to $(Q^*, P^*) \in \mathcal{M}$ is weakly dominated by $(Q^*, P^*) \in \mathcal{M}$. Recall from Theorem 1 that any robust mechanism must coincide with the efficient and revenue maximizing mechanism for type θ_1. Hence, for any robust mechanism $(Q, P) \in \mathcal{M}$,

$$Ev(Q^*(\theta_1), P^*(\theta_1)) = Ev(Q(\theta_1), P(\theta_1)).$$

Because $N = 2$, the weak dominance claim is thus equivalent to the claim that, for any robust mechanism $(Q, P) \neq (Q^*, P^*)$,

$$Ev(Q^*(\theta_2), P^*(\theta_2)) > Ev(Q(\theta_2), P(\theta_2)).$$

It suffices to show that (q_2^*, p_2^*) uniquely maximizes the seller's payoff from θ_2 subject to the downward incentive compatibility constraint. That is, it suffices to show that (q_2^*, p_2^*) solves the relaxed problem[13]

$$\begin{aligned} &\max_{q_2, p_2} \quad p_2 - \frac{1}{2} q_2^2 \\ &\text{subject to} \\ &u(q_2, \theta_2) - p_2 \geq u(q_1^*, \theta_2) - p_1^*. \end{aligned}$$

In any solution, the incentive compatibility constraint binds, yielding $p_2 = u(q_2, \theta_2) + p_1^* - u(q_1^*, \theta_2)$. Eliminating constants from the objective function, it follows that any optimal quality must solve

$$\max_{q_2} \quad \theta_2 q_2 - \frac{1}{2} q_2^2.$$

The unique solution is $q_2 = \theta_2$, which yields $p_2 = \theta_2^2 - \theta_1^2$. So, (q_2^*, p_2^*) uniquely maximizes the seller's payoff from θ_2.

In settings such as Example 2, Börgers (2017) argues that the seller should select a robust mechanism that is also *admissible*. Indeed, admissibility is

[13]Stochastic mechanisms are strictly suboptimal in the example given the linearity of u and the strict convexity of c.

often imposed as a choice axiom in decision theory.[14]

Definition 6
The mechanism $(Q, P) \in \mathcal{M}$ *is* ***admissible*** *if there does not exist a mechanism* $(Q', P') \in \mathcal{M}$ *that weakly dominates it.*[15]

The following Theorem establishes that requiring the seller to choose an optimal mechanism against *some* full support and adversarial LPS is equivalent to requiring the seller to choose a robust and admissible mechanism.

Theorem 2 (Characterization of perfectly robust mechanisms)
A mechanism $(Q, P) \in \mathcal{M}$ *is perfectly robust if and only if it is robust and admissible.*

Proof. It is first proven that if a mechanism $(Q, P) \in \mathcal{M}$ is perfectly robust, then it is robust and admissible. Suppose $(Q, P) \in \mathcal{M}$ is perfectly robust. Then, by definition, there exists an LPS, μ, that is adversarial with respect to (Q, P) and such that (Q, P) is μ-optimal. Hence, it is robust. To prove that such a mechanism is admissible, observe that if $(Q, P) \in \mathcal{M}$ is not admissible, then there exists a mechanism $(Q', P') \in \mathcal{M}$ that weakly dominates it. Hence, (Q', P') is lexicographically preferred to (Q, P) given any full support LPS, μ. It follows that (Q, P) cannot be a best-response to any full support LPS and, hence, cannot be perfectly robust.

To prove sufficiency, suppose that $(Q, P) \in \mathcal{M}$ is robust. By Theorem 1,

$$Ev(Q(\theta_1), P(\theta_1)) \leq Ev(Q(\theta), P(\theta)) \quad \text{for all } \theta \in \Theta.$$

So, any LPS, μ, with $\mu_1 = \delta_1$ is adversarial with respect to (Q, P). Moreover, because $(Q(\theta_1), P(\theta_1))$ is full-information optimal for type θ_1, for any $(Q', P') \in \mathcal{M}$,

$$Ev(Q'(\theta_1), P'(\theta_1)) \leq Ev(Q(\theta_1), P(\theta_1)).$$

[14] See, e.g., Luce and Raiffa (1957) p. 287, Axiom 5 and p. 77 for an argument for its imposition in the particular context of two-player, zero-sum games.

[15] If $\sigma \in \Delta(\mathcal{M})$ dominates $(Q, P) \in \mathcal{M}$, then there exists a pure strategy $(Q', P') \in \mathcal{M}$ that dominates (Q, P) by the reduction argument in footnote 9. So restricting attention to pure strategy domination in the definition of admissibility is without loss of generality.

So, no other mechanism obtains a higher first-order payoff than (Q, P) against any LPS with $\mu_1 = \delta_1$.

If, in addition, $(Q, P) \in \mathcal{M}$ is admissible, then there exists a full support distribution $\rho \in \Delta(\Theta)$ such that, for any $(Q', P') \in \mathcal{M}$,

$$\sum_{\theta \in \Theta} \rho(\theta) Ev(Q(\theta), P(\theta)) \geq \sum_{\theta \in \Theta} \rho(\theta) Ev(Q'(\theta), P'(\theta)). \tag{2}$$

The proof of this claim is immediate from a standard linear programming duality argument if $\mathcal{M}$ is assumed to be finite (see, e.g., Myerson (1991) Theorem 1.7). Appendix A shows that the claim continues to hold even though $\mathcal{M}$ is infinite. Thus, $(Q, P) \in \mathcal{M}$ is μ-optimal for $\mu = (\delta_1, \rho)$. Because μ has full support and is adversarial with respect to (Q, P), it follows that (Q, P) is perfectly robust. $\square$

An immediate implication of Theorem 2 is that, in Example 2 (and any setting with two types), the unique perfectly robust mechanism is the efficient and revenue maximizing mechanism. Observe that in Example 2 the efficient and revenue maximizing mechanism is robust and admissible. If it were to be weakly dominated, then it must be dominated by a robust mechanism — if a mechanism is *not* robust, then it attains a strictly lower payoff from θ_1. Moreover, it was shown that the efficient and revenue maximizing mechanism attains a strictly higher payoff from type θ_2 than any other robust mechanism.

With more than two types, perfectly robust mechanisms need not be efficient. The following example — a minimal departure from Example 2 — demonstrates that "intermediate" types can be maximally distorted subject to monotonicity constraints on the allocation rule.[16]

Example 3. Suppose $\Theta = \{\theta_1, \theta_2, \theta_3\}$, $\mathcal{Q} = [0, \max(\Theta)]$, $u(q, \theta) = \theta q$, and $c(q) = \frac{1}{2}q^2$. Consider an allocation in which both $Q(\theta_1)$ and $Q(\theta_2)$ are the Dirac measure on $q_1^* = \theta_1$ and $Q(\theta_3)$ is the Dirac measure on $q_3^* = \theta_3$; notice that the allocation to the middle type is completely distorted subject to monotonicity

[16] If the seller restricts herself to use deterministic mechanisms, then it is possible to show that any such mechanism entails efficiency for types θ_1 and θ_N. However, all other types may receive a distorted allocation.

of the allocation rule. Moreover, suppose that prices are determined by the individual rationality constraint for the lowest type and the downward adjacent incentive compatibility constraints:

$$
\begin{aligned}
&P(\theta_1) = P(\theta_2) = u(q_1^*, \theta_1) \quad \text{and} \\
&P(\theta_3) = u(q_1^*, \theta_1) + u(q_3^*, \theta_3) - u(q_1^*, \theta_3).
\end{aligned}
$$

It is easy to verify that $(Q, P) \in \mathcal{M}$. In addition, the LPS $\mu = (\delta_1, \delta_3, \delta_2)$ has full support and is adversarial with respect to $(Q, P) \in \mathcal{M}$ because the seller's payoff is increasing in the buyer's type.

It remains to show that $(Q, P) \in \mathcal{M}$ is μ-optimal and, therefore, perfectly robust. Observe first that $(Q(\theta_1), P(\theta_1))$ uniquely attains the highest possible payoff against δ_1. Second, from standard constraint simplification arguments, any mechanism that is full-information optimal for the lowest type cannot attain a payoff against δ_3 higher than the value of the following relaxed problem:

$$
\begin{aligned}
&\max_{(q_2,p_2),(q_3,p_3)} \quad p_3 - \frac{1}{2}q_3^2 \\
&\text{subject to} \\
&u(q_3, \theta_3) - p_3 \geq u(q_2, \theta_3) - p_2 \\
&u(q_2, \theta_2) - p_2 \geq u(q_1^*, \theta_2) - u(q_1^*, \theta_1) \\
&q_3 \geq q_2 \geq q_1^*.
\end{aligned}
$$

It is immediate that, in any solution, the second incentive constraint binds. If not, then p_2 can be strictly increased, slackening the first incentive constraint. Hence, p_3 can be strictly increased to strictly increase the objective function. So,

$$
p_2 = u(q_1^*, \theta_1) + u(q_2, \theta_2) - u(q_1^*, \theta_2)
$$

in any solution. The relaxed problem thus reduces to

$$\max_{q_2,(q_3,p_3)} \quad p_3 - \frac{1}{2}q_3^2$$

subject to

$$u(q_3,\theta_3) - p_3 \geq u(q_2,\theta_3) - u(q_2,\theta_2) - (u(q_1^*,\theta_1) - u(q_1^*,\theta_2))$$

$$q_3 \geq q_2 \geq q_1^*.$$

Notice that the incentive constraint must bind in any solution. Moreover, the right-hand side of the incentive constraint is minimized subject to $q_2 \geq q_1^*$ if and only if $q_2 = q_1^*$ because

$$u(q_2,\theta_3) - u(q_2,\theta_2) = (\theta_3 - \theta_2)q_2.$$

Hence, in any solution, it must be that $q_2 = q_1^*$ and

$$p_3 = u(q_1^*,\theta_1) + u(q_3,\theta_3) - u(q_1^*,\theta_3).$$

Substituting p_3 into the objective function and eliminating constants yields the surplus maximization problem

$$\max_{q_3} \quad p_3 - \frac{1}{2}q_3^2,$$

whose solution is $q_3^* = \theta_3$. It follows that $(Q,P) \in \mathcal{M}$ uniquely maximizes the seller's second-order payoff (it is uniquely optimal in the relaxed problem and is feasible in the constrained problem). Therefore, it is optimal against $(\delta_1,\delta_3,\delta_2)$.

3.4 Properly robust mechanisms

Are the beliefs sustaining the inefficient mechanism in Example 3 reasonable? If the seller were truly concerned with the robustness of her mechanism, then she might be relatively more concerned about a buyer type that is more harmful for her bottom line than a type that is less harmful. In particular, the LPS $\mu' = (\delta_1, \delta_2, \delta_3)$ seems more consistent with uncertainty aversion than $\mu = (\delta_1, \delta_3, \delta_2)$ because the seller obtains a strictly higher payoff from θ_3 than θ_2. The main result of the paper is that, when the seller must play a best-response to a full support and strongly adversarial LPS (such as $\mu' = (\delta_1, \delta_2, \delta_3)$ in Example 3), a sharp prediction arises: the seller's unique optimal mechanism is the efficient and revenue maximizing mechanism.

Theorem 3 (Characterization of properly robust mechanisms)
A mechanism is properly robust if and only if it is efficient and revenue maximizing.

Proof. To establish sufficiency, suppose $(Q, P) \in \mathcal{M}$ is the efficient and revenue maximizing mechanism. To prove that it is properly robust, consider the LPS $\mu = (\delta_1, \delta_2, \dots, \delta_N)$. By construction, μ has full support. Because the seller's payoff under (Q, P) is increasing in the buyer's type, μ is also strongly adversarial with respect to (Q, P). To show that (Q, P) is μ-optimal, proceed by induction on the vector of payoffs obtained from an arbitrary mechanism $(Q', P') \in \mathcal{M}$. For the base case, note that $(Q(\theta_1), P(\theta_1))$ is full-information optimal against δ_1. Hence, (Q', P') cannot obtain a higher first-order payoff than (Q, P) against μ. Now, suppose Q' is efficient for types $\theta_1, \dots, \theta_k$ and P' is revenue maximizing given Q' for types $\theta_1, \dots, \theta_k$. That is,

$$P'(\theta_1) = u(q_1^*, \theta_1) \quad \text{and}$$

$$P'(\theta_j) = u(q_1^*, \theta_1) + \sum_{i=2}^{j} u(q_i^*, \theta_i) - u(q_{i-1}^*, \theta_i) \quad \text{for } j \in \{2, ..., k\}.$$

Any mechanism in this class that maximizes the seller's $(k+1)$-th order payoff cannot yield a higher payoff than the solution to the relaxed problem

$$\begin{aligned}&\max_{Q\in\Delta(\mathcal{Q}),p\in\mathbb{R}} \quad p - Ec(Q)\\ &\text{subject to}\\ &Eu(Q,\theta_{k+1}) - p \geq \bar{u} := u(q_k^*,\theta_{k+1}) - P(\theta_k).\end{aligned} \tag{3}$$

In any solution to (3), p is determined by the binding constraint. Substituting the binding constraint into the objective function and eliminating the constant $\bar{u}$ yields the surplus maximization problem

$$\max_{Q\in\Delta(\mathcal{Q})} \quad Eu(Q,\theta_{k+1}) - Ec(Q),$$

whose value is no higher than what is attained under an efficient allocation for type θ_{k+1}. It follows that (Q',P') cannot attain a higher $(k+1)$-th order payoff than any mechanism that is efficient and revenue maximizing for types $\theta_1,\ldots,\theta_{k+1}$. The μ-optimality of (Q,P) follows from induction.

To establish necessity, let $\mathcal{M}'\supseteq\mathcal{M}$ denote the set of pairs (Q,P), where $Q:\Theta\to\Delta(\mathcal{Q})$ and $P:\Theta\to\mathbb{R}_+$, that are individually rational and satisfy the downward adjacent incentive compatibility constraints: for all $i=2,\ldots,N$,

$$Eu(Q(\theta_i),\theta_i) - P(\theta_i) \geq Eu(Q(\theta_{i-1}),\theta_i) - P(\theta_{i-1}).$$

It is shown that if a mechanism is properly robust in the relaxed mechanism space $\mathcal{M}'$ (i.e., there exists a full support LPS μ that is strongly adversarial with respect to the mechanism and the mechanism is lexicographically preferred to any other in $\mathcal{M}'$), then it is the efficient and revenue maximizing mechanism. Because the efficient and revenue maximizing mechanism is feasible in $\mathcal{M}$, the desired result follows.

Now, fix a properly robust mechanism $(Q,P)\in\mathcal{M}'$. It is first shown that $Ev(Q(\theta),P(\theta))$ must be non-decreasing in $\theta\in\Theta$. Suppose, towards contradiction, that $Ev(Q(\theta),P(\theta))$ is *not* non-decreasing. Then, there exists an integer $j\in\{1,\ldots,N-1\}$ such that $Ev(Q(\theta_{j+1}),P(\theta_{j+1})) < Ev(Q(\theta_j),P(\theta_j))$.

Let J denote the smallest such integer. Now, consider the pair of functions (Q', P') that is identical to (Q, P) for types $\theta_1, \ldots, \theta_J$, but sets $Q'(\theta) = Q(\theta_J)$ and $P'(\theta) = P(\theta_J)$ for each $\theta \in \{\theta_{J+1}, \ldots, \theta_N\}$. Notice that (Q', P') is individually rational and satisfies all downward adjacent incentive compatibility constraints, i.e., $(Q', P') \in \mathcal{M}'$. In addition, (Q, P) is not lexicographically preferred to (Q', P') against any full support LPS that is strongly adversarial with respect to (Q, P). To see why, fix such an LPS, μ, and let $\kappa := \min\{k : \mu_k(\theta_{J+1}) > 0\}$. For any type $\theta \leq \theta_J$, $Ev(Q'(\theta), P'(\theta)) = Ev(Q(\theta), P(\theta))$. In addition, if $\theta > \theta_J$ and $\mu_k(\theta) > 0$ for some $k \in \{1, \ldots, \kappa\}$, then $Ev(Q(\theta), P(\theta)) \leq Ev(Q(\theta_{J+1}), P(\theta_{J+1})) < Ev(Q(\theta_J), P(\theta_J)) = Ev(Q'(\theta), P'(\theta))$. So, by $Ev(Q(\theta_{J+1}), P(\theta_{J+1})) < Ev(Q'(\theta_{J+1}), P'(\theta_{J+1}))$,

$$\sum_{\theta \in \Theta} \mu_\kappa(\theta) Ev(Q(\theta), P(\theta)) < \sum_{\theta \in \Theta} \mu_\kappa(\theta) Ev(Q'(\theta), P'(\theta))$$

and, for any strictly positive integer $k < \kappa$,

$$\sum_{\theta \in \Theta} \mu_k(\theta) Ev(Q(\theta), P(\theta)) \leq \sum_{\theta \in \Theta} \mu_k(\theta) Ev(Q'(\theta), P'(\theta)).$$

It follows that (Q, P) is not lexicographically preferred to (Q', P'). Hence, it is not properly robust.

Observe now that if $(Q, P) \in \mathcal{M}'$ is properly robust, then $Q(\theta_1)$ is the Dirac measure on q_1^* and $P(\theta_1) = u(q_1^*, \theta_1)$ (because any properly robust mechanism is robust and, therefore, full-information optimal for type θ_1 by the argument in Theorem 1). It is shown that if $Q(\theta_i)$ is the Dirac measure on q_i^* and $P(\theta_i)$ is pinned down by (1) for types $\theta \in \{\theta_1, \ldots, \theta_k\}$, then $Q(\theta_{k+1})$ is the Dirac measure on q_{k+1}^* and $P(\theta_{k+1})$ is pinned down by (1) in any properly robust mechanism $(Q, P) \in \mathcal{M}'$. Suppose, towards contradiction, that the implication does not hold. Consider the mechanism $(Q', P') \in \mathcal{M}'$ that coincides with $(Q, P) \in \mathcal{M}'$ for types $\theta \in \{\theta_1, \ldots, \theta_k\}$, but which is efficient and revenue maximizing. Then, under $(Q', P') \in \mathcal{M}'$, $Ev(Q'(\theta_i), P'(\theta_i)) = Ev(Q(\theta_i), P(\theta_i))$ for $i = 1, \ldots, k$ and $Ev(Q'(\theta_i), P'(\theta_i)) > Ev(Q(\theta_{k+1}), P(\theta_{k+1}))$ for $i = k + 1, \ldots, N$. If $Ev(Q(\theta_{k+1}), P(\theta_{k+1})) < Ev(Q(\theta_{k+2}), P(\theta_{k+2}))$, then let $J =$

$k+1$. Otherwise, let $J \in \{k+2, \ldots, N\}$ be the largest integer such that $Ev(Q(\theta_{k+1}), P(\theta_{k+1})) = Ev(Q(\theta_{k+2}), P(\theta_{k+2})) = \cdots = Ev(Q(\theta_J), P(\theta_J))$. If $(Q, P) \in \mathcal{M}'$ is properly robust, then $Ev(Q(\theta), P(\theta))$ is non-decreasing in θ by the previous paragraph. So, for any type $\theta > \theta_J$, $\theta <_\mu \theta_j$ for all $j = 1, 2, \ldots, J$ under any full support LPS μ that is strongly adversarial with respect to (Q, P). Because $Ev(Q'(\theta_i), P'(\theta_i)) = Ev(Q(\theta_i), P(\theta_i))$ for $i = 1, \ldots, k$ and $Ev(Q'(\theta_i), P'(\theta_i)) > Ev(Q(\theta_i), P(\theta_i))$ for $i = k+1, \ldots, J$, it follows that (Q, P) is not lexicographically preferred to (Q', P') under μ. In particular, for $\kappa := \min\{\ell : \mu_\ell(\theta_{k+1}) > 0\}$,

$$\sum_{\theta \in \Theta} \mu_\kappa(\theta) Ev(Q(\theta), P(\theta)) < \sum_{\theta \in \Theta} \mu_\kappa(\theta) Ev(Q'(\theta), P'(\theta))$$

and, for any strictly positive integer $k < \kappa$,

$$\sum_{\theta \in \Theta} \mu_k(\theta) Ev(Q(\theta), P(\theta)) \leq \sum_{\theta \in \Theta} \mu_k(\theta) Ev(Q'(\theta), P'(\theta)).$$

Hence, (Q, P) is not properly robust. The proof of necessity follows from induction. $\square$

To conclude the analysis of Example 3, observe that the unique properly robust mechanism is the efficient and revenue maximizing mechanism; it is μ-optimal with respect to the full support and strongly adversarial LPS $\mu = (\delta_1, \delta_2, \delta_3)$. The logic is simple: against δ_1, any μ-optimal mechanism must have $Q(\theta_1)$ equal to the Dirac measure on q_1^* with prices extracting full surplus. Within the set of mechanisms with this property, any optimal mechanism against δ_2 must set $Q(\theta_2)$ equal to the Dirac measure on q_2^*, with prices extracting as much surplus as possible subject to the downward adjacent incentive compatibility constraint. Repeating the argument for the highest type establishes the result.

4 Alternative foundations

4.1 Bayesian foundations

The proof of Theorem 3 establishes that any properly robust mechanism is a best-response to the LPS $\mu = (\delta_1, \delta_2, \ldots, \delta_N)$. By construction, it is therefore Bayesian optimal against δ_1. Notice, however, that *any* mechanism that is full-information optimal for type θ_1 is Bayesian optimal against δ_1. For simplicity of exposition, suppose for now that the seller restricts herself to using deterministic mechanisms. Then, a (deterministic) allocation rule is a part of a ρ-optimal mechanism if and only if it solves

$$\begin{aligned}
&\max_{Q:\Theta\to\mathcal{Q}} \quad \sum_{i=1}^{N} \rho(\theta_i)\left[u(Q(\theta_i),\theta_i) - c(Q(\theta_i)) - d(\theta_i)\right] \\
&\text{subject to} \\
&d(\theta_i) = \begin{cases} h(\theta_i)\left(u(Q(\theta_i),\theta_{i+1}) - u(Q(\theta_i),\theta_i)\right) & \text{if } \rho(\theta_i) > 0 \\ 0 & \text{otherwise} \end{cases} \\
&Q(\cdot) \text{ non-decreasing,}
\end{aligned} \tag{4}$$

where $h(\theta_i) = \sum_{j=i+1}^{n} \rho(\theta_j)/\rho(\theta_i)$ is the inverse hazard rate for type θ_i and $d(\theta_i)$ is a term that distorts the allocation given to a buyer of type θ_i. Notice, for any full-support prior distribution, quality distortions must arise for all types below θ_N and these distortions depend on the shape of the corresponding inverse hazard rates. However, for $\rho = \delta_1$, there are no restrictions on the form of the optimal mechanism other than full-information optimality for the lowest type (because no other type appears in the objective function).

Due to the relationship between LPSs and the vanishing "trembles" used in the literature on equilibrium refinements, there is a precise sense in which properly robust mechanisms outperform others in the limit of particular sequences of Bayesian priors converging to δ_1. Specifically, given any LPS $\mu = (\mu_1, \ldots, \mu_K)$ and vector $r \in (0,1)^{K-1}$, define a probability measure on

Θ by the nested convex combination

$$r\Box\mu := (1-r_1)\mu_1 +$$

$$r_1\left[(1-r_2)\mu_2 + r_2\left[(1-r_3)\mu_3 + r_3\left[\cdots + r_{K-2}\left[(1-r_{K-1})\mu_{K-1} + r_{K-1}\mu_K\right]\cdots\right]\right]\right].$$

Now, suppose $(Q,P) \in \mathcal{M}$ is a properly robust mechanism and $(Q',P') \in \mathcal{M}$ is *not* properly robust. It is shown that, if $\mu = (\delta_1,\ldots,\delta_N)$ and $(r(\ell))_\ell \in ((0,1)^{K-1})^{\mathbb{N}}$ converges to the zero vector, then (Q,P) yields a higher expected payoff for the seller than (Q',P') along the tail of the sequence of prior beliefs $(r(\ell)\Box\mu)_\ell$. Moreover, a converse result holds.

Corollary 1 (Bayesian foundations)
Let $\mu = (\delta_1,\ldots,\delta_N)$. A mechanism $(Q,P) \in \mathcal{M}$ is properly robust if and only if for any mechanism $(Q',P') \in \mathcal{M}$ that is not properly robust and any sequence $(r(\ell))_\ell \in ((0,1)^{K-1})^{\mathbb{N}}$ with $r(\ell) \to (0,0,\ldots,0)$, there exists an $L \in \mathbb{N}$ such that

$$\sum_{\theta\in\Theta}(r(\ell)\Box\mu)(\theta)Ev(Q(\theta),P(\theta)) > \sum_{\theta\in\Theta}(r(\ell)\Box\mu)(\theta)Ev(Q'(\theta),P'(\theta)) \quad \textit{for all } \ell \geq L.$$

Proof. Immediate from the sufficiency direction of Theorem 3 and Proposition 4 of Mailath, Samuelson, and Swinkels (1997), which itself is almost immediate from Proposition 1 of Blume, Brandenburger, and Dekel (1991b). $\Box$

To better understand the intuition behind the result, suppose $|\Theta| = 3$ and consider a sequence $(r_1(\ell), r_2(\ell))_\ell \in ((0,1)^2)^{\mathbb{N}}$ for which $(r_1(\ell), r_2(\ell)) \to (0,0)$. Then, for $\mu = (\delta_1,\delta_2,\delta_3)$, the Bayesian prior $(r(\ell)\Box\mu)$ sets

$$\begin{aligned}
(r(\ell)\Box\mu)(\theta_1) &= 1 - r_1(\ell), \\
(r(\ell)\Box\mu)(\theta_2) &= r_1(\ell)\left(1 - r_2(\ell)\right), \quad \text{and} \\
(r(\ell)\Box\mu)(\theta_3) &= r_1(\ell) r_2(\ell).
\end{aligned}$$

The corresponding inverse hazard rates are

$$\begin{aligned} h_\ell(\theta_1) &= \frac{r_1(\ell)}{1 - r_1(\ell)}, \\ h_\ell(\theta_2) &= \frac{r_2(\ell)}{1 - r_2(\ell)}, \quad \text{and} \\ h_\ell(\theta_3) &= 0. \end{aligned}$$

If $r_1(\ell) \geq r_2(\ell)$ for each ℓ, then the inverse hazard rate is non-increasing in type and the monotonicity constraint on the allocation rule does not bind in the (deterministic) Bayesian screening problem (4). By Strausz (2006), restricting attention to deterministic mechanisms is thus without loss of optimality; there is a deterministic mechanism, (Q_ℓ, P_ℓ), where $Q_\ell : \Theta \to \mathcal{Q}$ and $P_\ell : \Theta \to \mathbb{R}_+$, that is $r(\ell)\square\mu$-optimal in the space of all mechanisms. Necessarily, it satisfies

$$Q_\ell(\theta_i) \in \underset{q \in \mathcal{Q}}{\operatorname{argmax}} \quad u(q, \theta_i) - c(q) - d_\ell(\theta_i),$$

where

$$d_\ell(\theta_i) = h_\ell(\theta_i) \left(u(q, \theta_{i+1}) - u(q, \theta_i) \right),$$

and P_ℓ satisfies (1) replacing $Q(\cdot)$ with $Q_\ell(\cdot)$ and eliminating expectations. Now, suppose (Q', P') is a deterministic mechanism that is *not* efficient and revenue maximizing. Then, because $\max_\theta h_\ell(\theta) \to 0$ as $\ell \to \infty$, the efficient and revenue maximizing mechanism, (Q, P), will attain expected profits closer to the $r(\ell)\square\mu$-optimal mechanism than (Q', P') for ℓ sufficiently large. This sketch formalizes the following "Bayesian" intuition: if lower value buyers are much more likely than higher value buyers, then inverse hazard rates converge to zero. Correspondingly, quality distortions converge to zero in any corresponding sequence of deterministic, Bayesian optimal mechanisms.

4.2 Leximin foundations

A related optimality criterion to proper robustness is the leximin criterion (Rawls (1971), Sen (1970)). Modifying Savage (1972)'s framework to allow for discontinuous preferences over subjective lotteries, Mackenzie (2024) recently showed that the leximin criterion reflects an axiom of "maximal risk aversion".[17] A definition of leximin optimality adapted to the setting studied in this paper is provided below.

Definition 7 (Leximin optimality)
*A mechanism $(Q, P) \in \mathcal{M}$ is **leximin preferred** to the mechanism $(Q', P') \in \mathcal{M}$ if there exist permutations $\pi : \{1, \ldots, N\} \to \{1, \ldots, N\}$ and $\pi' : \{1, \ldots, N\} \to \{1, \ldots, N\}$ such that*

1. *for any $i \in \{1, \ldots, N\}$ and $j \in \{1, \ldots, N\}$ satisfying $i \leq j$,*

$$Ev(Q(\theta_{\pi(i)}), P(\theta_{\pi(i)})) \leq Ev(Q(\theta_{\pi(j)}), P(\theta_{\pi(j)}))$$

   and

$$Ev(Q(\theta_{\pi'(i)}), P(\theta_{\pi'(i)})) \leq Ev(Q(\theta_{\pi'(j)}), P(\theta_{\pi'(j)}));$$

2. *and*

$$\left(Ev(Q(\theta_{\pi(k)}), P(\theta_{\pi(k)}))\right)_{k=1}^{N} \geq_L \left(Ev(Q'(\theta_{\pi'(k)}), P'(\theta_{\pi'(k)}))\right)_{k=1}^{N}.$$

*A mechanism $(Q, P) \in \mathcal{M}$ is **leximin optimal** if it is leximin preferred to any mechanism $(Q', P') \in \mathcal{M}$.*

In general decision problems, leximin optimality does not imply perfect (and, hence, proper) robustness; leximin optimal strategies need not be a best-response to *any* full support LPS.

Example 4. Consider the decision problem in Figure 2; again, the decision-maker chooses a row and each column corresponds to a state. U is associated

[17] I thank Andy Mackenzie for introducing me to the leximin criterion and for conjecturing that Corollary 2 holds.

	L	C	R
U	6	3	3
M	3	3	6
D	4	3	4

Figure 2: *Leximin optimality versus proper robustness.*

with the sorted payoff vector $(3,3,6)$; M is associated with the sorted payoff vector $(3,3,6)$; and D is associated with the sorted payoff vector $(3,4,4)$. Hence, D is leximin preferred to U and M. On the other hand, D is neither perfectly nor properly robust. To see why, fix any full support LPS, μ, of length K and let $\kappa := \min\{k \in \{1,\dots,K\} : \mu_k(L) > 0 \quad \text{or} \quad \mu_k(R) > 0\}$. Then, for any strictly positive integer $k < \kappa$, $\mu_k(C) = 1$. So, for all strictly positive integers $k < \kappa$,

$$\mathbb{E}_{\mu_k}[f(U,\tilde{\theta})] = \mathbb{E}_{\mu_k}[f(M,\tilde{\theta})] = \mathbb{E}_{\mu_k}[f(D,\tilde{\theta})] = 3,$$

where $f : \{U,M,D\} \times \{L,C,R\} \to \mathbb{R}$ is the decision-maker's payoff function. Moreover,

$$\mathbb{E}_{\mu_\kappa}[f(D,\tilde{\theta})] < \max\{\mathbb{E}_{\mu_\kappa}[f(U,\tilde{\theta})], \mathbb{E}_{\mu_\kappa}[f(M,\tilde{\theta})]\}.$$

Hence, D is *not* lexicographically preferred to either U or M. So, D is not a best-response to any full-support LPS, i.e., it is neither perfectly nor properly robust.

Nevertheless, the proof of Theorem 3 can be modified to show that the unique leximin optimal mechanism is also the efficient and revenue maximizing mechanism.

Corollary 2 (Leximin foundations)
A mechanism $(Q,P) \in \mathcal{M}$ is leximin optimal if and only if it is efficient and revenue maximizing.

Proof. Let $\mathcal{M}' \supseteq \mathcal{M}$ denote the set of pairs (Q,P), where $Q : \Theta \to \Delta(\mathcal{Q})$ and $P : \Theta \to \mathbb{R}_+$, that are individually rational and satisfy the downward adjacent incentive compatibility constraints: for all $i = 2, \ldots, N$,

$$Eu(Q(\theta_i), \theta_i) - P(\theta_i) \geq Eu(Q(\theta_{i-1}), \theta_i) - P(\theta_{i-1}).$$

The third paragraph of the proof of Theorem 3 establishes that, for any $(Q, P) \in \mathcal{M}'$ for which $Ev(Q(\cdot), P(\cdot))$ is *not* non-decreasing, there exists a $(Q', P') \in \mathcal{M}'$ for which $Ev(Q'(\cdot), P'(\cdot))$ is non-decreasing that is strictly leximin preferred to (Q, P). The first paragraph of the proof of Theorem 3 establishes that the efficient and revenue maximizing mechanism $(Q, P) \in \mathcal{M}$ is strictly leximin preferred to any $(Q', P') \in \mathcal{M}' \backslash \{(Q, P)\}$ for which $Ev(Q'(\cdot), P'(\cdot))$ is non-decreasing. Hence, the efficient and revenue maximizing mechanism is strictly leximin preferred to all other mechanisms in $\mathcal{M}'$. Because the efficient and revenue maximizing mechanism belongs to $\mathcal{M}$ and $\mathcal{M}' \supseteq \mathcal{M}$, it is therefore the unique leximin optimal mechanism. $\square$

An interesting feature of Example 4 is that leximin optimality, perfect robustness, and proper robustness coincide if the choice space is extended to allow for lotteries; the lottery $\frac{1}{2} \circ U + \frac{1}{2} \circ M$ is the unique leximin optimal, perfectly robust, and properly robust strategy. So, Corollary 2 is perhaps unsurprising given that randomization has no "hedging" advantage in the screening environment considered (see footnote 9).[18] Nevertheless, it provides further justification for the use of the efficient and revenue maximizing mechanism. The Blume, Brandenburger, and Dekel (1991a) framework is used in this paper in order to illuminate the connection between the seller's lexicographic considerations, the literature on equilibrium refinements, and Bayesian optimality. Moreover, the criterion produces an arguably more cognitively straightforward foundation for the efficient and revenue maximizing mechanism; it is the unique lexicographically optimal mechanism under the LPS $\mu = (\delta_1, \ldots, \delta_N)$.

[18]Restricting attention to pure strategies is often of interest in robust mechanism design problems even if hedging has value in a simultaneous-move game with Nature; any hedging benefit of randomization is eliminated under beliefs that Nature moves after the decision-maker's strategy is realized (see, e.g., Ke and Zhang (2020) for a discussion of the issues concerning beliefs about the timing of resolution of uncertainty). Identifying general conditions for finite and infinite-dimensional decision problems under which proper robustness and leximin optimality coincide is thus an interesting avenue for future research.

5 Discussion

This paper proposed a refinement of the maxmin criterion — proper robustness — and characterized its predictions in a canonical screening environment. It was shown that a mechanism is properly robust if and only if it is efficient and revenue maximizing. While the screening model studied in this paper is important in its own right, the optimality criterion developed can be used to study other problems of economic interest, such as multi-dimensional screening and mechanism design with multiple agents. The analysis of these important extensions is left for future research.

References

Akerlof, G.A. "The Market for "Lemons": Quality Uncertainty and the Market Mechanism." *The Quarterly Journal of Economics*, Vol. 84 (1970), pp. 488–500.

Anderson, E.J. and Nash, P. *Linear programming in infinite-dimensional spaces: theory and applications*. John Wiley & Sons, 1987.

Arrow, K.J. "Alternative approaches to the theory of choice in risk-taking situations." *Econometrica*, (1951), pp. 404–437.

Bergemann, D. and Schlag, K. "Robust monopoly pricing." *Journal of Economic Theory*, Vol. 146 (2011), pp. 2527–2543.

Bergemann, D. and Schlag, K.H. "Pricing without priors." *Journal of the European Economic Association*, Vol. 6 (2008), pp. 560–569.

Blume, L., Brandenburger, A., and Dekel, E. "Lexicographic probabilities and choice under uncertainty." *Econometrica*, (1991a), pp. 61–79.

Blume, L., Brandenburger, A., and Dekel, E. "Lexicographic probabilities and equilibrium refinements." *Econometrica*, (1991b), pp. 81–98.

Börgers, T. "(No) Foundations of dominant-strategy mechanisms: a comment on Chung and Ely (2007)." *Review of Economic Design*, Vol. 21 (2017), pp. 73–82.

Carrasco, V., Farinha Luz, V., Kos, N., Messner, M., Monteiro, P., and Moreira, H. "Optimal selling mechanisms under moment conditions." *Journal of Economic Theory*, Vol. 177 (2018), pp. 245–279.

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

Carroll, G. "Robustness in mechanism design and contracting." *Annual Review of Economics*, Vol. 11 (2019), pp. 139–166.

Che, E. "Robustly Optimal Auction Design under Mean Constraints." In "Proceedings of the 23rd ACM Conference on Economics and Computation," EC '22, pp. 153–181. New York, NY, USA: Association for Computing Machinery, 2022.

Che, Y.K. and Zhong, W. "Robustly Optimal Mechanisms for Selling Multiple Goods." *Review of Economic Studies (forthcoming)*, (2024).

Chernoff, H. "Rational selection of decision functions." *Econometrica*, (1954), pp. 422–443.

Chung, K.S. and Ely, J.C. "Foundations of dominant-strategy mechanisms." *Review of Economic Studies*, Vol. 74 (2007), pp. 447–476.

Dresher, M. *Games of strategy: Theory and Applications*. Prentice-Hall, 1961.

Dworczak, P. and Pavan, A. "Preparing for the worst but hoping for the best: Robust (bayesian) persuasion." *Econometrica*, Vol. 90 (2022), pp. 2017–2051.

Kambhampati, A. "Randomization is optimal in the robust principal-agent problem." *Journal of Economic Theory*, Vol. 207 (2023), p. 105585.

Kambhampati, A., Toikka, J., and Vohra, R. "Randomization and the Robustness of Linear Contracts." In "North American Winter Meeting Econometric Society," 2024.

Ke, S. and Zhang, Q. "Randomization and ambiguity aversion." *Econometrica*, Vol. 88 (2020), pp. 1159–1195.

Luce, R.D. and Raiffa, H. *Games and decisions: Introduction and critical survey.* John Wiley and Sons, Inc., 1957.

Mackenzie, A. "Subjective Lexicographic Expected Utility." *Working paper, Rutgers University*, (2024).

Madarász, K. and Prat, A. "Sellers with misspecified models." *The Review of*

Economic Studies, Vol. 84 (2017), pp. 790–815.

Mailath, G.J., Samuelson, L., and Swinkels, J.M. "How proper is sequential equilibrium?" *Games and Economic Behavior*, Vol. 18 (1997), pp. 193–218.

Maskin, E. and Riley, J. "Monopoly with incomplete information." *The RAND Journal of Economics*, Vol. 15 (1984), pp. 171–196.

Milgrom, P. and Shannon, C. "Monotone comparative statics." *Econometrica*, (1994), pp. 157–180.

Mussa, M. and Rosen, S. "Monopoly and product quality." *Journal of Economic Theory*, Vol. 18 (1978), pp. 301–317.

Myerson, R.B. "Refinements of the Nash equilibrium concept." *International Journal of Game Theory*, Vol. 7 (1978), pp. 73–80.

Myerson, R.B. "Optimal auction design." *Mathematics of Operations Research*, Vol. 6 (1981), pp. 58–73.

Myerson, R.B. *Game theory.* Harvard university press, 1991.

Rawls, J. *A Theory of Justice: Original Edition.* Harvard University Press, 1971.

Rochet, J.C. "A necessary and sufficient condition for rationalizability in a quasi-linear context." *Journal of Mathematical Economics*, Vol. 16 (1987), pp. 191–200.

Rothschild, M. and Stiglitz, J. "Equilibrium in Competitive Insurance Markets: An Essay on the Economics of Imperfect Information." *The Quarterly Journal of Economics*, Vol. 90 (1976), pp. 629–649.

Savage, L.J. "The theory of statistical decision." *Journal of the American Statistical association*, Vol. 46 (1951), pp. 55–67.

Savage, L.J. *The foundations of statistics.* Courier Corporation, 1972.

Selten, R. "Reexamination of the perfectness concept for equilibrium points in extensive games." *International Journal of Game Theory*, Vol. 4 (1975), pp. 25–55.

Sen, A. *Collective Choice and Social Welfare.* San Francisco: Holden Day, 1970.

Simon, L.K. and Stinchcombe, M.B. "Equilibrium refinement for infinite normal-form games." *Econometrica*, (1995), pp. 1421–1443.

Spence, M. "Job Market Signaling." *The Quarterly Journal of Economics*, Vol. 87 (1973), pp. 355–374.

Strausz, R. "Deterministic versus stochastic mechanisms in principal–agent models." *Journal of Economic Theory*, Vol. 128 (2006), pp. 306–314.

Van Damme, E. *Refinements of the Nash equilibrium concept.* Springer-Verlag, 1983.

Vohra, R.V. *Mechanism design: a linear programming approach*, Vol. 47. Cambridge University Press, 2011.

A Weak dominance and optimality

It is shown here that, for any admissible mechanism $(Q, P) \in \mathcal{M}$, there exists a full support distribution $\rho \in \Delta(\Theta)$ under which (Q, P) is a best-response in $\mathcal{M}$. Observe that the following linear program has value strictly less than zero if and only if there exists a full support distribution against which $(Q, P) \in \mathcal{M}$ is a best-response:

$$\begin{aligned}
&\min_{\delta \in \mathbb{R}, \rho \in \mathbb{R}^N} \delta \\
&\text{subject to} \\
&-\sum_{\theta \in \Theta} \rho(\theta) \geq -1 \\
&\rho(\theta) + \delta \geq 0 \quad \text{for all } \theta \in \Theta \\
&\sum_{\theta \in \Theta} \rho(\theta) \left(Ev(Q(\theta), P(\theta)) - Ev(Q'(\theta), P'(\theta)) \right) \geq 0 \quad \text{for all } (Q', P') \in \mathcal{M}.
\end{aligned} \tag{5}$$

To see why, suppose that the program has value strictly less than zero. Then, there exists a feasible (δ, ρ) with $\rho(\theta) \geq -\delta > 0$ for all $\theta \in \Theta$ and $\sum_{\theta \in \Theta} \rho(\theta) \leq 1$ such that

$$\sum_{\theta \in \Theta} \rho(\theta) \left(Ev(Q(\theta), P(\theta)) - Ev(Q'(\theta), P'(\theta)) \right) \geq 0 \quad \text{for all } (Q', P') \in \mathcal{M}.$$

If $\sum_{\theta\in\Theta}\rho(\theta)=1$, then $\rho\in\Delta(\Theta)$ is a full support distribution against which (Q,P) is a best-response. If not, then (Q,P) is a best-response against the full support distribution $\rho'\in\Delta(\Theta)$ defined by

$$\rho'(\theta)=\frac{\rho(\theta)}{\sum_{\hat{\theta}\in\Theta}\rho(\hat{\theta})}.$$

For the other direction, suppose that $\rho\in\Delta(\Theta)$ is a full support distribution against which (Q,P) is a best-response and let $-\delta=\min_{\theta}\rho(\theta)$. Then, (δ,ρ) is feasible and $-\delta=\rho(\theta')$ for some $\theta'\in\Theta$. Because $\rho(\theta')>0$ by hypothesis, it has thus been shown that there is a feasible solution with $\delta<0$. Hence, the value of (5) must be strictly less than zero.

Let $\mathcal{B}_+$ denote the space of bounded, regular, and positive Borel measures on $\mathcal{M}$. Then, the dual of (5) can be formulated as

$$\begin{aligned}
&\max_{\epsilon\in\mathbb{R}_+,\eta\in\mathbb{R}_+^N,\sigma\in\mathcal{B}_+} -\epsilon\\
&\text{subject to}\\
&\sum_{\theta\in\Theta}\eta(\theta)=1\\
&\eta(\theta)-\epsilon+\int_{\mathcal{M}}\left(Ev(Q(\theta),P(\theta))-Ev(Q'(\theta),P'(\theta))\right)d\sigma(Q',P')=0 \quad \text{for all } \theta\in\Theta.
\end{aligned} \tag{6}$$

Note also that (5) is the dual of (6). It is now shown that (6) has value greater than or equal to zero if and only if there exists a $\sigma\in\Delta(\mathcal{M})$ that weakly dominates $(Q,P)\in\mathcal{M}$. Suppose that the dual program has value weakly greater than zero. Then, there exists a feasible (ϵ,η,σ) with $\epsilon\leq 0$ (in fact, feasibility implies $\epsilon=0$). Hence, for all $\theta\in\Theta$,

$$\eta(\theta)+\int_{\mathcal{M}}\left(Ev(Q(\theta),P(\theta))-Ev(Q'(\theta),P'(\theta))\right)d\sigma(Q',P')\leq\epsilon\leq 0,$$

or

$$\eta(\theta)+\int_{\mathcal{M}}Ev(Q(\theta),P(\theta))d\sigma(Q',P')\leq\int_{\mathcal{M}}Ev(Q'(\theta),P'(\theta))d\sigma(Q',P').$$

Because $\eta(\theta) \geq 0$ for $\theta \in \Theta$ and $\sum_{\theta \in \Theta} \eta(\theta) = 1$ (so that $\eta(\theta') > 0$ for some $\theta' \in \Theta$), it follows that

$$\int_{\mathcal{M}} Ev(Q(\theta), P(\theta)) d\sigma(Q', P') \leq \int_{\mathcal{M}} Ev(Q'(\theta), P'(\theta)) d\sigma(Q', P')$$

for all $\theta \in \Theta$ and

$$\int_{\mathcal{M}} Ev(Q(\theta'), P(\theta')) d\sigma(Q', P') < \int_{\mathcal{M}} Ev(Q'(\theta'), P'(\theta')) d\sigma(Q', P') \tag{7}$$

for some $\theta' \in \Theta$. Thus, if $\sigma \in \mathcal{B}_+$ is a probability measure on $\mathcal{M}$, i.e., $\sigma \in \Delta(\mathcal{M})$ and

$$Ev(Q(\theta), P(\theta)) = \int_{\mathcal{M}} Ev(Q(\theta'), P(\theta')) d\sigma(Q', P'),$$

then σ dominates (Q, P). If $\sigma \in \mathcal{B}_+$ is not a probability measure, then the re-scaled measure $\sigma' \in \Delta(\mathcal{M})$ defined by[19]

$$\sigma'(E) = \frac{\sigma(E)}{\sigma(\mathcal{M})} \quad \text{for each Borel set } E$$

dominates (Q, P). For the other direction, suppose there exists $\sigma \in \Delta(\mathcal{M}) \subset \mathcal{B}_+$ that weakly dominates (Q, P). Set $\epsilon = 0$ and define $\eta(\theta)$ for each $\theta \in \Theta$ using the second constraint:

$$\eta(\theta) = -\int_{\mathcal{M}} \left(Ev(Q(\theta), P(\theta)) - Ev(Q'(\theta), P'(\theta)) \right) d\sigma(Q', P').$$

Because σ weakly dominates (Q, P), we have $\eta(\theta) \geq 0$ for all $\theta \in \Theta$ and $\eta(\theta') > 0$ for some $\theta' \in \Theta$. If $\sum_\theta \eta(\theta) = 1$, then (ϵ, η, σ) is feasible, implying that the value of the dual is at least zero. Otherwise, define $\eta' \in \mathbb{R}^N_+$ and $\sigma' \in \mathcal{B}_+$ by

$$\eta'(\theta) = \frac{\eta(\theta)}{\sum_{\hat{\theta} \in \Theta} \eta(\hat{\theta})} \text{ for each } \theta \in \Theta$$

[19] Notice that the denominator is strictly positive by (7).

and

$$\sigma'(E) = \frac{\sigma(E)}{\sum_{\hat{\theta} \in \Theta} \eta(\hat{\theta})} \quad \text{for each Borel set } E.$$

Then $(\epsilon, \eta', \sigma')$ is feasible, again implying that the value of the dual is at least zero.

It is now shown that the value of (6) equals the value of (5), i.e., strong duality holds for the pair of linear programs. Note that all constraints in (6) can be satisfied: set $\epsilon = \frac{1}{N}$, $\eta(\theta) = \frac{1}{N}$ for all $\theta \in \Theta$, and σ to be the Dirac measure on (Q, P). Moreover, the value of (6) is finite: $\epsilon \geq 0$ by feasibility. If $\mathcal{M}$ were finite, the equivalence of the values of (5) and (6) would then be immediately satisfied by strong duality for finite linear programs. Strong duality need not hold, in general, for linear programs with an infinite number of constraints (see Section 4.2.1 of Anderson and Nash (1987) for an example). Notice, however, that (6) is of the form (EP) in Section 3.3 of Anderson and Nash (1987) when the max operator is replaced with the min operator in the objective function and the minus sign is removed. Hence, by Theorem 3.10 of Anderson and Nash (1987), it is sufficient to observe that the following set is closed to establish the equivalence of the values of (6) and (5):

$$D = \{x(\epsilon, \eta, \sigma) \in \mathbb{R}^{N+2} : \epsilon \in \mathbb{R}_+, \eta \in \mathbb{R}^N_+, \sigma \in \mathcal{B}_+\},$$

where

$$\begin{aligned}
x_1(\epsilon, \eta, \sigma) &= \sum_{\theta \in \Theta} \eta(\theta) \\
x_2(\epsilon, \eta, \sigma) &= \eta(\theta_1) - \epsilon + \int_{\mathcal{M}} (Ev(Q(\theta_1), P(\theta_1)) - Ev(Q'(\theta_1), P'(\theta_1)))\, d\sigma(Q', P') \\
&\vdots \\
x_{N+1}(\epsilon, \eta, \sigma) &= \eta(\theta_N) - \epsilon + \int_{\mathcal{M}} (Ev(Q(\theta_N), P(\theta_N)) - Ev(Q'(\theta_N), P'(\theta_N)))\, d\sigma(Q', P') \\
x_{N+2}(\epsilon, \eta, \sigma) &= \epsilon.
\end{aligned}$$

Closedness follows from continuity of $x(\epsilon, \eta, \sigma)$ in $(\epsilon, \eta, \sigma) \in \mathbb{R}_+ \times \mathbb{R}^N_+ \times \mathcal{B}_+$ when

$\mathbb{R}_+ \times \mathbb{R}_+^N \times \mathcal{B}_+$ is equipped with the product topology inducing the Euclidean topology on $\mathbb{R}_+$ and $\mathbb{R}_+^N$ and the weak topology on $\mathcal{B}_+$.

To conclude the proof, notice that admissibility ensures that (6) has value strictly smaller than zero. So, strong duality implies that (5) has strictly negative value. Hence, for any admissible mechanism $(Q, P) \in \mathcal{M}$, there exists a full support distribution $\rho \in \Delta(\Theta)$ under which (Q, P) is a best-response in $\mathcal{M}$.