Robust Robustness*

Ian Ball Deniz Kattwinkel

29 August 2024

Abstract

The maxmin approach to distributional robustness evaluates each mechanism according to its payoff guarantee over all priors in an ambiguity set. We propose a refinement: the guarantee must be approximately satisfied at priors near the ambiguity set (in the weak topology). We call such a guarantee *robust*. The payoff guarantees from some maxmin-optimal mechanisms in the literature are not robust. We show, however, that over certain standard ambiguity sets (such as continuous moment sets), every mechanism's payoff guarantee is robust. We give a behavioral characterization of our refined robustness notion by imposing a new continuity axiom on maxmin preferences.

Keywords: robust mechanism design; maxmin expected utility; distributional robustness

*We thank audiences at SAET, the Penn Mini-Conference on Economic Theory, and the North American Summer Meeting of the Econometric Society for useful feedback. For helpful discussions, we thank Ben Brooks, Eddie Dekel, Jeff Ely, Piotr Dworzack, Drew Fudenberg, Yingni Guo, Tibor Heumann, Alex Jakobsen, Philippe Jehiel,Peter Klibanoff, Kristóf Madarász, George Mailath, Suraj Malladi, Stephen Morris, Pietro Ortoleva, Alessandro Pavan, Harry Pei, Jacopo Perego, Doron Ravid, Phil Reny, Larry Samuelson, Marciano Siniscalchi, Rani Spiegler, Juuso Toikka, and Juuso Välimäki. This paper was initially written while Ian Ball was visiting Northwestern University.

1 Introduction

The standard Bayesian approach to mechanism design assumes that the designer has a prior over the relevant set of states. In practice, the designer may not have enough information to formulate an exact prior. This raises the concern that a Bayesian-optimal mechanism may perform poorly under a slightly different state distribution. To address this concern, the maxmin approach models the designer's uncertainty as a set of priors called the *ambiguity set*. In the maxmin model, the designer evaluates each mechanism according to its *payoff guarantee*, i.e., its worst-case expected payoff over the ambiguity set.

In this paper, we propose a refinement of the maxmin approach. The ambiguity set, like the prior in the Bayesian model, is an exogenous input to the design problem that is subject to error. Our refinement demands that a mechanism's payoff guarantee is itself *robust* in the following sense: the expected payoff from a mechanism does not drop far below its guarantee at priors just outside the ambiguity set. The validity of a mechanism's payoff guarantee rests on the state distribution being inside the ambiguity set. Under our refinement, this guarantee extends continuously to nearby priors.

To illustrate how our refined notion of robustness can be violated, consider the standard monopoly pricing problem. The designer (seller) has a single good, and she is uncertain of the buyer's valuation. Suppose that the seller does not have enough information to formulate an exact prior over the buyer's valuation. She knows only that the median valuation is λ, where $\lambda > 0$. She evaluates each implementable social choice function according to its revenue guarantee over all valuation distributions with median λ. It can be verified that the best possible revenue guarantee is $\lambda/2$. This guarantee is uniquely achieved by posting a price of λ. But the guarantee from this posted price is not robust: For any positive ε there exists a valuation distribution with median $\lambda-\varepsilon$ under which this posted price yields revenue 0 (because the good is never purchased).[1]

We operationalize and axiomatize our refined notion of robustness. To show

[1]For example, consider the distribution that puts all mass on the point $\lambda - \varepsilon$.

that our refinement has bite, we give examples from the literature of proposed maxmin-optimal mechanisms whose payoff guarantees are not robust. On the other hand, we show that if the designer's ambiguity set comes from certain widely used classes, then every mechanism's payoff guarantee is necessarily robust. Finally, we give a behavioral characterization of robustness in terms of a new continuity property of the maxmin preference relation.

Formally, we consider an Anscombe–Aumann setting, enriched with a Polish topology on the state space. The state represents any aspects of the environment that are unknown to the designer, such as agents' preferences or technology. The topology on the state space reflects which states the designer finds difficult to distinguish. Therefore, the associated weak topology on the space of state distributions captures which perturbations of a state distribution are difficult for the designer to rule out. For example, in the monopoly pricing problem, the designer's partial information may not allow her to confidently distinguish between the valuations θ and $\theta + \varepsilon$, for ε sufficiently small. If the designer considers the distribution that puts probability 1/2 each on valuations 0 and θ, then it is difficult to rule out the distribution that puts probability 1/2 each on valuations 0 and $\theta + \varepsilon$.

The designer's uncertainty about the state is represented by a set of priors called the ambiguity set. The designer has a state-dependent utility function over decisions. A social choice function is an Anscombe–Aumann act, i.e., a map from states to decision lotteries. The designer evaluates each social choice function according to its payoff guarantee over all priors in the ambiguity set. The payoff guarantee from a social choice function depends only on the induced *value function*, which specifies the designer's utility in each state. Therefore, the designer's problem can be reduced to directly choosing a value function from a feasible set. For example, in a standard adverse selection problem, this feasible set contains every value function that is induced by some incentive-compatible social choice function.

The payoff guarantee from a value function v over an ambiguity set Π is *robust* if the expected payoff from v approximately satisfies the guarantee at priors sufficiently close to Π in the weak topology. An ambiguity set Π is *globally robust* if for every bounded value function v, the payoff guarantee from v over Π is robust. If the designer uses a globally robust ambiguity set, then she is assured that whichever decision environment she faces and whichever social choice function she implements, the associated payoff guarantee will be robust.

Theorem 1 shows that the following widely used ambiguity sets are globally robust: continuous moment sets (which restrict the expectation of continuous functions of the state) and balls defined with respect to the Wasserstein or Prokhorov metrics. We show that any ambiguity set taking one of these forms has the following richness property. Any prior close to the ambiguity set (in the weak topology) can be modified with small probability to obtain a prior inside the ambiguity set. Such a modification has a small effect on the expectation of any bounded value function.

Conversely, the following commonly used ambiguity sets do not have this richness property: relative entropy and total variation balls, singletons, and sets defined by restrictions on a distribution's support, quantiles, or marginals. These restrictions do not recognize the topology on the state space. Theorem 2 shows that these ambiguity sets are not globally robust.

We next characterize the behavioral content of our refinement by giving an axiomatization within the framework of decision-making under uncertainty. For the axiomatization, we explicitly model the designer's primitive preferences over Anscombe–Aumann acts (rather than her induced preferences over value functions).[2] The designer's utility function and ambiguity set together induce a maxmin preference relation over acts. Theorem 3 shows that the payoff guarantee from an act over a closed ambiguity set Π is robust if and only if the associated preference relation satisfies an upper semicontinuity property at that act. This characterization relates our notion of robustness, which concerns nearby priors in a particular utility representation, with a continuity axiom, which concerns preferences over nearby acts.

Theorem 4 axiomatizes the global robustness of an ambiguity set. We in-

[2]In this part, we use the decision theory terminology of "acts" rather than the mechanism design terminology of "social choice functions."

troduce continuity axioms with respect to a new mode of convergence for acts. This mode of convergence is inspired by Γ-convergence of real-valued functions, which is the standard notion of convergence in the analysis of minimization problems. Assuming state-independent utility, we show that an ambiguity set is globally robust and tight (a topological property) if and only if the associated preference relation satisfies Γ-continuity and tightness axioms.

The rest of the paper is organized as follows. Section 2 introduces the setting and defines our refined notion of robustness. Section 3 gives examples from the literature of maxmin-optimal mechanisms whose payoff guarantees are not robust. Section 4 classifies which commonly used ambiguity sets are globally robust and which are not. Section 5 provides axiomatizations of robustness. Section 6 discusses related literature. Section 7 is the conclusion. Proofs omitted from the main text are in Appendix A. Additional results and proofs are in Appendix B.

2 Model

We introduce the setting and then we present our refined notion of robustness.

2.1 Maxmin mechanism design setting

Consider a designer in the following environment. There is a state space Θ, which is a Polish topological space endowed with its Borel σ-algebra $\mathcal{B}(\Theta)$.[3] The state represents any aspects of the environment that are unknown to the designer, such as agents' preferences or technology. There is a decision space X, which is endowed with a σ-algebra. Denote by $\Delta(\Theta)$ (respectively, $\Delta(X)$) the space of probability measures on Θ (respectively, X). A social choice function is a measurable function $f\colon \Theta \to \Delta(X)$. The designer has a bounded, measurable utility function $u\colon X \times \Theta \to \mathbf{R}$, which we extend linearly to $\Delta(X) \times \Theta$. The designer's state-dependent utility function can

[3] That is, Θ is homeomorphic to a complete metric space that has a countable dense subset.

capture many different objectives, even regret-minimization. For example, in an auction setting, u can equal negative regret, i.e., the maximal realized valuation minus revenue.

What distinguishes our setting from the classical Anscombe–Aumann framework is the topology on the state space.[4] This topology reflects which states the designer finds difficult to distinguish. The topology will be important for our refined notion of robustness below.

The designer evaluates each social choice function f according to the objective

$$\inf_{\pi \in \Pi} \int_{\Theta} u(f(\theta), \theta) \, \mathrm{d}\pi(\theta), \tag{1}$$

where Π is a nonempty subset of $\Delta(\Theta)$ called the *ambiguity set*. We take an infimum rather than a minimum because we have not made assumptions on Π to guarantee the existence of a minimizer.[5] In a maxmin design problem, the designer maximizes the objective in (1) over a feasible set $\mathcal{F}$ of social choice functions. For example, in a standard adverse selection problem, the state θ is the type profile of the agents and the set $\mathcal{F}$ contains all social choice functions satisfying incentive compatibility and participation constraints. Our abstract formulation takes $\mathcal{F}$ as a primitive; we do not explicitly model any agents other than the designer. In summary, a maxmin design problem is represented by a tuple $(\Theta, X, u, \Pi, \mathcal{F})$.

The designer's objective in (1) depends on the social choice function f only through the induced *value function* v_f, defined by $v_f(\theta) = u(f(\theta), \theta)$ for each θ in Θ. The function v_f specifies the designer's utility in each state. Therefore, the designer's problem can be reduced to directly choosing among induced value functions.[6]

[4]Here we adopt the terminology of mechanism design. In the language of decision theory, the designer is the decision-maker; decisions are consequences; and social choice functions are Anscombe–Aumann acts.

[5]To be sure, if u is state-independent, then this objective also has a Gilboa and Schmeidler (1989) representation as a minimum over a set of finitely additive probability measures. Let Π' denote the closed convex hull of Π in the space of finitely additive probability measures, endowed with the topology of setwise convergence. For any simple social choice function f, the infimum over Π' is achieved and the minimum value agrees with the infimum in (1).

[6]For the axiomatizations in Section 5, we consider the designer's primitive preferences

Formally, we reduce a maxmin design problem to a triple $(\Theta, \Pi, \mathcal{V})$, where $\mathcal{V}$ is a nonempty subset of $B(\Theta)$, the space of bounded, measurable real-valued functions on Θ. Given a value function v in $B(\Theta)$ and a prior π in $\Delta(\Theta)$, let $\langle v, \pi \rangle$ denote the integral of v with respect to π. The designer therefore maximizes over $\mathcal{V}$ the objective

$$W_\Pi(v) = \inf_{\pi \in \Pi} \langle v, \pi \rangle. \tag{2}$$

We call $W_\Pi(v)$ the designer's *payoff guarantee* from v over Π. This guarantee is the worst-case expected payoff from value function v over all priors in the ambiguity set Π. The solution set of the maxmin design problem is $\operatorname{argmax}_{v \in \mathcal{V}} W_\Pi(v)$. A social choice function (or mechanism) that induces a value function in this solution set is *maxmin optimal* with respect to the ambiguity set Π.

2.2 Robustness

Now we state our refined notion of robustness. In the space $\Delta(\Theta)$, a sequence (π_n) *weakly converges* to π if $\langle h, \pi_n \rangle \to \langle h, \pi \rangle$ for each bounded, continuous function $h\colon \Theta \to \mathbf{R}$. Crucially, weak convergence in $\Delta(\Theta)$ reflects the topology on the state space Θ. For example, a sequence (δ_{θ_n}) of unit masses weakly converges to the unit mass δ_θ if and only if the sequence (θ_n) converges to θ in the space Θ.[7] Unless otherwise indicated, the topology on $\Delta(\Theta)$ is assumed to be the topology of weak convergence. In $\Delta(\Theta)$, convergence refers to weak convergence.

Definition 1 (Robustness)**.** Let Π be a nonempty subset of $\Delta(\Theta)$.

1. Given v in $B(\Theta)$, the payoff guarantee from v over Π is *robust* if for every sequence (π_n) in $\Delta(\Theta)$ that converges to a prior in the closure of Π,

$$\liminf_n \langle v, \pi_n \rangle \geq W_\Pi(v). \tag{3}$$

over Anscombe–Aumann acts rather than her induced preferences over value functions.

[7] Given θ in Θ, the unit mass δ_θ in $\Delta(\Theta)$ is defined by $\delta_\theta(A) = 1$ if θ is in A and $\delta_\theta(A) = 0$ otherwise.

2. The set Π is *globally robust* if for each v in $B(\Theta)$, the payoff guarantee from v over Π is robust.

In words, the payoff guarantee from a value function over an ambiguity set is robust if the expected payoff from the value function approximately satisfies the guarantee at priors near the ambiguity set (in the weak topology). The inequality in (3) can be violated only if infinitely many of the priors in the sequence (π_n) are outside Π. In particular, if $\Pi = \Delta(\Theta)$, then there are no priors outside Π, so the ambiguity set $\Delta(\Theta)$ is globally robust.

The weak topology on $\Delta(\Theta)$ determines which priors are "near" the ambiguity set. For example, the discrete probability measures $\pi = p_1\delta_{\theta_1} + \cdots + p_n\delta_{\theta_n}$ and $\pi' = p'_1\delta_{\theta'_1} + \cdots + p'_n\delta_{\theta'_n}$ are close in the weak topology if for each i, the difference $|p_i - p'_i|$ is small, and the states θ_i and θ'_i are close according to the topology on Θ. The topology on Θ reflects which states the designer finds difficult to distinguish. If there are certain states that the designer can confidently distinguish from all other states, this can be represented by a topology under which these states are isolated points. In the applications that we consider below, the state space Θ has a natural topology without isolated points.

Remark 1 (Robustness in discrete models)**.** Even in discrete models, our notion of robustness has bite. For example, consider a binary model in which a buyer's valuation is assumed to be either $\theta_L = 1$ or $\theta_H = 2$. Unless the designer can confidently distinguish between the valuations 2 and 2.001, say, then it is natural to view θ_L and θ_H as points in a continuous subset Θ of the real line, with the usual topology. In this case, robustness will take into consideration small perturbations of θ_L and θ_H, even if the ambiguity set contains only priors that concentrate on $\{\theta_L, \theta_H\}$.

If a designer is seeking assurance against distributional misspecification, an alternative approach is to enlarge the ambiguity set. But unless the ambiguity set is enlarged to the full space $\Delta(\Theta)$ (in which case payoff guarantees are often trivial), there will generally still be priors just outside the ambiguity set. The designer may be concerned if the expected payoff from a mechanism drops far below its payoff guarantee at those nearby priors.[8] If the payoff guarantee is robust, then the designer can be assured that the guarantee will not drop dramatically if the ambiguity set is enlarged slightly further. We will show below that our notion of robustness generally depends on the "richness" of the ambiguity set, not its size.

3 Robustness in applications

In this section, we give examples from the literature of maxmin-optimal mechanisms that perform very poorly if some prior in the ambiguity set is slightly perturbed. Motivated by these examples, we then establish preliminary results about our robustness notion.

3.1 Non-robustness of maxmin-optimal mechanisms

We consider monopoly pricing, Bayesian persuasion, and delegated project choice. In each case, there is a simple maxmin-optimal mechanism with respect to a particular ambiguity set, but this mechanism's payoff guarantee is not robust.

Monopoly pricing In the robust monopoly pricing problems studied in Bergemann and Schlag (2008) and Carroll (2017), the payoff guarantee from the maxmin-optimal mechanism sometimes fails to be robust, as we discuss in Section 6 and Section 4, respectively. Here we formalize the simple monopoly pricing problem from the introduction. The state $\theta \in \Theta = \mathbf{R}_+$ is the buyer's valuation for the good. The ambiguity set Π contains all priors with median λ.[9] The maxmin solution is the posted price $p^* = \lambda$. At this price, the good is sold if and only if the buyer's valuation is at least λ.[10] The induced value function

[8] Recall that the maxmin representation does not express different levels of confidence in different priors in the ambiguity set; see Remark 6 for a discussion of variational preferences, which can express different confidence in different priors.

[9] Formally, a prior π is in Π if and only if $\mathbf{P}_\pi(\theta \leq \lambda) \geq 1/2$ and $\mathbf{P}_\pi(\theta \geq \lambda) \geq 1/2$.

[10] We assume that ties are broken in the designer's favor, as is standard.

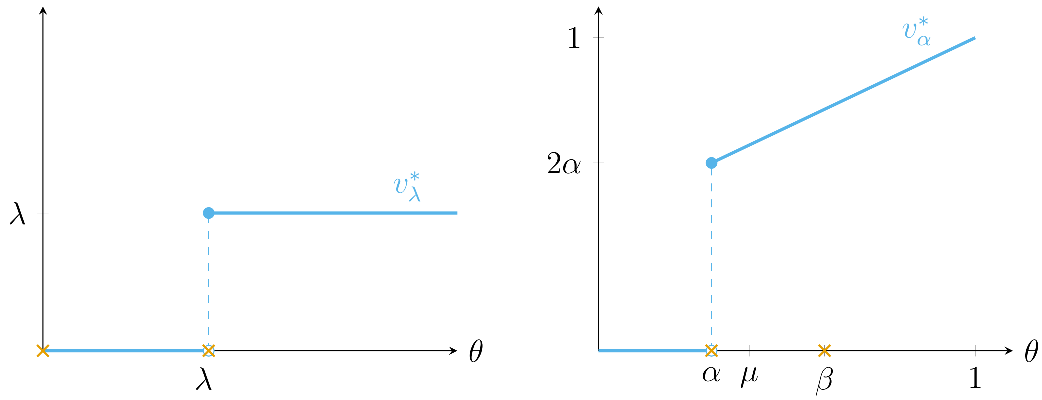

Figure 1. Non-robust payoff guarantees: posted price (left) and KG experiment (right)

v^*_λ is shown in the left panel of Figure 1. The payoff guarantee $W_\Pi(v^*_\lambda)$ equals $\lambda/2$. This worst-case payoff from v^*_λ is achieved at the prior $\pi = \delta_0/2 + \delta_\lambda/2$.[11] These two point masses are indicated on the plot. The payoff guarantee from v^*_λ is not robust. If the prior π is perturbed to $\pi_\varepsilon = \delta_0/2 + \delta_{\lambda-\varepsilon}/2$, for any $\varepsilon > 0$, then the seller's expected revenue drops to 0.

Persuasion Consider Hu and Weng's (2021) maxmin version of the persuasion problem in Kamenica and Gentzkow (2011).[12] The sender commits to a Blackwell experiment about a binary fundamental $\omega \in \Omega = \{0, 1\}$. The receiver observes the realization of the experiment and chooses a binary action $a \in \{0, 1\}$. Payoffs for the sender and receiver are given by $u_S(a, \omega) = a$ and $u_R(a, \omega) = -(a - \omega)^2$. The sender is uncertain of the receiver's belief over Ω. Thus, the state $\theta \in \Theta = [0, 1]$ is the receiver's belief, i.e., the probability assigned to $\omega = 1$. The ambiguity set Π contains all priors with fixed mean μ that are supported on the interval $[\alpha, \beta]$, where $0 < \alpha < \mu < \beta \leq 1$ and $\alpha < 1/2$. The interpretation is that the sender and receiver initially have common belief μ, but the sender is uncertain of what additional information

[11]This worst-case payoff is achieved at any prior in Π that assigns probability 1/2 to the set $[\lambda, \infty)$.

[12]We describe a special case of Hu and Weng's (2021) analysis. Kosterina (2022) studies a maxmin version of a continuous persuasion problem.

the receiver gets. The sender knows that the receiver's belief remains in the interval $[\alpha, \beta]$.

Hu and Weng (2021) show that if μ is sufficiently close to α, then a maxmin solution is the KG α-experiment, i.e., the binary experiment defined by the property that belief α is split between 0 and 1/2.[13] This experiment induces the value function v^*_α shown in the right panel of Figure 1. The probability of the "high" realization of the experiment is affine in θ. This realization induces the receiver to choose action $a = 1$ if and only if $\theta \geq \alpha$. The expected payoff from v^*_α is constant over Π, but the payoff guarantee from v^*_α over Π is not robust. Consider the prior $\pi = p\delta_\alpha + (1-p)\delta_\beta$, where $p\alpha + (1-p)\beta = \mu$. These two point masses are indicated on the plot. The prior π is in Π, but if π is perturbed to $\pi_\varepsilon = p\delta_{\alpha-\varepsilon} + (1-p)\delta_\beta$, for any ε in $(0, \alpha)$, then the sender's expected utility drops by $2\alpha p$.

Project choice Consider Guo and Shmaya's (2023) maxmin version of the delegated project choice problem in Armstrong and Vickers (2010). There are two players: a principal and an agent. The principal must select a feasible project. Each project is represented by a pair $u = (u_A, u_P) \in \mathbf{R}^2$ indicating the agent's and principal's respective payoffs. Payoffs are normalized relative to the status quo $(0, 0)$. The agent privately knows the set $\mathcal{A}$ of available projects. The principal's loss from project u in "state" $\mathcal{A}$ is her regret $\max_{u' \in \mathcal{A} \cup \{(0,0)\}} u'_P - u_P$.

The agent can propose a project u from $\mathcal{A} \cup \{(0, 0)\}$.[14] The principal commits to a mechanism $\alpha\colon \mathbf{R}^2 \to [0, 1]$. Under mechanism α, the principal adopts proposal u with probability $\alpha(u)$. With complementary probability, the principal keeps the status quo $(0, 0)$. The principal evaluates each mechanism according to its worst-case regret over all finite sets $\mathcal{A} \subset [\underline{u}_A, 1] \times [0, 1]$, where $\underline{u}_A$ is a fixed parameter in $[0, 1]$. The optimal regret guarantee is $R = (1 -$

[13] See Proposition 3 (p. 928) and Proposition 4 (p. 930). Further, this is the unique maxmin solution if $\alpha < \beta \leq 1/2$ or $\alpha < 1 - \beta < 1/2 < \beta$.

[14] It is assumed that it is infeasible for the agent to propose a project outside $\mathcal{A} \cup \{(0, 0)\}$. The single-proposal protocol is the simplest case considered in Guo and Shmaya (2023). They also solve the case of multiple proposals.

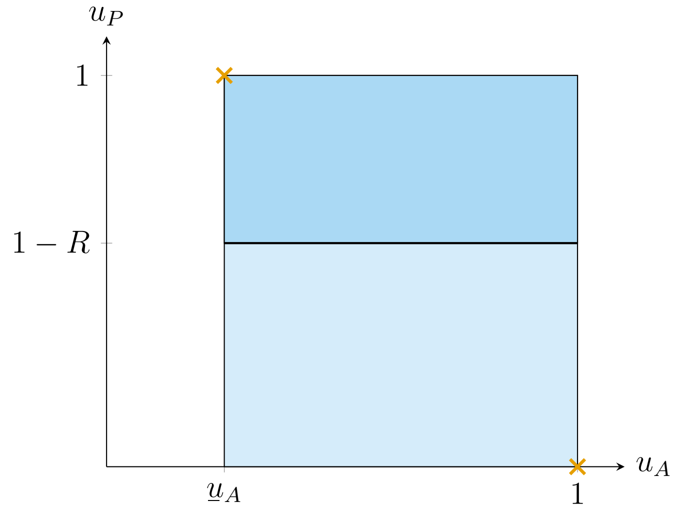

Figure 2. Worst-case prior under project choice mechanism

$\underline{u}_A)/(2-\underline{u}_A)$. This guarantee is achieved by the following mechanism, which is illustrated in Figure 2. A project u is *top-tier* if $u_P \geq 1-R$. If the agent proposes a top-tier project (in the dark shaded region), then the principal adopts the proposal with certainty. Any other proposal u (in the light shaded region) is adopted with probability $\underline{u}_A/u_A$. If the agent proposes a project that is not top-tier, then his expected utility is exactly $\underline{u}_A$. Thus, the agent finds it optimal to propose some top-tier project if any are feasible.[15]

The regret guarantee from the proposed mechanism is not robust. Consider the set $\mathcal{A} = \{(\underline{u}_A, 1), (1, 0)\}$ marked on the graph. In state $\mathcal{A}$, it is optimal for the agent to propose the top-tier project $(\underline{u}_A, 1)$, so the principal's regret is 0. If the set $\mathcal{A}$ is perturbed to $\mathcal{A}_\varepsilon = \{(\underline{u}_A - \varepsilon, 1), (1, 0)\}$, for any $\varepsilon > 0$, then the agent proposes project $(1, 0)$, and the principal's regret jumps up to 1, which is strictly worse than the regret guarantee of R.

[15]It is assumed that the principal can select the agent's best response to each mechanism. To derive this payoff guarantee, observe that if some top-tier project is feasible, then the principal's regret is at most R (from adopting a suboptimal top-tier project). If no top-tier project is feasible, then the principal's regret is at most $(1-\underline{u}_A)(1-R)$ (from adopting the status quo). By the definition of R, these bounds agree.

3.2 Disciplining ambiguity sets

In the above examples and in most of the distributionally robust mechanism design literature, maxmin-optimal mechanisms are derived for ambiguity sets in some special parametric family. Here we show that essentially any Bayesian optimal mechanism is maxmin optimal with respect to *some* non-singleton ambiguity set defined by a simple inequality. Call a prior *nondegenerate* if it is not equal to a unit mass.

Proposition 1 (Bayesian solutions are maxmin optimal)
Fix $\mathcal{V} \subset B(\Theta)$. Let v_0 be a value function in $\mathcal{V}$ and let π_0 be a nondegenerate prior in $\Delta(\Theta)$. If v_0 is in $\operatorname{argmax}_{v \in \mathcal{V}} \langle v, \pi_0 \rangle$, then v_0 is in $\operatorname{argmax}_{v \in \mathcal{V}} W_{\Pi_0}(v)$, where Π_0 is the non-singleton ambiguity set defined by

$$\Pi_0 = \{\pi \in \Delta(\Theta) : \langle v_0, \pi \rangle \geq \langle v_0, \pi_0 \rangle\} .$$

Proof. For any v in $\mathcal{V}$, we have

$$W_{\Pi_0}(v) \leq \langle v, \pi_0 \rangle \leq \langle v_0, \pi_0 \rangle = W_{\Pi_0}(v_0),$$

where the first inequality holds because π_0 is in Π_0; the second inequality holds because v_0 is in $\operatorname{argmax}_{v \in \mathcal{V}} \langle v, \pi_0 \rangle$; and the equality holds by the definition of Π_0.[16] We check that Π_0 is not a singleton. Choose θ in Θ such that $v(\theta) \geq \langle v_0, \pi_0 \rangle$. Since $\langle v_0, \cdot \rangle$ is linear, it follows that $[\pi_0, \delta_\theta] \subset \Pi_0$. The interval $[\pi_0, \delta_\theta]$ is nondegenerate because $\pi_0 \neq \delta_\theta$. □

Using the construction in Proposition 1, the ambiguity set can be tailored to essentially any desired Bayesian-optimal mechanism. Moreover, it may not be apparent that the ambiguity set has been constructed in this way. For example, in the monopoly pricing problem, the Bayesian solution under the prior $\delta_0/2 + \delta_\lambda/2$ is a posted price of λ. The associated ambiguity set Π_0 from Proposition 1 consists of all distributions with median at least λ. This statistical constraint does not appear related to a posted price. Our takeaway

[16] The same argument goes through if Π_0 is replaced with any subset Π_0' of Π_0 that contains π_0.

is that the simplicity of a derived optimal mechanism cannot itself justify the form of the ambiguity set. The motivation for the ambiguity set must be external to the model.

3.3 Continuous value functions and robustness

The examples in Section 3.1 illustrate how discrete choices by an agent can induce discontinuities in the value function, which can in turn lead to non-robustness. Here, we relate the continuity of a value function to the robustness of its payoff guarantee. For any function $v\colon \Theta \to \mathbf{R}$, let $\operatorname{lsc} v$ denote the lower semicontinuous envelope of v, i.e., the pointwise greatest lower semicontinuous function that is pointwise smaller than v. Geometrically, the epigraph of $\operatorname{lsc} v$ is the closure of the epigraph of v.

Proposition 2 (Robustness and continuity)
Let v be a value function in $B(\Theta)$.

1. *The payoff guarantee from v over Π is robust for every ambiguity set Π if and only if v is continuous.*
2. *The payoff guarantee from v over Π is robust for every closed ambiguity set Π if and only if v is lower semicontinuous.*
3. *Given a prior π_0 in $\Delta(\Theta)$, the payoff guarantee from v over the singleton $\{\pi_0\}$ is robust if and only if $\langle v, \pi_0 \rangle = \langle \operatorname{lsc} v, \pi_0 \rangle$.*

The proof of Proposition 2 uses the portmanteau theorem. If a value function v is continuous (respectively, lower semicontinuous), then by the portmanteau theorem, the map $\langle v, \cdot \rangle$ on $\Delta(\Theta)$ is continuous (respectively, lower semicontinuous). If $\langle v, \cdot \rangle$ is continuous, then perturbing any prior in the ambiguity set has a small effect on the expectation of v, so the payoff guarantee from v over any ambiguity set is robust. If $\langle v, \cdot \rangle$ is lower semicontinuous, then the expectation of v can jump down, but not up, at the limit of a sequence of priors. Therefore, the robustness inequality (3) cannot be violated by any sequence (π_n) converging to a prior in Π. This implies robustness if Π is closed.

Simple maxmin-optimal mechanisms such as posted prices can induce value functions that are discontinuous. If a value function is discontinuous, then the robustness of its payoff guarantee over an ambiguity set depends on the structure of the ambiguity set.

The last part of Proposition 2 considers subjective expected utility, which corresponds to maxmin expected utility with a singleton ambiguity set. The expected payoff from v under prior π_0 is robust if and only if π_0 puts zero probability on the set of states at which $\operatorname{lsc} v$ lies strictly below v. (If those states are slightly perturbed, the payoff from v can jump down.) We conclude that if a value function has at most countably many discontinuities, then its expected payoff under any continuous prior is robust.

4 Classification of ambiguity sets

In this section, we identify which commonly used ambiguity sets are globally robust and which are not.

4.1 Ambiguity sets that are globally robust

We formally define standard ambiguity sets that will prove to be globally robust.

Moment sets When the state is a real number, it is common to seek a payoff guarantee over all distributions with specified mean and variance; see Scarf (1958) for a classical application to an inventory problem and Azar and Micali (2012), Auster (2018), Bachrach et al. (2022) and Carrasco et al. (2018, 2019) for applications to auctions.

We define moment restrictions for priors on the arbitrary Polish space Θ. For any measurable function $g\colon \Theta \to \mathbf{R}^m$ and any subset Y of $\mathbf{R}^m$, let

$$M(g, Y) = \{\pi \in \Delta(\Theta) : \mathbf{E}_{\theta\sim\pi}[g(\theta)] \in Y\}.$$

Whenever we constrain the expectation $\mathbf{E}_{\theta\sim\pi}[g(\theta)]$, we are implicitly requiring integrability: $\mathbf{E}_{\theta\sim\pi}[|g_j(\theta)|] < \infty$ for each $j = 1, \ldots, m$. For any prior π, the expectation $\mathbf{E}_{\theta\sim\pi}[g(\theta)]$ lies in $\operatorname{conv} g(\Theta)$, the convex hull of the image of g. Thus, $M(g, Y) = M(g, Y \cap \operatorname{conv} g(\Theta))$ for any g and Y.

We are interested in restrictions on *continuous* moment functions. Even if g is continuous, a moment set $M(g, Y)$ can encode a restriction on a *discontinuous* moment function. For example, with $\Theta = \mathbf{R}$, let $g(\theta) = (\theta - \theta_0)^2$ and $Y = (-\infty, 0]$. Then $M(g, Y)$ equals $\{\delta_{\theta_0}\}$, which should not count as a continuous moment set. The problem here is that Y does not contain any points in the relative interior of $\operatorname{conv} g(\Theta)$. We impose conditions on Y and g to rule out pathological examples of this form.

A subset Y of $\operatorname{conv} g(\Theta)$ is *uniformly g-interior* if there exists $\delta > 0$ such that $Y^\delta \cap \operatorname{aff} g(\Theta) \subset \operatorname{conv} g(\Theta)$, where $Y^\delta = \{y' \in \mathbf{R}^m : \inf_{y\in Y} \|y - y'\| \leq \delta\}$. In Appendix A.3, we show that our results go through with a weaker interiority condition.[17] An ambiguity set Π is a *continuous moment set* if $\Pi = M(g, Y)$ for some dimension $m \geq 1$, some continuous function $g \colon \Theta \to \mathbf{R}^m$, and some subset Y of $\mathbf{R}^m$ such that $Y \cap \operatorname{conv} g(\Theta)$ is uniformly g-interior.

Metric balls It is natural to seek a payoff guarantee over all distributions in a ball around some reference prior; Pinar and Kizilkale (2017) take this approach in a monopoly screening problem. Given a function $D \colon \Delta(\Theta) \times \Delta(\Theta) \to [0, \infty]$, radius $r > 0$, and prior π_0, let

$$B_D(\pi_0, r) = \{\pi \in \Delta(\Theta) : D(\pi_0, \pi) \leq r\}.$$

We consider two standard metrics D on $\Delta(\Theta)$. To define these metrics, we assume that a compatible metric d on Θ has been chosen.[18]

The *Wasserstein metric* W is defined by[19]

$$W(\mu, \nu) = \inf_{\gamma} \mathbf{E}_{(\theta,\theta')\sim\gamma} [d(\theta, \theta')],$$

[17] The weaker condition is satisfied by every moment set we have come across in applications; see Appendix B.3 for some examples.

[18] That is, the metric d is complete and induces the topology on Θ.

[19] If d is not bounded, then W can take the value ∞, so technically it is not a metric.

where the infimum is over all probability measures γ in $\Delta(\Theta\times\Theta)$ with $\operatorname{marg}_1 \gamma = \mu$ and $\operatorname{marg}_2 \gamma = \nu$. Thus, $W(\mu, \nu)$ is the infimal expected moving distance when transporting mass from μ to ν.

The *Prokhorov metric* P is defined by

$$P(\mu, \nu) = \inf\{\varepsilon : \mu(A) \leq \nu(A^\varepsilon) + \varepsilon \text{ and } \nu(A) \leq \mu(A^\varepsilon) + \varepsilon,\ \forall A \in \mathcal{B}(\Theta)\},$$

where $A^\varepsilon = \{\theta' \in \Theta : \inf_{\theta\in A} d(\theta, \theta') \leq \varepsilon\}$ and $\mathcal{B}(\Theta)$ is the Borel σ-algebra on Θ. The Prokhorov metric induces the weak topology on $\Delta(\Theta)$.

Theorem 1 (Globally robust)
Continuous moment sets, Wasserstein balls, and Prokhorov balls are globally robust.

If the designer uses an ambiguity set taking one of these forms, then she is assured that whichever decision environment she faces and whichever social choice function she implements, the associated payoff guarantee will be robust.

Remark 2 (Other metrics)**.** In the proof, we show that any D-ball is robust, provided that D is a metric that induces the weak topology on $\Delta(\Theta)$ and is convex in each of its arguments (which holds if D is induced by a norm).

Remark 3 (Continuous value functions)**.** Consider a continuous value function v_0 that is Bayesian optimal with respect to a nondegenerate prior π_0. Proposition 1 constructs an ambiguity set Π_0 such that v_0 is maxmin optimal with respect to Π_0. This set Π_0 is actually the moment set $M(v_0, Y)$ with $Y = [\langle v_0, \pi_0\rangle, \infty)$. Provided that π_0 does not concentrate on $\operatorname{argmax}_{\theta\in\Theta} v_0(\theta)$, this set is a continuous moment set, and hence is robust by Theorem 1.

To prove Theorem 1, we show that any ambiguity set Π taking one of the specified forms has the following richness property. Consider a sequence (π_n) of priors outside Π that converges to a prior in the closure of Π. We show that for some sequence (ε_n) converging to 0, each prior π_n can be modified with probability ε_n to get a prior inside Π. These modifications have a vanishing effect on the expectation of any bounded, measurable value function v, so the robustness inequality (3) must hold.

We illustrate these modifications through the monopoly pricing example in Section 3.1. Here, let Π contain all priors over $\mathbf{R}_+$ with mean $\lambda/2$. Let $\pi = \delta_0/2 + \delta_\lambda/2$. Note that π is in Π. Let (θ_n) be a strictly increasing sequence that converges to λ. For each n, let $\pi_n = \delta_0/2 + \delta_{\theta_n}/2$. The sequence (π_n) converges to π, but each π_n is outside Π. For each n, the prior π_n can be modified with probability $\varepsilon_n = (1/2)(\lambda - \theta_n)/(\lambda + 1 - \theta_n)$ to form the prior $\pi'_n = \delta_0/2 + (1/2 - \varepsilon_n)\delta_{\theta_n} + \varepsilon_n \delta_{\lambda+1}$. This prior π'_n has mean $\lambda/2$ and hence is in Π. Note that $\varepsilon_n \downarrow 0$. On the other hand, let Π' contain all priors over $\mathbf{R}_+$ with *median* λ. The prior π is also in Π, but π_n cannot be modified with small probability to get a prior in Π'. If π_n is modified with probability strictly less than 1/2, then the median of the resulting distribution will still be strictly below λ.

4.2 Ambiguity sets that are not globally robust

Here, we formally define standard ambiguity sets that will prove not to be globally robust.

- A *probability set* is defined by

$$P(A, \alpha, \beta) = \{\pi \in \Delta(\Theta) : \alpha \leq \pi(A) \leq \beta\},$$

  for some measurable proper subset A of Θ and some $\alpha, \beta \in [0, 1]$ with $\alpha \leq \beta$. In particular $P(A, 1, 1)$ is the *support set* that contains all priors π with $\pi(A) = 1$. A support set is used in the project choice problem of Guo and Shmaya (2023) and in the monopoly pricing problem of Bergemann and Schlag (2008).

- A *support–moment set* is defined by

$$M(S; g, Y) = \{\pi \in \Delta(\Theta) : \pi(S) = 1 \text{ and } \mathbf{E}_{\theta \sim \pi}[g(\theta)] \in Y\},$$

  for some measurable proper subset S of Θ, some continuous function $g \colon \Theta \to \mathbf{R}^m$, and some subset Y of $\mathbf{R}^m$ that intersects the relative interior of conv $g(S)$. Support–moment ambiguity sets have been used in robust versions of the multi-good monopoly problem (Che and Zhong, 2021) and auction design (Bachrach et al., 2022).

- Given priors $\mu, \nu \in \Delta(\Theta)$, write $\mu \ll \nu$ if μ is absolutely continuous with respect to ν. The *relative entropy* (Kullback–Leibler divergence) is defined by

$$R(\mu \parallel \nu) = \begin{cases} \langle \mu, \log(\frac{\mathrm{d}\mu}{\mathrm{d}\nu}) \rangle & \text{if } \mu \ll \nu, \\ \infty & \text{otherwise,} \end{cases}$$

  where $\frac{\mathrm{d}\mu}{\mathrm{d}\nu}$ denotes the Radon–Nikodym derivative of μ with respect to ν. The *relative entropy ball* of radius $r > 0$ about the reference prior ν contains all priors μ satisfying $R(\mu \parallel \nu) \leq r$. Relative entropy balls are used in a variant of multiplier preferences called *constraint preferences* (Hansen and Sargent, 2001).

- Suppose that $\Theta = \prod_{j=1}^{k} \Theta_j$, for some Polish spaces $\Theta_1, \ldots, \Theta_k$. A *marginal set* is defined by

$$\Gamma((\pi_j)_{j \in J}) = \{\pi \in \Delta(\Theta) : \operatorname{marg}_j \pi = \pi_j \text{ for all } j \in J\},$$

  for some nonempty subset J of $\{1, \ldots, k\}$ and some probability measures $\pi_j \in \Delta(\Theta_j)$ for each j in J. Carroll (2017) studies a robust version of the multi-good monopoly problem in which the ambiguity set is the marginal set that fixes the valuation distribution for each good.[20]

- Suppose that Θ is a convex subset of $\mathbf{R}$. For any π in $\Delta(\Theta)$ and $\alpha \in [0, 1]$, let $Q_\alpha(\pi)$ denote the set of α-quantiles of π.[21] A *quantile set* is defined by

$$Q((x_j, \alpha_j)_{j=1}^{m}) = \{\pi \in \Delta(\Theta) : x_j \in Q_{\alpha_j}(\pi) \text{ for all } j = 1, \ldots, m\},$$

[20] Carroll (2017) shows that it is optimal to screen the agent independently along each dimension. Thus, posting separate prices for each good is maxmin optimal. This mechanism's payoff guarantee is not robust if the marginal valuation distribution for any good has an atom at the price posted for that good; see Proposition 2.3.

[21] That is, $Q_\alpha(\pi)$ contains all x for which $\pi(-\infty, x] \geq \alpha$ and $\pi[x, \infty) \geq 1 - \alpha$.

for some positive integer m and some $x_1, \ldots, x_m \in \mathbf{R}$ and $\alpha_1, \ldots, \alpha_m \in [0,1]$ satisfying $\inf \Theta < x_1 < \cdots < x_m < \sup \Theta$ and $\alpha_1 < \cdots < \alpha_m$.[22] The monopoly pricing example in Section 3.1 uses an ambiguity set of this form.

Some topological assumptions on the state space Θ are needed to show that these ambiguity sets are not globally robust. Indeed, if Θ has the discrete topology, then all ambiguity sets are globally robust.[23] The space Θ is *perfect* if it has no isolated points. The space Θ is *connected* if it cannot be expressed as a disjoint union of two nonempty open sets. If Θ is connected, then it is perfect.

Theorem 2 (Not globally robust)
Nonempty, proper subsets of $\Delta(\Theta)$ taking the following forms are not globally robust:

1. *probability sets and support–moment sets, provided that Θ is connected;*
2. *singletons, relative entropy balls, and total variation balls, provided that Θ is perfect;*
3. *marginal sets, provided that Θ is a product of perfect sets;*
4. *quantile sets, provided that Θ is a convex subset of $\mathbf{R}$.*

Suppose that the designer uses an ambiguity set taking one of these forms. Unless she can confidently rule out arbitrarily small state perturbations, she must independently check that the payoff guarantee from her proposed mechanism is robust.

The ambiguity sets in Theorem 2 do not reflect the topology on the state space. Consider an ambiguity set Π taking one of these forms. In the proof, we construct a set C that is assigned low probability under every prior in Π. We then construct a sequence (π_n) of priors concentrating on C that converges to a prior in Π.

[22] Here, $\inf \Theta \in [-\infty, \infty)$ and $\sup \Theta \in (-\infty, \infty]$.

[23] This does not mean that robustness has no bite in models with discrete types; see Remark 1.

5 Behavioral foundation for robustness

In this section, we axiomatize our notions of robustness and global robustness.

5.1 Maxmin preferences over acts

First, we formally define acts in a way that is suitable for the axiomatizations. (Here, we use the decision theory terminology of "acts" rather than "social choice functions.") Let $\Delta_0(X)$ denote the set of simple lotteries on the decision space X. An act is a measurable simple function $f\colon \Theta \to \Delta_0(X)$. Let $\mathcal{F}_0$ denote the set of acts. We identify each lottery in $\Delta_0(X)$ with the associated constant act.

A utility function $u\colon X \times \Theta \to \mathbf{R}$ is *state-measurable* (*state-continuous*, *state-bounded*) if, for each fixed x in X, the function $u(x, \cdot)$ is measurable (continuous, bounded) on Θ. We extend u linearly to the domain $\Delta_0(X) \times \Theta$. Given a state-measurable utility function u, each act f in $\mathcal{F}_0$ induces a measurable value function $v_f\colon \Theta \to \mathbf{R}$ defined by $v_f(\theta) = u(f(\theta), \theta)$. Let $\mathcal{F}_0(u)$ denote the set of acts f for which v_f is bounded. If u is state-bounded (in particular, if u is state-independent), then $\mathcal{F}_0(u) = \mathcal{F}_0$.

Given a state-measurable utility function u and a nonempty subset Π of $\Delta(\Theta)$, define the maxmin preference relation $\succsim_{(u,\Pi)}$ on $\mathcal{F}_0(u)$ by

$$f \succsim_{(u,\Pi)} g \iff W_\Pi(v_f) \geq W_\Pi(v_g),$$

where W_Π is defined in (2). Note that the value functions v_f and v_g depend on u.

5.2 Robustness of payoff guarantees

We axiomatize the robustness of a payoff guarantee. For this axiomatization, we allow for state-dependent utility. We first define a notion of limit for acts.

Definition 2. Given acts $f, g \in \mathcal{F}_0$, act g is a *graphical limit* of f if for each state θ in Θ, there exists a sequence (θ_n) converging to θ such that $f(\theta_n) = g(\theta)$

for all n.

Each act f can be expressed as $\sum_{j=1}^m x_j A_j$, for some $x_1, \ldots, x_m \in \Delta_0(X)$ and some measurable partition $(A_1, \ldots, A_m)$ of Θ; this means that $f(\theta) = x_j$ if θ is in A_j. The graphical limits of f are precisely the acts of the form $\sum_{j=1}^m x_j A'_j$ for some measurable partition $(A'_1, \ldots, A'_m)$ of Θ satisfying $A'_j \subset \bar{A}_j$ for each j in J.[24] If Θ is connected, then each nonconstant act f has a graphical limit g with $g \neq f$.

Theorem 3 (Robustness)
Let $\succsim = \succsim_{(u,\Pi)}$ for some state-continuous utility function u and some nonempty closed subset Π of $\Delta(\Theta)$. For each act f in $\mathcal{F}_0(u)$, the following are equivalent:

1. *the payoff guarantee from v_f over Π is robust;*

2. *for every g in $\mathcal{F}_0$, if g is a graphical limit of f, then $g \succsim f$.*

In this equivalence, condition 1 concerns the payoff from f at priors near Π; condition 2 concerns the preference between f and nearby acts. To build intuition, recall the robust monopoly pricing problem illustrated in the left panel of Figure 1. There, the ambiguity set Π contains all distributions with median λ. Let f (respectively, g) denote the act under which the good is sold at price λ if $\theta \geq \lambda$ (respectively, $\theta > \lambda$) and otherwise the good is not sold. Act f induces the value function $v_f = v^*_\lambda$ and is maxmin optimal, but the payoff guarantee from v_f over Π is not robust, as we showed in Section 3.1. Act g is a graphical limit of f and it induces the value function v_g, which agrees with v^*_λ in all states except $\theta = \lambda$, where $v_g(\lambda) = 0$. Thus, $W_\Pi(v_f) = \lambda/2$ and $W_\Pi(v_g) = 0$, so f is strictly preferred to g, contrary to condition 2 of Theorem 3.

It can be shown in general that for each graphical limit g of an act f, we have $v_g \geq \operatorname{lsc} v_f$, with equality for some graphical limit of f. Therefore, to prove Theorem 3, it suffices to show that for any act f and any closed ambiguity set Π, the payoff guarantee from v_f over Π is robust if and only if

[24] Here, $\bar{A}_j$ denotes the closure of A_j. In this definition, we allow elements of the partition $(A'_1, \ldots, A'_m)$ to be empty.

$W_\Pi(v_f) = W_\Pi(\operatorname{lsc} v_f)$. Intuitively, in each state θ, the value $\operatorname{lsc} v_f(\theta)$ reflects the lowest payoffs from f in states arbitrarily close to θ. The payoff guarantee from v_f is robust if and only if these low payoffs in nearby states are already taken into consideration when evaluating f.

Remark 4 (Continuity with respect to graphical limits)**.** Condition 2 of Theorem 3 is an upper semicontinuity axiom. One might expect a full continuity axiom requiring *indifference* between an act and any of its graphical limits. But this property is too restrictive. If Θ is perfect, then Θ can be partitioned into two dense sets. For any decisions $x, y \in X$, consider the act that equals x on one such dense set and y on the other. The constant acts x and y are each graphical limits of this act. Therefore, full continuity would require indifference between x and y, and hence over all of X.

5.3 Global robustness of ambiguity sets

We now axiomatize global robustness of an ambiguity set, under the assumption that the utility function is state-independent.[25] We introduce continuity axioms inspired by Γ-convergence. In the analysis of minimization problems, Γ-convergence has proven to be the most useful mode of convergence of real-valued functions (Braides, 2002, pp. 1–2). Here we define an analogous notion of Γ-convergence for acts.

We define Γ-convergence with respect to a relation $\succsim$ on $\mathcal{F}_0$ that is *monotone* in the following sense: for all $f, g \in \mathcal{F}_0$, if $f(\theta) \succsim g(\theta)$ for all θ in Θ, then $f \succsim g$. The maxmin preference relation $\succsim_{(u,\Pi)}$ is monotone for any state-independent utility function u and any ambiguity set Π.

Definition 3. Let $\succsim$ be a monotone relation on $\mathcal{F}_0$. Given acts $g, f_1, f_2, \ldots \in \mathcal{F}_0$, act g is a Γ-*limit* of the sequence (f_n) if for each state θ, the following hold:

[25] With maxmin preferences, state-dependent utility creates difficulties because the class of value functions induced by acts may not be well-behaved. Hill (2019) avoids this difficulty by working with a finite state space.

1. for some sequence (θ_n) converging to θ, there exists m such that $f_n(\theta_n) = g(\theta)$ for all $n \geq m$.

2. for any sequence (θ_n) converging to θ and any h in $\mathcal{F}_0$ with $g(\theta) \succ h$, there exists m such that $f_n(\theta_n) \succ h$ for all $n \geq m$.

We say that an act g is a Γ-limit of an act f if g is a Γ-limit of the constant sequence (f_n) with $f_n = f$ for each n. A Γ-limit is defined relative to a relation $\succsim$ on $\mathcal{F}_0$. The relation $\succsim$ should always be clear from context. Intuitively, part 2 of Definition 3 requires that $g(\theta)$ is the worst decision that is taken in states arbitrarily close to state θ arbitrarily far along the sequence (f_n). The worst decision is selected because this definition is tailored to minimization problems. A sequence (f_n) can have more than one Γ-limit, but all Γ-limits must be in the same indifference class.

Given a relation $\succsim$ on $\mathcal{F}_0$, a sequence (f_n) in $\mathcal{F}_0$ is *bounded* if there exist constant acts $x, y \in \Delta_0(X)$ such that $x \succsim f_n \succsim y$ for all n. Boundedness is defined with respect to a relation $\succsim$. The relation $\succsim$ should be clear from context. We next state two continuity axioms for a monotone relation $\succsim$ on $\mathcal{F}_0$.

Axiom 1 (Weak upper Γ-semicontinuity)**.** For any $f, g \in \mathcal{F}_0$, if g is a Γ-limit of f, then $g \succsim f$.

Axiom 2 (Lower Γ-semicontinuity)**.** For any bounded sequence (f_n) in $\mathcal{F}_0$ and any $g, h \in \mathcal{F}_0$, if $h \succsim f_n$, for all n, and g is a Γ-limit of (f_n), then $h \succsim g$.

Weak upper Γ-semicontinuity requires that each upper contour set of $\succsim$ is closed under Γ-limits of each act (i.e., each constant sequence of acts). Lower Γ-semicontinuity requires that each lower contour set of $\succsim$ is closed under Γ-limits of each bounded sequence of acts. It can be shown that a monotone relation $\succsim$ is weakly upper Γ-semicontinuous if and only if $\succsim$ satisfies condition 2 in Theorem 3 for each act f in $\mathcal{F}_0$.

A subset Π of $\Delta(\Theta)$ is *tight* if for every positive ε there exists a compact subset K of Θ such that for every π in Π, we have $\pi(K) \geq 1 - \varepsilon$. If Θ is compact, then every subset of $\Delta(\Theta)$ is trivially tight. Given acts $f, g \in \mathcal{F}_0$

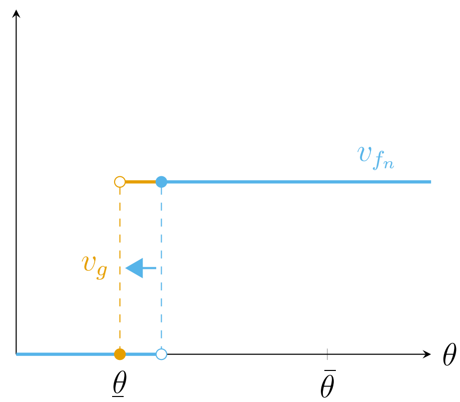

Figure 3. Failure of lower Γ-semicontinuity

and a subset A of Θ, let fAg denote the act that agrees with f on A and with g on $\Theta \setminus A$. Now we give an axiom for a relation $\succsim$ that will characterize tightness.

Axiom 3 (Tightness)**.** For any $f, g \in \mathcal{F}_0$ and any $x \in \Delta_0(X)$, if $f \succ g$, then there exists a compact subset K of Θ such that $fKx \succ g$ and $f \succ gKx$.

In words, for any strict preference relation between acts, there exists a sufficiently large compact set K such that the strict preference is preserved if one of the acts is replaced with a fixed decision in all states outside K.

Now we state our axiomatization of global robustness. Recall that if u is state-independent, then $\mathcal{F}_0 = \mathcal{F}_0(u)$, so the maxmin relation $\succsim_{(u,\Pi)}$ is defined on $\mathcal{F}_0$.

Theorem 4 (Global robustness)
Let $\succsim = \succsim_{(u,\Pi)}$ *for some nonconstant state-independent utility function* u *and some nonempty subset* Π *of* $\Delta(\Theta)$*. The following are equivalent:*

1. Π *is globally robust and tight;*

2. $\succsim$ *is weakly upper* Γ*-semicontinuous, lower* Γ*-semicontinuous, and tight.*

Figure 3 illustrates a failure of lower Γ-semicontinuity. Let $\Theta = \mathbf{R}_+$. Let Π be the set of priors that assign probability 1 to the *open* interval $(\underline{\theta}, \bar{\theta})$, where

$0 < \underline{\theta} < \bar{\theta}$. This set Π is tight, as can be seen by taking the compact set $K = [\underline{\theta}, \bar{\theta}]$. Let (θ_n) be a sequence converging downward to $\underline{\theta}$. Fix a price p. Let f_n (respectively, g) specify that the good is sold at price p if and only if $\theta \geq \theta_n$ (respectively, $\theta > \underline{\theta}$). Thus, g is a Γ-limit of the sequence (f_n). The corresponding value functions v_{f_n} and v_g are plotted in Figure 3. Note that v_g is lower semicontinuous even though each v_{f_n} is not. The payoff guarantee over Π jumps up at this Γ-limit because the undesirable consequence—no sale—escapes the set Π, which is not closed.

Remark 5 (Γ-continuity)**.** Theorem 4 requires *weak* upper Γ-semicontinuity but full lower Γ-semicontinuity. This asymmetry arises because Γ-limits are asymmetric and favor downward jumps. Appendix B.2 gives an alternative axiomatization in which weak upper Γ-semicontinuity is strengthened to upper Γ-semicontinuity, and global robustness is strengthened to *uniform robustness*. Uniform robustness requires the inequality (3) to hold uniformly over bounded sets of value functions. Moreover, the ambiguity sets in Theorem 1 are uniformly robust, as we show in the proof.

Remark 6 (Monotone continuity and variational preferences)**.** In Appendix B.1, we show that weak upper Γ-continuity is incompatible with Arrow's (1970) monotone continuity axiom. This provides a direct way to check that certain classes of preferences violate weak upper Γ-semicontinuity. For example, Maccheroni et al. (2006, Theorem 13, pp. 1460–1461) axiomatize the subclass of variational preferences satisfying monotone continuity. Preferences in this subclass violate weak upper Γ-semicontinuity. This subclass includes the multiplier preferences of Hansen and Sargent (2001), which are axiomatized in Strzalecki (2011).

Stanca (2023) is the only other paper we are aware of that provides a behavioral foundation for a notion of robustness in a setting with a topological state space.[26] In his model, the decision space is Euclidean and the decision-maker (DM) maximizes subjective expected utility over a menu of continuous

[26] Prasad (2003) gives examples of Bayesian decision problems in which the maximal payoff is discontinuous in the prior, with respect to the weak topology.

acts. Stanca (2023) defines such a menu to be robust if for any sequence of priors converging to the reference prior, the DM's maximal expected utility converges to the DM's maximal expected utility under the reference prior. Crucially, the DM's choice from the menu can vary with the prior. By contrast, we are interested in the the designer's payoff guarantee from a fixed mechanism when the priors in the ambiguity set are perturbed.

6 Related literature

Our paper refines the maxmin approach to distributional robustness in economic design.[27] Carroll (2019) provides a taxonomy of different robustness concepts.[28] Some of the related literature is discussed in the main text. In particular, Section 3.1 gives examples from the literature of mechanisms whose payoff guarantees are not robust, and Section 4 references ambiguity sets from the literature that are globally robust. Here, we discuss examples of maxmin design problems in which the ambiguity set is not globally robust but the payoff guarantee from the maxmin-optimal mechanism is robust in at least some cases.

In Bergemann and Schlag's (2008) seminal paper on robust monopoly pricing, the ambiguity set is a support set, which is not globally robust (by Theorem 2). If the left endpoint of the support is high enough, then the regret-minimizing price distribution has an atom on this left endpoint. In this case, the policy's payoff guarantee is not robust.[29] Brooks and Du (2021a)

[27] There is also an extensive literature on distributional robustness in statistics and operations research; see Rahimian and Mehrotra (2022) for a survey.

[28] In particular, there is large literature on robustness to uncertainty about beliefs; see Bergemann et al. (2013) for a survey. Our focus on small perturbations of the prior is related to work on small perturbations of beliefs. Jehiel et al. (2012) and Chen et al. (2023) require robustness only for beliefs that are close to some benchmark belief. Meyer-ter Vehn and Morris (2011), Oury and Tercieux (2012), Oury (2015), and Chen et al. (2022) also allow for payoff-type uncertainty.

[29] If the left endpoint of the support is low enough, then the regret-minimizing pricing policy is atomless, so its payoff guarantee is robust. In Bergemann and Schlag (2011), the payoff guarantee from the maxmin-optimal policy is robust. More recent work has explored dynamic extensions of Bergemann and Schlag's (2008) static framework. In the two period model of Handel and Misra (2015), the seller updates her ambiguity set in the second period

and Brooks and Du (2021b) study robust multi-agent auction settings with uncertainty over the information structure,[30] the value distribution, and the equilibrium selection. In Brooks and Du (2021a), the ambiguity set restricts the support of the common value. In Brooks and Du (2021b), the ambiguity set restricts the mean of each agent's private valuation distribution. In each paper, the maxmin-optimal mechanism takes the form of a "proportional auction." The designer's induced value function is continuous, so the resulting payoff guarantee is robust, by Proposition 2. In a similar set-up, Brooks and Du (2023) derive a proportional cost-sharing mechanism as the robust solution of a public goods problem. Their ambiguity set restricts the support of the sum of the agents' valuations.

Carroll (2015) studies a robust moral hazard problem. The principal knows that certain actions are feasible, but she is uncertain about which additional actions are feasible. Thus, the "state" is the realized set of feasible actions, and the ambiguity set is a support set, where the support contains all supersets of the known-action set. By Theorem 2, this ambiguity set is not globally robust. Carroll (2015) restricts attention to contracts that specify a single mapping from output to wages. It can be shown that the payoff guarantee from such a contract is robust. Suppose instead that the principal could offer a menu of wage contracts to screen the agent's privately known feasible set. The payoff guarantee from such a menu can be non-robust because the agent's choice from the menu can change when his feasible action set is perturbed.

Next, we discuss alternative notions of robustness against misspecification of the environment.

In a nonlinear pricing setting, Madarász and Prat (2017) study the effect of local perturbations of the type distribution. They formulate and parameterize a notion of closeness between the true model and the designer's misspecified

based on the buyer's behavior in the first period. In Ilut et al. (2020), the ambiguity set is not updated, but nature selects a new worst-case prior each period after new information is observed. A similar updating rule is analyzed in Auster et al. (2024).

[30] In Hinnosaar and Kawai (2020) and Libgober and Mu (2021), the designer is also uncertain about the agent's information structure, but the ambiguity set comprises the whole space.

model. They show how a mechanism that is optimal with respect to a misspecified model can perform poorly under the true model, no matter how small the misspecification. They propose an alternative mechanism that performs well under the perturbations they consider. By contrast, we study a general decision problem, and we seek a principled foundation for maxmin preferences that are robust to small perturbations.

Pei and Strulovici (2024) propose a mechanism that implements social choice functions in a way that is robust to a small risk of large preference perturbations. They focus on uncertainty about agents' higher order beliefs.

Finally, Cerreia-Vioglio et al. (2024) axiomatize a class of variational preferences in which the cost function measures proximity to a fixed set of priors.[31] They interpret such a cost function as providing a "protective belt" against misspecification.[32] Unlike our paper, Cerreia-Vioglio et al. (2024) do not consider a topology on the state space. Their cost functions infinitely penalize priors that are not absolutely continuous with respect to one of the fixed priors. As a result, the protective belt does not generally cover perturbations of the prior that result from perturbations of the states in the support of the prior. Analyzing the implications of our axioms within the framework of variational preferences is an interesting direction for future work.

7 Conclusion

This paper refines the maxmin approach to distributionally robust mechanism design. We argue for payoff guarantees that are approximately preserved if the ambiguity set is slightly misspecified. Our main innovation is to take into account the topology on the state space. As an illustration of our results, consider a monopolist who engages a consultant to develop a new pricing mechanism to robustly maximize revenue. The consultant must gather some partial

[31] Lanzani (2024) also uses variational preferences to model misspecification. In his model, the agent dynamically adjusts his concern for misspecification in response to observed data.

[32] The most tractable examples of such cost functions are Hausdorff distances to the fixed set of priors, where the distance between priors is given by a divergence. In Hansen and Sargent (2022), the cost function is the KL-distance to a single fixed prior.

information about the distribution of consumer valuations. Our results offer a justification for seeking a payoff guarantee over all distributions consistent with the consultant's estimates of certain moments of the consumer valuation distribution, rather than over all distributions consistent with estimated quantiles or bounds on the support.

A Main proofs

A.1 Mathematical preliminaries

Signed measures All signed measures on Θ are defined on the Borel σ-algebra $\mathcal{B}(\Theta)$. For any signed measure μ on Θ, define the total variation norm $\|\mu\|_{\text{TV}}$ by

$$\|\mu\|_{\text{TV}} = \frac{1}{2} \sup_{E \in \mathcal{P}} \sum_{E \in \mathcal{P}} |\mu(E)|,$$

where the supremum is over all finite measurable partitions $\mathcal{P}$ of Θ. It can be shown that

$$\|\mu\|_{\text{TV}} \leq \sup_{A} |\mu(A)|, \tag{4}$$

where the supremum is over all sets A in $\mathcal{B}(\Theta)$. If $\mu(\Theta) = 0$, then (4) holds with equality (and the equality still holds if $|\mu(A)|$ is replaced with $\mu(A)$ or $-\mu(A)$).

By the Jordan decomposition theorem, each signed measure μ can be uniquely expressed as $\mu = \mu_+ - \mu_-$ for some nonnegative measures μ_+ and μ_- that are mutually singular.[33] Here, we extend the notation $\langle \cdot, \cdot \rangle$ to integration against signed measures. Given a measurable function $v \colon \Theta \to \mathbf{R}$ and a signed measure μ on Θ, define the integral of v against μ by $\langle v, \mu \rangle = \langle v, \mu_+ \rangle - \langle v, \mu_- \rangle$, provided that v is absolutely integrable with respect to both μ_+ and μ_-. The *support* of μ, denoted $\operatorname{supp} \mu$, is defined to be the support of the associated nonnegative measure $\mu_+ + \mu_-$.[34]

[33] That is, there exists Borel subset A of Θ such that $\mu_+(A) = 0$ and $\mu_-(\Theta \setminus A) = 0$.

[34] Recall that the support of a nonnegative measure on a Polish space is the complement of the largest open set with measure 0 (which can be shown to exist).

A signed measure μ on Θ is *bounded* if $\|\mu\|_{\mathrm{TV}} < \infty$. A function $v\colon \Theta \to \mathbf{R}$ is bounded if $\|v\|_\infty \coloneqq \sup_{\theta \in \Theta} |v(\theta)| < \infty$. For any bounded, measurable function $v\colon \Theta \to \mathbf{R}$ and any bounded signed measure μ, the integral $\langle v, \mu \rangle$ is well-defined and satisfies $\langle v, \mu \rangle \leq 2\|v\|_\infty \|\mu\|_{\mathrm{TV}}$.

Probability kernels A *probability kernel* on Θ is a map $\kappa\colon \Theta \times \mathcal{B}(\Theta) \to [0,1]$ such that (i) for each θ in Θ, the map $A \mapsto \kappa(\theta, A)$ is a probability measure; and (ii) for each A in $\mathcal{B}(\Theta)$, the map $\theta \mapsto \kappa(\theta, A)$ is measurable. Given a probability measure π in $\Delta(\Theta)$ and probability kernel κ on Θ, define the push-forward measure $\pi\kappa$ in $\Delta(\Theta)$ by

$$(\pi\kappa)(A) = \int_\Theta \kappa(\theta, A)\, \mathrm{d}\pi(\theta), \qquad A \in \mathcal{B}(\Theta).$$

Given a measurable function $v\colon \Theta \to \mathbf{R}$, define the function $\kappa v\colon \Theta \to \mathbf{R}$ by

$$(\kappa v)(\theta) = \int_\Theta v(\theta')\, \mathrm{d}\kappa_\theta(\theta'), \qquad \theta \in \Theta,$$

where κ_θ denotes the measure $\kappa(\theta, \cdot)$. If Θ is a finite set with n elements, then we can represent measures as row n-vectors, functions as column n-vectors, and kernels as $n \times n$ matrices. In this case, our notation is consistent with matrix multiplication.

Weak convergence and transportation Let d be a compatible metric on Θ. Let $B_\varepsilon(\theta) = \{\theta' \in \Theta : d(\theta, \theta') \leq \varepsilon\}$. For any probability kernel κ, let

$$\|\kappa\|_d = \inf\{\varepsilon : \kappa(\theta, B_\varepsilon(\theta)) = 1 \text{ for all } \theta \in \Theta\}.$$

Lemma 1 (Transport kernels)
Let d be a bounded, compatible metric on Θ. Let (π_n) be a sequence in $\Delta(\Theta)$ and let π be in $\Delta(\Theta)$. If (π_n) weakly converges to π, then the following hold:

1. *there exists a sequence (κ_n) of probability kernels such that $\|\kappa_n\|_d \to 0$ and $\|\pi_n - \pi\kappa_n\|_{\mathrm{TV}} \to 0$;*

2. *there exists a sequence* (κ'_n) *of probability kernels such that* $\|\kappa'_n\|_d \to 0$ *and* $\|\pi_n \kappa'_n - \pi\|_{\mathrm{TV}} \to 0$.

The proofs of all lemmas appear in Appendix B.

Γ-convergence of functions When analyzing minimization problems, the most convenient mode of convergence for objective functions is Γ-convergence; see Braides (2002) for a textbook treatment.

Definition 4 (Γ-convergence)**.** Let X be an arbitrary metric space. Let $c, c_1, c_2, \dots$ be real-valued functions on X. The sequence (c_n) Γ-*converges* to c, and we write $\Gamma\text{-}\lim c_n = c$, if for each x in X, the following hold:

1. for every sequence (x_n) converging to x, we have $\liminf_n c_n(x_n) \geq c(x)$;

2. for some sequence (x_n) converging to x, we have $\limsup_n c_n(x_n) \leq c(x)$.

Recall that the map $\theta \mapsto \delta_\theta$ embeds the space Θ in $\Delta(\Theta)$. Under this embedding, any bounded, measurable function $v\colon \Theta \to \mathbf{R}$ can be extended to the domain $\Delta(\Theta)$ via the map $\langle v, \cdot \rangle$. The portmanteau theorem says that the topological properties of a function v on Θ transfer to its extension $\langle v, \cdot \rangle$ on $\Delta(\Theta)$. Our next result says that the Γ-convergence of functions on Θ similarly transfers to their extensions on $\Delta(\Theta)$.

A sequence (v_n) of real-valued functions on Θ is *bounded* if $\sup_n \|v_n\|_\infty < \infty$.

Lemma 2 (Γ-portmanteau)
Let (v_n) *be a bounded sequence in* $B(\Theta)$*. If* $\Gamma\text{-}\lim v_n = v$*, then* $\Gamma\text{-}\lim \langle v_n, \cdot \rangle = \langle v, \cdot \rangle$.

Recall that $\operatorname{lsc} v$ denotes the lower semicontinuous envelope of v. It can be checked that $\Gamma\text{-}\lim_n v = \operatorname{lsc} v$. By Lemma 2, $\Gamma\text{-}\lim_n \langle v, \cdot \rangle = \langle \operatorname{lsc} v, \cdot \rangle$.

Convergence of acts We relate the convergence of acts to the Γ-convergence of their induced value functions.

Lemma 3 (Graphical limits)
Let u be state-continuous. For any act f in $\mathcal{F}_0(u)$, there exists an act g in $\mathcal{F}_0(u)$ such that g is a graphical limit of f and $v_g = \operatorname{lsc} v_f$.

Given a relation $\succsim$ on $\Delta_0(X)$, a subset F of $\Delta_0(X)$ is *indifference-free* if for all x, y in F, we have $x \not\sim y$.

Lemma 4 (Γ-limits of acts)
Let $\succsim = \succsim_{(u,\Pi)}$ for some nonconstant, state-independent utility function u and some nonempty subset Π of $\Delta(\Theta)$. Let (f_n) be a sequence in $\mathcal{F}_0$.

(i) Suppose that $\cup_n \{f_n(\theta) : \theta \in \Theta\}$ is finite and indifference-free. If $(u \circ f_n)$ is Γ-convergent, then (f_n) has a Γ-limit in $\mathcal{F}_0$.

(ii) If g in $\mathcal{F}_0$ is a Γ-limit of (f_n), then $\Gamma\text{-}\lim_n u \circ f_n = u \circ g$.

A.2 Proof of Proposition 2

Let v be a value function in $B(\Theta)$.

1. First, suppose that v is continuous. Fix an ambiguity set Π. Let (π_n) be a sequence in $\Delta(\Theta)$ that converges to some prior π in the closure of Π. By the definition of the closure, we can choose a sequence (π'_n) in Π that converges to π. By the portmanteau theorem,

$$\lim_n \langle v, \pi_n \rangle = \langle v, \pi \rangle = \lim_n \langle v, \pi'_n \rangle \geq W_\Pi(v).$$

Thus, the payoff guarantee from v over Π is robust.

Conversely, suppose that v is discontinuous. Then for some $\varepsilon > 0$, there exists a point θ in Θ and a sequence (θ_n) converging to θ such that either (i) $v(\theta_n) \geq v(\theta) + \varepsilon$ for all n, or (ii) $v(\theta_n) \leq v(\theta) - \varepsilon$ for all n. If (i) holds, let $\Pi_1 = \{\delta_{\theta_n} : n \geq 1\}$. In this case, the payoff guarantee from v over Π_1 is not robust because δ_θ is in the closure of Π_1. If (ii) holds, let $\Pi_2 = \{\delta_\theta\}$. In this case, the payoff guarantee from v over Π_2 is not robust because the sequence (δ_{θ_n}) converges to δ_θ.

2. First, suppose that v is lower semicontinuous. Fix a closed ambiguity set Π. Let (π_n) be a sequence in $\Delta(\Theta)$ that converges to some prior π in the closure of Π. Since Π is closed, π is in Π. By the portmanteau theorem,

$$\liminf_n \langle v, \pi_n \rangle \geq \langle v, \pi \rangle \geq W_\Pi(v).$$

Thus, the payoff guarantee from v over Π is robust.

Conversely, suppose that v is not lower semicontinuous. Then for some $\varepsilon > 0$, there exists a point θ in Θ and a sequence (θ_n) converging to θ such that $v(\theta_n) \leq v(\theta) - \varepsilon$ for all n. The payoff guarantee from v over the singleton $\{\delta_\theta\}$ is not robust at v because the sequence (δ_{θ_n}) converges to δ_θ.

3. Fix π_0 in $\Delta(\Theta)$. First suppose that $\langle v, \pi_0 \rangle = \langle \operatorname{lsc} v, \pi_0 \rangle$. By Lemma 2, $\Gamma\text{-}\lim_n \langle v, \cdot \rangle = \langle \operatorname{lsc} v, \cdot \rangle$, so for any sequence (π_n) in $\Delta(\Theta)$ that converges to π_0, we have

$$\liminf_n \langle v, \pi_n \rangle \geq \langle \operatorname{lsc} v, \pi_0 \rangle = \langle v, \pi_0 \rangle.$$

Thus, the payoff guarantee from v over $\{\pi_0\}$ is robust.

Conversely, suppose that $\langle v, \pi_0 \rangle > \langle \operatorname{lsc} v, \pi_0 \rangle$. By Lemma 2, $\Gamma\text{-}\lim_n \langle v, \cdot \rangle = \langle \operatorname{lsc} v, \cdot \rangle$, so there exists a sequence (π_n) in $\Delta(\Theta)$ converging to π_0 such that

$$\limsup_n \langle v, \pi_n \rangle \leq \langle \operatorname{lsc} v, \pi_0 \rangle < \langle v, \pi_0 \rangle.$$

Thus, the payoff guarantee from v over $\{\pi_0\}$ is not robust.

A.3 Proof of Theorem 1

An ambiguity set Π has the *total variation approximation property* if for each sequence (π_n) in $\Delta(\Theta)$ converging to a prior in the closure of Π and for each $\varepsilon > 0$, there exists a sequence (ρ'_n) in Π such that

$$\limsup_n \|\rho'_n - \pi_n\|_{\mathrm{TV}} \leq \varepsilon. \tag{5}$$

We check that if Π has the total variation approximation property, then Π is uniformly robust (see Appendix B.2). Let (π_n) be a sequence in $\Delta(\Theta)$ that converges to a prior in the closure of Π. Fix $\varepsilon > 0$. By the total variation approximation property, we can choose a sequence (ρ'_n) in Π satisfying (5). For each v in $B(\Theta)$, we have

$$\begin{aligned}\langle v, \pi_n \rangle &= \langle v, \rho'_n \rangle - \langle v, \rho'_n - \pi_n \rangle \\ &\geq W_\Pi(v) - 2\|v\|_\infty \|\rho'_n - \pi_n\|_{\mathrm{TV}}.\end{aligned}$$

Rearranging and then taking the infimum over all v in $B(\Theta)$ with $\|v\|_\infty \leq 1$, we have

$$\inf_v \left[\langle v, \pi_n \rangle - W_\Pi(v)\right] \geq -2\|\rho'_n - \pi'_n\|_{\mathrm{TV}}.$$

Take the limit infimum in n and apply (5) to get

$$\liminf_n \left(\inf_v \left[\langle v, \pi_n \rangle - W_\Pi(v)\right]\right) \geq -2\varepsilon.$$

Since ε was arbitrary, we conclude that Π is uniformly robust.

We now prove that any ambiguity set taking one of the forms in the theorem statement has the total variation approximation property. We use the following result to deal with unbounded functions.

Lemma 5 (Unbounded moment approximation)
Let (π_n) and (π'_n) be sequences in $\Delta(\Theta)$ that weakly converge to the same prior π in $\Delta(\Theta)$. Let $H\colon \Theta \to \mathbf{R}_+$ be continuous. For each $\varepsilon > 0$, there exists a sequence (ρ_n) in $\Delta(\Theta)$ weakly converging to π such that

(i) $\limsup_n \|\rho_n - \pi_n\|_{\mathrm{TV}} \leq \varepsilon$;

(ii) H *is bounded on* $\cup_n \operatorname{supp}(\rho_n - \pi'_n)$;

(iii) for any continuous function $h\colon \Theta \to \mathbf{R}$ *satisfying* $|h| \leq H$, *we have* $\langle h, \rho_n - \pi'_n \rangle \to 0$.

Roughly, we choose each ρ_n to agree with π'_n when H is very large and with π_n otherwise.

Continuous moment sets Let $\Pi = M(g, Y)$ for some continuous function $g\colon \Theta \to \mathbf{R}^m$ and some subset Y of $\mathbf{R}^m$. Without loss, we may assume that $Y \subset \operatorname{conv} g(\Theta)$. Assume first that Y is uniformly g-interior, i.e., for some $\delta > 0$, we have $Y^\delta \cap \operatorname{aff} g(\Theta) \subset \operatorname{conv} g(\Theta)$. Let (π_n) be a sequence in $\Delta(\Theta)$ that weakly converges to a prior π in the closure of Π. Choose a sequence (π'_n) in Π that weakly converges to π.

Let $H(\theta) = \|g(\theta)\|$, where $\|\cdot\|$ denotes the Euclidean norm on $\mathbf{R}^m$. Fix $\varepsilon > 0$. It follows from Lemma 5 that there exists a sequence (ρ_n) in $\Delta(\Theta)$ converging to π such that (i) $\limsup_n \|\rho_n - \pi_n\|_{\mathrm{TV}} \leq \varepsilon$; (ii) H is bounded on $\operatorname{supp}(\rho_n - \pi'_n)$ for each n; and (iii) $\lim_n \langle g_j, \rho_n - \pi'_n \rangle = 0$ for each $j = 1, \ldots m$. Each prior π'_n is in Π, so each function g_j is absolutely integrable with respect to π'_n and hence, by (ii), also with respect to ρ_n. We conclude from (iii) that $\lim_n |\langle g_j, \rho_n \rangle - \langle g_j, \pi'_n \rangle| = 0$, for each $j = 1, \ldots, m$.

We construct a sequence (ρ'_n) in Π by modifying the sequence (ρ_n). For each n, let $x_n = \langle g, \rho_n \rangle$ and $y_n = \langle g, \pi'_n \rangle$, and let[35]

$$z_n = y_n + \delta \frac{y_n - x_n}{\|y_n - x_n\|} \in B(y_n, \delta) \cap \operatorname{aff} g(\Theta) \subset \operatorname{conv} g(\Theta). \tag{6}$$

By Carathéodory's theorem, there exists a probability measure ζ_n in $\Delta(\Theta)$ supported on at most $m + 1$ points of Θ such that $\langle g, \zeta_n \rangle = z_n$. Let

$$\rho'_n = (1 - \alpha_n)\rho_n + \alpha_n \zeta_n, \quad \text{where} \quad \alpha_n = \frac{\|y_n - x_n\|}{\delta + \|y_n - x_n\|}.$$

Since $\|x_n - y_n\| \to 0$, we have $\alpha_n \to 0$. For each n, some algebra shows that[36]

$$\langle g, \rho'_n \rangle = (1 - \alpha_n)x_n + \alpha_n z_n = y_n \in Y, \tag{7}$$

so ρ'_n is in Π. For each n, the triangle inequality gives

$$\|\rho'_n - \pi_n\|_{\mathrm{TV}} \leq \alpha_n + \|\rho_n - \pi_n\|_{\mathrm{TV}},$$

[35] Here, we use the convention that $0/\|0\| = 0$. Thus, $z_n = y_n$ if $x_n = y_n$.

[36] Here, we are extending the notation $\langle \cdot, \cdot \rangle$ to integration of $\mathbf{R}^m$-valued functions.

so $\limsup_n \|\rho'_n - \pi_n\|_{\mathrm{TV}} \leq \varepsilon$, as desired.

Next, we prove the same conclusion under a different interiority assumption on Y. A set Y is *star g-interior* if there exists a relative interior point y_0 of $\operatorname{conv} g(\Theta)$ such that $[y_0, y] \subset Y$ for all y in Y. If we assume that Y is star g-interior (rather than uniformly g-interior), then the proof above goes through if we make a few modifications: (6) becomes

$$z_n = y_0 + \delta \frac{y_n - x_n}{\|y_n - x_n\|} \in B(y_0, \delta) \cap \operatorname{aff} g(\Theta) \subset \operatorname{conv} g(\Theta),$$

and (7) becomes

$$\langle g, \rho'_n \rangle = (1 - \alpha_n) x_n + \alpha_n z_n = (1 - \alpha_n) y_n + \alpha_n y_0 \in [y_0, y_n] \subset Y.$$

Since the finite union of robust ambiguity sets is robust, we conclude that $M(g, Y)$ is robust if g is continuous and the relevant constraint set $Y \cap \operatorname{conv} g(\Theta)$ can be expressed as a finite union of sets, each of which is uniformly g-interior or star g-interior.

Metric balls Given a compatible metric d on Θ, let $\Pi = B_W(\pi_0, r)$, for some prior π_0 in $\Delta(\Theta)$ and some radius $r > 0$. Let (π_n) be a sequence in $\Delta(\Theta)$ that weakly converges to a prior π in the closure of Π. Choose a sequence (π'_n) in Π that weakly converges to π.

Choose θ_0 in Θ. Let $H(\theta) = d(\theta_0, \theta)$. Fix $\varepsilon > 0$. By Lemma 5, there exists a sequence (ρ_n) in $\Delta(\Theta)$ converging to π such that (i) $\limsup_n \|\rho_n - \pi_n\|_{\mathrm{TV}} \leq \varepsilon$, and (ii) H is bounded on $\cup_n \operatorname{supp}(\rho_n - \pi'_n)$. By (ii), it can be shown that $W(\rho_n, \pi'_n) \to 0$.[37]

[37] Recall that a compatible metric d on Θ has been chosen. Following Bogachev (2018, p. 109), define the Kantorovich–Rubinstein norm $\|\cdot\|_{\mathrm{KR}}$ on the space of bounded signed measures by $\|\mu\|_{\mathrm{KR}} = \sup_f \langle f, \mu \rangle$, where the supremum is over all 1-Lipschitz functions $f \colon \Theta \to \mathbf{R}$ with $\|f\|_\infty \leq 1$. By Bogachev (2018, 3.2.2 Theorem, p. 111), the norm $\|\cdot\|_{\mathrm{KR}}$ induces the weak topology on the space of probability measures. Using Bogachev (2018, 3.2.7 Theorem, p. 114), it can be shown that for any priors $\mu, \nu \in \Delta(\Theta)$ that both concentrate on a subset S of Θ, we have $\|\mu - \nu\|_{\mathrm{KR}} \leq \max\{\operatorname{diam} S, 1\} W(\mu, \nu)$, where $\operatorname{diam} S = \sup_{\theta, \theta' \in S} d(\theta, \theta')$.

By (ii), the supremum of H over $\cup_n \operatorname{supp}(\rho_n - \pi'_n)$ is finite. Denote this supremum by L.

We construct a sequence (ρ'_n) in Π by modifying the sequence (ρ_n). For each n, let

$$\rho'_n = (1-\alpha_n)\rho_n + \alpha_n \pi_0, \quad \text{where} \quad \alpha_n = \min\{r^{-1} W(\rho_n, \pi'_n), 1\}.$$

Since $W(\rho_n, \pi'_n) \to 0$, it follows that $\alpha_n \to 0$. We claim that ρ'_n is in Π for each n. If $\alpha_n = 1$, this is immediate, so suppose that $\alpha_n < 1$. It can be checked that the Wasserstein metric W is convex in each of its arguments, so

$$\begin{aligned} W(\rho'_n, \pi_0) &\leq (1-\alpha_n) W(\rho_n, \pi_0) \\ &\leq (1-\alpha_n)[W(\rho_n, \pi'_n) + W(\pi'_n, \pi_0)] \\ &\leq (1-\alpha_n)[W(\rho_n, \pi'_n) + r] \\ &\leq r + W(\rho_n, \pi'_n) - \alpha_n r \\ &\leq r. \end{aligned}$$

Hence, ρ'_n is in Π. For each n, the triangle inequality gives

$$\|\rho'_n - \pi_n\|_{\mathrm{TV}} \leq \alpha_n + \|\rho_n - \pi_n\|_{\mathrm{TV}},$$

so $\limsup_n \|\rho'_n - \pi_n\|_{\mathrm{TV}} \leq \varepsilon$, as desired.

Now, let $\Pi = B_P(\pi_0, r)$ for some prior π_0 in $\Delta(\Theta)$ and some radius $r > 0$. The Prokhorov metric induces the weak topology on $\Delta(\Theta)$, so it is immediate that $P(\pi_n, \pi'_n) \to 0$. We can take $\rho_n = \pi_n$ for each n, and the rest of the proof goes through as above, with P in place of W, since the Prokhorov metric is convex in each of its arguments.

For each n, consider the Jordan decomposition $\rho_n - \pi'_n = \mu^n_+ - \mu^n_-$. For each n, the measures μ^n_+ and μ^n_- concentrate on the ball $B(\theta_0, L)$, which has diameter at most $2L$. Therefore,

$$\begin{aligned} W(\rho_n, \pi'_n) &= W(\mu^n_+, \mu^n_-) \\ &\leq \max\{2L, 1\} \|\mu^n_+ - \mu^n_-\|_{\mathrm{KR}} \\ &\leq \max\{2L, 1\}(\|\rho_n - \pi\|_{\mathrm{KR}} + \|\pi'_n - \pi\|_{\mathrm{KR}}), \end{aligned}$$

where the last line uses the triangle inequality. Since (ρ_n) and (π'_n) each weakly converge to π, the right side tends to 0 as $n \to \infty$. We conclude that $W(\rho_n, \pi'_n) \to 0$, as desired.

A.4 Proof of Theorem 2

We separate the proof into parts. Some parts use the following approximation result. For any subset S of Θ, we view $\Delta(S)$ as a subset of $\Delta(\Theta)$.

Lemma 6 (Concentrated approximation)
Suppose that Θ is perfect. For any prior π in $\Delta(\Theta)$ there exist a subset A of Θ with $\pi(A) = 1$ and a sequence (π_n) in $\Delta(\Theta \setminus A)$ that weakly converges to π.

Probability sets Suppose that Θ is connected. Let $\Pi = P(A, \alpha, \beta)$ some measurable subset A of Θ and some $\alpha, \beta \in [0, 1]$ with $\alpha \leq \beta$. Suppose that Π is a nonempty proper subset of $\Delta(\Theta)$. It follows that A is a nonempty proper subset of Θ and $(\alpha, \beta) \neq (0, 1)$. We may assume without loss that $\alpha > 0$; otherwise, we must have $\beta < 1$, and we can express $P(A, \alpha, \beta)$ as $P(A^c, 1 - \beta, 1 - \alpha)$. Let $v = 1_A$. Thus, $W_\Pi(v) = \alpha$. We show that the payoff guarantee from v over Π is not robust. Since Θ is connected, A cannot be both open and closed. We consider two (overlapping) cases.

First, suppose that A is not open. Choose a sequence (θ_n) in $\Theta \setminus A$ that converges to some point θ in A. Fix $\theta' \in \Theta \setminus A$. For each n, let $\pi_n = \alpha\delta_{\theta_n} + (1 - \alpha)\delta_{\theta'}$. The sequence (π_n) converges to $\pi = \alpha\delta_\theta + (1 - \alpha)\delta_{\theta'} \in \Pi$, but for all n, we have $\langle v, \pi_n \rangle = 0 < \alpha$.

Second, suppose that A is not closed. Choose a sequence (θ_n) in A that converges to some point θ in $\Theta \setminus A$. For each n, let $\pi_n = \alpha\delta_{\theta_n} + (1 - \alpha)\delta_\theta \in \Pi$. The sequence (π_n) converges to δ_θ, but $\langle v, \delta_\theta \rangle = 0 < \alpha$.

Support–moment sets Suppose that Θ is connected. Let $\Pi = M(S; g, Y)$ for some measurable proper subset S of Θ, some continuous function $g \colon \Theta \to \mathbf{R}^m$, and some subset Y of $\mathbf{R}^m$ that intersects the relative interior of $\operatorname{conv} g(S)$. Choose $y_0 \in Y \cap \operatorname{relint}(\operatorname{conv} g(S))$. It can be shown that there exists $\delta > 0$ such that (a) $D := B(y_0, \delta) \cap \operatorname{aff} g(S) \subset \operatorname{conv} g(S)$, and (b) there exists a continuous function $\tilde{g} \colon D \to \Delta(S)$ such that $\langle g, \tilde{g}(z) \rangle = z$, for each z in D.[38]

[38]By the definition of the relative interior, we may choose $\delta' > 0$ such that $B(y_0, \delta') \cap \operatorname{aff} g(S) \subset \operatorname{conv} g(S)$. Choose a maximal collection $z_1, \ldots, z_k$ of affinely independent vectors

Using $\tilde{g}$, we define a map $\tilde{\pi}\colon \Theta \to \Delta(\Theta)$ as follows. For each θ in Θ, let[39]

$$z(\theta) = y_0 + \delta \frac{y_0 - g(\theta)}{\|y_0 - g(\theta)\|} \in B(y_0, \delta) \cap \operatorname{aff} g(S) \subset \operatorname{conv} g(S),$$

and let

$$\tilde{\pi}(\theta) = (1 - \alpha(\theta))\delta_\theta + \alpha(\theta)\tilde{g}(z(\theta)), \quad \text{where} \quad \alpha(\theta) = \frac{\|y_0 - g(\theta)\|}{\delta + \|y_0 - g(\theta)\|}.$$

For each θ in Θ, some algebra shows that

$$\langle g, \pi \rangle = (1 - \alpha)g(\theta) + \alpha z(\theta) = y_0 \in Y.$$

Thus, $\tilde{\pi}(\theta)$ is in Π if θ is in S. The map $\tilde{\pi}$ is continuous because g and $\tilde{g}$ are continuous and $\alpha(\theta_n) \to 0$ for any sequence (θ_n) with $g(\theta_n) \to y_0$.

We now complete the proof. Let $v = 1_S$. We have $W_\Pi(v) = 1$. We show that the payoff guarantee from v over Π is not robust. Since Θ is connected, S cannot be both open and closed. We consider two (overlapping) cases.

First, suppose that S is not open. Choose a sequence (θ_n) in $\Theta \setminus S$ that converges to some point θ in S. Let

$$\begin{aligned} \pi &= (1 - \alpha(\theta))\delta_\theta + \alpha(\theta)\tilde{g}(z(\theta)), \\ \pi_n &= (1 - \alpha(\theta))\delta_{\theta_n} + \alpha(\theta)\tilde{g}(z(\theta)), \end{aligned}$$

for each n. Note that $\pi = \tilde{\pi}(\theta)$. Since θ is in S, we know that π is in Π. The sequence (π_n) converges to π, but for all n, we have $\langle v, \pi_n \rangle = \alpha(\theta) < 1$.

---

in $B(y_0, \delta') \cap \operatorname{aff} g(S)$. Thus, there exists δ in $(0, \delta')$ such that

$$D \coloneqq B(y_0, \delta) \cap \operatorname{aff} g(S) \subset \operatorname{conv}(z_1, \ldots, z_k) \subset B(y_0, \delta') \cap \operatorname{aff} g(S).$$

Each vector z_j is in $\operatorname{conv} g(S)$, so by Carathéodory's theorem, we may select a probability measure ζ_j in $\Delta(S)$ concentrating on at most $m+1$ points such that $\langle g, \zeta_j \rangle = z_j$. There exists a continuous coordinate mapping $\hat{p}$ from $\operatorname{conv}(z_1, \ldots, z_k)$ to the probability simplex in $\mathbf{R}^k$ such that $z = \sum_{j=1}^k \hat{p}_j(z) z_j$, for each z in $\operatorname{conv}(z_1, \ldots, z_k)$. For each z in D, let $\tilde{g}(z) = \sum_{j=1}^k \hat{p}_j(z)\zeta_j$. By linearity, $\langle g, \tilde{g}(z) \rangle = \sum_{j=1}^k \hat{p}_j(z) z_j = z$. The function $\tilde{g}\colon D \to \Delta(S)$ is continuous because the coordinate map $\hat{p}$ is continuous.

[39] Here, we adopt the convention that $0/\|0\| = 0$. If $g(\theta) = y_0$, then $z(\theta) = y_0$.

Second, suppose that S is not closed. Choose a sequence (θ_n) in S that converges to some point θ in $\Theta \backslash S$. For each n, let $\pi_n = \tilde{\pi}(\theta_n)$ and let $\pi = \pi(\theta)$. Since (θ_n) is in S, the sequence (π_n) is in Π. Since $\tilde{\pi}$ is continuous, the sequence (π_n) converges to π, but $\langle v, \pi \rangle \leq \alpha(\theta) < 1$.

Singletons and relative entropy and total variation balls Suppose that Θ is perfect. Fix π_0 in $\Delta(\Theta)$. We consider ambiguity sets of the following three forms: (a) $\Pi = \{\pi_0\}$; (b) $\Pi = B_R(\pi_0, \beta)$ for some $\beta > 0$; and (c) $\Pi = B_{\mathrm{TV}}(\pi_0, \gamma)$ for some $\gamma \in (0, 1)$. By Lemma 6, there exists a set A with $\pi_0(A) = 1$ and a sequence (π_n) in $\Delta(\Theta \setminus A)$ that converges to π_0. Let $v = 1_A$. For each n, we have $\langle v, \pi_n \rangle = 0$. On the other hand, $W_\Pi(v) = 1$ in cases (a) and (b). In case (c), $W_\Pi(v) \geq 1 - \gamma$ because for each $\pi' \in \Pi$,

$$\langle v, \pi' \rangle = \pi'(A) \geq \pi_0(A) - \|\pi' - \pi_0\|_{\mathrm{TV}} \geq 1 - \gamma.$$

Marginal sets Let $\Theta = \prod_{j=1}^m \Theta_j$, where $\Theta_1, \ldots, \Theta_m$ are perfect. Let $\Pi = \Gamma((\pi_j)_{j \in J})$ for some nonempty subset J of $\{1, \ldots, m\}$ and some priors π_j in $\Delta(\Theta_j)$ for each j in J. Without loss, suppose 1 is in J. Apply Lemma 6 at the prior π_1 in $\Delta(\Theta_1)$ to obtain a subset A_1 of Θ_1 with $\pi_1(A_1) = 1$ and a sequence (π_1^n) in $\Delta(\Theta \setminus A_1)$ that converges to π_1. For $j \notin J$, arbitrarily choose $\pi_j \in \Delta(\Theta_j)$. For each n, let $\pi^n = \pi_1^n \otimes \pi_{-1}$, where $\pi_{-1} = (\pi_j)_{j \neq 1}$. The sequence (π^n) converges to $\pi_1 \otimes \pi_{-1} \in \Pi$. Let $v = 1_{A_1 \times \Theta_{-1}}$. Thus, $W_\Pi(v) = 1$, but for each n, we have $\langle v, \pi^n \rangle = 0$.

Quantile sets Let Θ be a convex subset of $\mathbf{R}$. Let $\Pi = Q((x_j, \alpha_j)_{j=1}^m)$, for some positive integer m and some $x_1, \ldots, x_m \in \mathbf{R}$ and $\alpha_1, \ldots, \alpha_m \in [0, 1]$ satisfying $\inf \Theta < x_1 < \cdots < x_m < \sup \Theta$ and $\alpha_1 < \cdots < \alpha_m$. If $\alpha_m = 0$, then $m = 1$ and Π is a support set, which is a special case of a probability set. Therefore, we may assume $\alpha_m > 0$. Set $\alpha_0 = 0$. Let

$$\pi = (1 - \alpha_{m-1})\delta_{x_m} + \sum_{j=1}^{k-1} (\alpha_j - \alpha_{j-1})\delta_{x_j}.$$

By construction, π is in Π. Set $v = 1_{(-\infty, x_m]}$. We have $W_\Pi(v) = \alpha_m$. Choose a strictly decreasing sequence (θ_n) in Θ that converges to x_m. For each n, let $\pi_n = \pi + (1 - \alpha_{m-1})(\delta_{\theta_n} - \delta_{x_m})$. The sequence (π_n) converges to π, but $\langle v, \pi_n \rangle = \alpha_{m-1} < \alpha_m$ for all n.

A.5 Proof of Theorem 3

Let $u \colon X \times \Theta$ be state-continuous. Let Π be a nonempty closed subset of $\Delta(\Theta)$. Let $\succsim = \succsim_{(u,\Pi)}$. Fix f in $\mathcal{F}_0(u)$.

First, suppose that the payoff guarantee from v_f over Π is robust. Let g in $\mathcal{F}_0$ be a graphical limit of f. We prove that $W_\Pi(v_g) \geq W_\Pi(v_f)$. For each θ in Θ, there exists a sequence (θ_n) converging to θ such that $g(\theta) = f(\theta_n)$ for all n. Since u is state-continuous, it follows that $v_g(\theta) = \lim_n v_f(\theta_n)$, so v_g is bounded and we have $v_g \geq \operatorname{lsc} v_f$. Therefore, it suffices to show that $W_\Pi(\operatorname{lsc} v_f) \geq W_\Pi(v_f)$. Fix π in Π. By Lemma 2, we have $\Gamma\text{-}\lim_n \langle v, \cdot \rangle = \langle \operatorname{lsc} v, \cdot \rangle$. Thus, there exists a sequence (π_n) in $\Delta(\Theta)$ converging to π such that

$$\langle \operatorname{lsc} v_f, \pi \rangle \geq \limsup_n \langle v_f, \pi_n \rangle \geq \liminf_n \langle v_f, \pi_n \rangle \geq W_\Pi(v_f), \tag{8}$$

where the last inequality holds because v_f is robust over Π. Since π is an arbitrary prior in Π, we conclude that $W_\Pi(\operatorname{lsc} v_f) \geq W_\Pi(v_f)$, as desired.

For the converse, suppose that the payoff guarantee from v_f over Π is not robust. Thus, there exists a sequence (π_n) in $\Delta(\Theta)$ converging to a prior in Π (which equals the closure of Π) such that

$$W_\Pi(v_f) > \liminf_n \langle v_f, \pi_n \rangle \geq \langle \operatorname{lsc} v_f, \pi \rangle \geq W_\Pi(\operatorname{lsc} v_f), \tag{9}$$

where the second inequality holds because $\Gamma\text{-}\lim_n \langle v_f, \cdot \rangle = \langle \operatorname{lsc} v_f, \cdot \rangle$ by Lemma 2. To complete the proof, we apply Lemma 3 to construct an act g in $\mathcal{F}_0(u)$ such that g is a graphical limit of f and $v_g = \operatorname{lsc} v_f$. Thus, $f \succ g$.

A.6 Proof of Theorem 4

Let u be state-independent and nonconstant. We write $u(x)$ for $u(x, \theta)$. Thus, the value function induced by a social choice function f is the composition $u \circ f$. Let Π be a nonempty subset of $\Delta(\Theta)$. Let $\succsim = \succsim_{(u,\Pi)}$. We separate the proof into parts. First, we prove that Π is tight if and only if $\succsim$ is tight.

Tightness Suppose that Π is tight. We show that $\succsim$ is tight. Fix $f, g \in \mathcal{F}_0$ with $W_\Pi(u \circ f) > W_\Pi(u \circ g)$. Fix $x \in \Delta_0(X)$. Let

$$\varepsilon = \frac{W_\Pi(u \circ f) - W_\Pi(u \circ g)}{\max\{1, 2\|u \circ f - u \circ x\|_\infty, 2\|u \circ g - u \circ x\|_\infty\}}.$$

Since Π is tight, we may select a compact subset K of Θ such that $\pi(K) \geq 1-\varepsilon$ for every π in Π. Since $u \circ (fKx) = (u \circ f)K + (u \circ x)K^c$, we have

$$\begin{aligned} W_\Pi(u \circ (fKx)) &\geq W_\Pi(u \circ f) - \sup\nolimits_{\pi \in \Pi} \pi(K^c) \|u \circ f - u \circ x\|_\infty \\ &\geq W_\Pi(u \circ f) - \varepsilon \|u \circ f - u \circ x\|_\infty \\ &> W_\Pi(u \circ g). \end{aligned}$$

A symmetric argument shows that $W_\Pi(u \circ f) > W_\Pi(u \circ (gKx))$.

Next, suppose that $\succsim$ is tight. We show that Π is tight. Fix $\varepsilon > 0$. Since u is nonconstant, we can choose $x, y \in X$ such that $y \succ x$. So $u(y) > u(x)$. Let $g = \varepsilon x + (1-\varepsilon)y$. By construction, $y \succ g$. Since $\succsim$ is tight, there exists a compact subset K of Θ such that $yKx \succ g$. Since $u \circ (yKx) = (u \circ y)K + (u \circ x)K^c$, we have

$$\begin{aligned} u(y) + (u(x) - u(y)) \sup_{\pi \in \Pi} \pi(K^c) &= W_\Pi(u \circ (yKx)) \\ &> W_\Pi(g) \\ &= u(y) + \varepsilon(u(x) - u(y)). \end{aligned}$$

Hence, $\sup_{\pi \in \Pi} \pi(K^c) < \varepsilon$. We conclude that Π is tight.

Forward implication Suppose that Π is globally robust and tight. We may assume, without loss, that Π is closed.[40] It follows from Theorem 3 that $\succsim$ is weakly upper Γ-semicontinuous. We show that $\succsim$ is lower Γ-semicontinuous. Fix acts $g, h \in \mathcal{F}_0$ and a bounded sequence (f_n) in $\mathcal{F}_0$. Suppose that $h \succsim f_n$ for all n and that g is a Γ-limit of (f_n). To check that $h \succsim g$, it suffices to prove that $W_\Pi(u \circ g) \leq \liminf_n W_\Pi(u \circ f_n)$.

Let $v_g = u \circ g$, and for each n, let $v_n = u \circ f_n$. By Lemma 4.ii, $v_g = \Gamma\text{-}\lim v_n$. By Lemma 2, $\langle v_g, \cdot \rangle = \Gamma\text{-}\lim \langle v_n, \cdot \rangle$. Since Π is tight and closed, Π is compact by Prokhorov's theorem (Billingsley, 1999, Theorem 5.1, p. 59). By Braides (2002, Proposition 1.18, p. 28), $W_\Pi(v_g) \leq \liminf_n W_\Pi(v_n)$, as desired.

Backward implication Suppose that $\succsim$ is weakly upper Γ-semicontinuous and lower Γ-semicontinuous. We prove that Π is globally robust. Let $Z = u(\Delta_0(X))$, which is a convex subset of $\mathbf{R}$ with nonempty interior (since u is nonconstant). Let $B_0(\Theta, Z)$ denote the space of measurable simple functions from Θ to Z. The space $B_0(\Theta, Z)$ is uniformly dense in the space $B(\Theta, Z)$ of bounded, measurable functions from Θ to Z. Moreover, $W_\Pi(\alpha v + \beta) = \alpha W_\Pi(v) + \beta$ for any real $\alpha > 0$ and any real β. Therefore, it suffices to prove that for each value function v in $B_0(\Theta, Z)$, the payoff guarantee from v over Π is robust.

Fix v in $B_0(\Theta, Z)$. Let (π_n) be a sequence in $\Delta(\Theta)$ that converges to a prior π in the closure of Π. We prove that $\liminf_n \langle v, \pi_n \rangle \geq W_\Pi(v)$. By Lemma 2, $\Gamma\text{-}\lim_n \langle v, \cdot \rangle = \langle \operatorname{lsc} v, \cdot \rangle$, so $\liminf_n \langle v, \pi_n \rangle \geq \langle \operatorname{lsc} v, \pi \rangle$. Choose f in $\mathcal{F}_0$ such that $u \circ f = v$. Using Lemma 3, it can be shown that f has a Γ-limit g in $\mathcal{F}_0$ with $u \circ g = \operatorname{lsc} v$. Since $\succsim$ is upper Γ-semicontinuous, we have $g \succsim f$, hence $W_\Pi(\operatorname{lsc} v) \geq W_\Pi(v)$.

Therefore, it suffices to prove that $\langle \operatorname{lsc} v, \pi \rangle \geq W_\Pi(\operatorname{lsc} v)$. To simplify notation, let $w = \operatorname{lsc} v$. Recall that $w = u \circ g$ for the act g defined above. Write $g = \sum_{j=1}^m x_j A_j$ for some $x_1, \ldots, x_m \in \Delta_0(X)$ and some measurable partition $(A_1, \ldots, A_m)$ of Θ. Since π is in the closure of Π, there exists a

[40] Otherwise, replace Π with its closure $\bar{\Pi}$, which is also robust. Since Π is robust, $W_\Pi = W_{\bar{\Pi}}$, hence $\succsim_{(u,\Pi)} = \succsim_{(u,\bar{\Pi})}$.

sequence (π'_n) in Π that converges to π. Choose a bounded, compatible metric d on Θ. By Lemma 1, there exists a sequence (κ_n) of probability kernels such that $\|\kappa_n\|_d \to 0$ and $\|\pi - \pi'_n \kappa_n\|_{\mathrm{TV}} \to 0$.

Fix $\varepsilon > 0$. We claim (see proof below) that there exists a sequence $(\bar{\kappa}_n)$ of probability kernels such that for each state θ and each n, we have

$$\|(\kappa_n)_\theta - (\bar{\kappa}_n)_\theta\|_{\mathrm{TV}} \leq \varepsilon, \tag{10}$$

and the set $\mathcal{K} \coloneqq \cup_n \{(\bar{\kappa}_n(\theta, A_j))_{j=1}^m : \theta \in \Theta\}$ is finite. For each n, let $g_n = \bar{\kappa}_n g \in \mathcal{F}_0$. Thus, $u \circ g_n = \bar{\kappa}_n w$, so $\langle u \circ g_n, \pi'_n \rangle = \langle w, \pi'_n \bar{\kappa}_n \rangle$. Therefore, for each n,

$$\begin{aligned} \langle w, \pi \rangle &= \langle u \circ g_n, \pi'_n \rangle + \langle w, \pi - \pi'_n \bar{\kappa}_n \rangle \\ &\geq W_\Pi(u \circ g_n) - 2\|w\|_\infty \|\pi - \pi'_n \bar{\kappa}_n\|_{\mathrm{TV}}. \end{aligned}$$

Take the limit supremum as $n \to \infty$. Since $\|\pi - \pi'_n \kappa_n\|_{\mathrm{TV}} \to 0$, it follows from (10) that

$$\langle w, \pi \rangle \geq \limsup_n W_\Pi(u \circ g_n) - 2\varepsilon \|w\|_\infty. \tag{11}$$

Since $\mathcal{K}$ is finite, the set $\cup_n \{g_n(\theta) : \theta \in \Theta\}$ is finite. We may assume without loss that $\cup_n \{g_n(\theta) : \theta \in \Theta\}$ is indifference-free.[41] By Braides (2002, Proposition 1.42, p. 35), the sequence $(u \circ g_n)$ has a Γ-convergent subsequence $(u \circ g_{n_k})$. By Lemma 4.i, the sequence (g_{n_k}) has a Γ-limit in $\mathcal{F}_0$, which we denote by $\bar{g}$. By the properties of the limit supremum and by the lower Γ-semicontinuity of $\succsim$, we have

$$\limsup_n W_\Pi(u \circ g_n) \geq \limsup_k W_\Pi(u \circ g_{n_k}) \geq W_\Pi(u \circ \bar{g}). \tag{12}$$

By Lemma 4.ii, $u \circ \bar{g} = \Gamma\text{-}\lim_k \bar{\kappa}_{n_k} w$. For each θ in Θ, there exists a sequence

[41] Otherwise, select a representative of each indifference class of $\cup_n \{g_n(\theta) : \theta \in \Theta\}$. For each θ and n, replace $g_n(\theta)$ with the representative of its indifference class. This procedure does not change the induced value functions.

(θ_k) converging to θ such that

$$\begin{aligned}(u \circ \bar{g})(\theta) &\geq \limsup_k (\bar{\kappa}_{n_k} w)(\theta_k) \\ &\geq \limsup_k (\kappa_{n_k} w)(\theta_k) - 2\varepsilon \|w\|_{\mathrm{TV}} \\ &\geq w(\theta) - 2\varepsilon \|w\|_{\mathrm{TV}},\end{aligned} \tag{13}$$

where the second inequality follows from (10) and the last inequality holds because $\|\kappa_{n_k}\|_d \to 0$ and w is lower semicontinuous.[42] Since ε was arbitrary, piecing together (11), (12), and (13) gives $\langle w, \pi \rangle \geq W_\Pi(w)$, as desired.

Proof of claim Fix $\varepsilon > 0$. Let Δ denote the probability simplex in $\mathbf{R}^m$. For each p_0 in $\operatorname{int} \Delta$, let

$$U(p_0) = \{p \in \Delta : p_0 + \varepsilon^{-1}(1-\varepsilon)(p_0 - p) \in \operatorname{int} \Delta\}.$$

The sets $U(p_0)$, for p_0 in $\operatorname{int} \Delta$, form an open cover of Δ. Since Δ is compact, there exists a finite subset Δ_0 of $\operatorname{int} \Delta$ such that $\cup_{p_0 \in \Delta_0} U(p_0) = \Delta$. Therefore, for each p in Δ, there exists some p_0 in Δ_0 such that $p_0 + \varepsilon^{-1}(1-\varepsilon)(p_0 - p)$ is in $\operatorname{int} \Delta$. Let $q = p_0 + \varepsilon^{-1}(1-\varepsilon)(p_0 - p)$. Some algebra shows that $p_0 = (1-\varepsilon)p + \varepsilon q$. Therefore, we can choose a measurable map $\hat{q} \colon \Delta \to \Delta$ such that for every p in Δ, we have $(1-\varepsilon)p + \varepsilon \hat{q}(p) \in \Delta_0$.

We now define the sequence $(\bar{\kappa}_n)$. For each $j = 1, \ldots, m$, choose $\theta_j \in A_j$. For each θ and n, let $p_{\theta,n} = (\kappa_n(\theta, A_j))_{j=1}^m \in \Delta$, and let

$$(\bar{\kappa}_n)_\theta = (1-\varepsilon)(\kappa_n)_\theta + \varepsilon \sum_{j=1}^m \hat{q}_j(p_{\theta,n}) \delta_{\theta_j}.$$

[42] For each n, the measure $\kappa_n(\theta_n, \cdot)$ concentrates on $B(\theta_n, \|\kappa_n\|_d)$, so there exists $\tilde{\theta}_n$ in $B(\theta_n, \|\kappa_n\|_d)$ such that $(\kappa_n w)(\theta_n) \geq w(\tilde{\theta}_n)$. Since $\theta_n \to \theta$ and $\|\kappa_n\|_d \to 0$, we have $\tilde{\theta}_n \to \theta$. Therefore, since w is lower semicontinuous,

$$\limsup_n (\kappa_{n_k} w)(\theta_{n_k}) \geq \limsup_k w(\tilde{\theta}_{n_k}) \geq \liminf_k w(\tilde{\theta}_{n_k}) \geq w(\theta).$$

By construction, $\|(\kappa_n)_\theta - (\bar{\kappa}_n)_\theta\|_{\mathrm{TV}} \leq \varepsilon$. For each θ and n, we have

$$(\bar{\kappa}_n(\theta, A_j))_{j=1}^m = (1-\varepsilon)p_{\theta,n} + \varepsilon\hat{q}(p_{\theta,n}) \in \Delta_0.$$

Thus, $\mathcal{K} \coloneqq \cup_n\{(\bar{\kappa}_n(\theta, A_j))_{j=1}^m : \theta \in \Theta\} \subset \Delta_0$, so $\mathcal{K}$ is finite.

References

ARMSTRONG, M. AND J. VICKERS (2010): "A Model of Delegated Project Choice," *Econometrica*, 78, 213–244. [10]

ARROW, K. J. (1970): *Essays in the Theory of Risk-Bearing*, North-Holland. [25, 52]

AUSTER, S. (2018): "Robust Contracting under Common Value Uncertainty," *Theoretical Economics*, 13, 175–204. [14]

AUSTER, S., Y.-K. CHE, AND K. MIERENDORFF (2024): "Prolonged Learning and Hasty Stopping: The Wald Problem with Ambiguity," *American Economic Review*, 114, 426–461. [27]

AZAR, P. AND S. MICALI (2012): "Optimal Parametric Auctions," Computer Science and Artificial Intelligence Laboratory Technical Report MIT-CSAIL-TR-2012-011. [14]

BACHRACH, N., Y.-C. CHEN, I. TALGAM-COHEN, X. YANG, AND W. ZHANG (2022): "Distributionally Robust Auction Design," Working paper. [14, 18]

BERGEMANN, D., S. MORRIS, ET AL. (2013): "An Introduction to Robust Mechanism Design," *Foundations and Trends in Microeconomics*, 8, 169–230. [26]

BERGEMANN, D. AND K. SCHLAG (2011): "Robust Monopoly Pricing," *Journal of Economic Theory*, 146, 2527–2543. [26]

BERGEMANN, D. AND K. H. SCHLAG (2008): "Pricing without Priors," *Journal of the European Economic Association*, 6, 560–569. [8, 17, 26]

BILLINGSLEY, P. (1999): *Convergence of Probability Measures*, John Wiley & Sons, 2 ed. [43, 62]

BOGACHEV, V. I. (2018): *Weak Convergence of Measures*, vol. 234 of *Mathematical Surveys and Monographs*, American Mathematical Society. [36]

Braides, A. (2002): *Gamma-Convergence for Beginners*, Oxford University Press. [22, 31, 43, 44, 56]

Brooks, B. and S. Du (2021a): "Maxmin Auction Design with Known Expected Values," Technical report, University of Chicago and University of California–San Diego. [26, 27]

——— (2021b): "Optimal Auction Design with Common Values: An Informationally Robust Approach," *Econometrica*, 89, 1313–1360. [27]

——— (2023): "Robust Mechanisms for the Financing of Public Goods," *Available at SSRN 4482541*. [27]

Carrasco, V., V. Farinha Luz, N. Kos, M. Messner, P. Monteiro, and H. Moreira (2018): "Optimal Selling Mechanisms under Moment Conditions," *Journal of Economic Theory*, 177, 245–279. [14]

Carrasco, V., V. Farinha Luz, P. Monteiro, and H. Moreira (2019): "Robust Mechanisms: The Curvature Case," *Economic Theory*, 68, 203–222. [14]

Carroll, G. (2015): "Robustness and Linear Contracts," *American Economic Review*, 105, 536–63. [27]

——— (2017): "Robustness and Separation in Multidimensional Screening," *Econometrica*, 85, 453–488. [8, 18]

——— (2019): "Robustness in Mechanism Design and Contracting," *Annual Review of Economics*, 11, 139–166. [26]

Cerreia-Vioglio, S., L. P. Hansen, F. Maccheroni, and M. Marinacci (2024): "Making Decisions under Model Misspecification," Working paper. [28]

Chateauneuf, A., F. Maccheroni, M. Marinacci, and J.-M. Tallon (2005): "Monotone Continuous Multiple Priors," *Economic Theory*, 26, 973–982. [52]

Che, Y.-K. and W. Zhong (2021): "Robustly-Optimal Mechanism for Selling Multiple Goods," in *Proceedings of the 22nd ACM Conference on Economics and Computation*, 314–315. [18]

Chen, Y.-C., T. Kunimoto, and Y. Sun (2023): "Continuous Implementation with Payoff Knowledge," *Journal of Economic Theory*, 209, 105624. [26]

Chen, Y.-C., M. Mueller-Frank, and M. M. Pai (2022): "Continuous Implementation with Direct Revelation Mechanisms," *Journal of Economic Theory*, 201, 105422. [26]

Gilboa, I. and D. Schmeidler (1989): "Maxmin Expected Utility with Non-Unique Prior," *Journal of Mathematical Economics*, 18, 141–153. [5, 52]

Guo, Y. and E. Shmaya (2023): "Regret-Minimizing Project Choice," *Econometrica*, 91, 1567–1593. [10, 17]

Handel, B. R. and K. Misra (2015): "Robust New Product Pricing," *Marketing Science*, 34, 864–881. [26]

Hansen, L. P. and T. J. Sargent (2001): "Robust Control and Model Uncertainty," *American Economic Review: Papers and Proceedings*, 91, 60–66. [18, 25]

——— (2022): "Structured Ambiguity and Model Misspecification," *Journal of Economic Theory*, 199, 105165. [28]

Hill, B. (2019): "A Non-Bayesian Theory of State-Dependent Utility," *Econometrica*, 87, 1341–1366. [22]

Hinnosaar, T. and K. Kawai (2020): "Robust Pricing with Refunds," *RAND Journal of Economics*, 51, 1014–1036. [27]

Hu, J. and X. Weng (2021): "Robust Persuasion of a Privately Informed Receiver," *Economic Theory*, 72, 909–953. [9, 10]

Ilut, C., R. Valchev, and N. Vincent (2020): "Paralyzed by Fear: Rigid and Discrete Pricing under Demand Uncertainty," *Econometrica*, 88, 1899–1938. [27]

Jehiel, P., M. Meyer-ter Vehn, and B. Moldovanu (2012): "Locally Robust Implementation and its Limits," *Journal of Economic Theory*, 147, 2439–2452. [26]

Kallenberg, O. (2021): *Foundations of Modern Probability*, Springer, 3 ed. [57]

Kamenica, E. and M. Gentzkow (2011): "Bayesian Persuasion," *American Economic Review*, 101, 2590–2615. [9]

Kosterina, S. (2022): "Persuasion with Unknown Beliefs," *Theoretical Economics*, 17, 1075–1107. [9]

Lanzani, G. (2024): "Dynamic Concern for Misspecification," Available at SSRN 4454504. [28]

Libgober, J. and X. Mu (2021): "Informational Robustness in Intertemporal Pricing," *The Review of Economic Studies*, 88, 1224–1252. [27]

Maccheroni, F., M. Marinacci, and A. Rustichini (2006): "Ambiguity Aversion, Robustness, and the Variational Representation of Preferences," *Econometrica*, 74, 1447–1498. [25]

Madarász, K. and A. Prat (2017): "Sellers with Misspecified Models," *Review of Economic Studies*, 84, 790–815. [27]

Meyer-ter Vehn, M. and S. Morris (2011): "The Robustness of Robust Implementation," *Journal of Economic Theory*, 146, 2093–2104. [26]

Oury, M. (2015): "Continuous Implementation with Local Payoff Uncertainty," *Journal of Economic Theory*, 159, 656–677. [26]

Oury, M. and O. Tercieux (2012): "Continuous Implementation," *Econometrica*, 80, 1605–1637. [26]

Pei, H. and B. Strulovici (2024): "Robust Implementation with Costly Information," *Review of Economic Studies*, rdae016. [28]

Pinar, M. C. and C. Kizilkale (2017): "Robust Screening under Ambiguity," *Mathematical Programming: Series A*, 163, 273–299. [15]

Prasad, K. (2003): "Non-Robustness of Some Economic Models," *Topics in Theoretical Economics*, 3. [25]

Rahimian, H. and S. Mehrotra (2022): "Frameworks and Results in Distributionally Robust Optimization," *Open Journal of Mathematical Optimization*, 3. [26]

Scarf, H. (1958): "A Min-max Solution of an Inventory Problem," in *Studies in the Mathematical Theory of Inventory and Production*, ed. by K. Arrow, S. Karlin, and H. Scarf, Stanford University Press, chap. 12, 201–209. [14]

Stanca, L. (2023): "Robust Bayesian Choice," *Mathematical Social Sciences*, 126, 94–106. [25, 26]

Strzalecki, T. (2011): "Axiomatic Foundations of Multiplier Preferences," *Econometrica*, 79, 47–73. [25]

Villani, C. (2009): *Optimal Transport: Old and New*, Springer-Verlag. [57, 64]

B Online Appendix

B.1 Incompatibility with Arrow's monotone continuity

In this subsection, we show that weak upper Γ-continuity is incompatible with Arrow's (1970) monotone continuity axiom. Under subjective expected utility, the prior is countably additive if and only if the induced preferences satisfy the following monotone continuity axiom. Recall that $\mathcal{B}(\Theta)$ denotes the collection of Borel subsets of Θ.

Axiom 4 (Monotone continuity)**.** For each $f, g \in \mathcal{F}_0$, each $x \in \Delta_0(X)$, and each sequence (E_n) in $\mathcal{B}(\Theta)$ with $E_n \downarrow \varnothing$,[43] if $f \succ g$, then there exists m such that $xE_m f \succ g$ and $f \succ xE_m g$.

Following the mechanism design literature, our framework considers only countably additive priors. For maxmin preferences, however, the behavioral foundation for countably additive priors is more subtle. Chateauneuf et al. (2005) show that a preference relation satisfies Arrow's (1970) monotone continuity axiom together with the axioms of Gilboa and Schmeidler (1989) if and only if it has a maxmin representation in which the ambiguity set is a collection of countably additive priors that satisfies a compactness property.

We next show that weak upper Γ-semicontinuity is generally inconsistent with monotone continuity. A relation $\succsim$ on $\mathcal{F}_0$ is *nontrivial* if there exist $f, g \in \mathcal{F}_0$ such that $f \succ g$.

Theorem 5 (Incompatible continuity)
Suppose that Θ is perfect. A nontrivial, monotone, complete, transitive relation on $\mathcal{F}_0$ cannot be both weakly upper Γ-semicontinuous and monotone continuous.

Proof. Suppose for a contradiction that such a relation $\succsim$ exists. Then there exist $x, y \in \Delta_0(X)$ such that $x \succ y$; otherwise, monotonicity implies that $\succsim$ is trivial. Since Θ is separable, we can enumerate a countable dense subset $\{\theta_1, \theta_2, \ldots\}$ of Θ. For each n, let $E_n = \{\theta_j : j \geq n\}$. By construction, $E_n \downarrow \varnothing$.

[43] That is, $E_1 \supset E_2 \supset \cdots$ and $\cap_n E_n = \varnothing$.

By monotone continuity, there exists m such that $yE_m x \succ y$. Since Θ is perfect, E_m is a dense subset of Θ, and hence y is a Γ-limit of $yE_m x$. By weak upper Γ-semicontinuity, $y \succsim yE_m x$, giving the contradiction $y \succ y$. $\square$

These two continuity axioms have different implications for an act that yields a bad consequence on a countable, dense subset of the state space. Weak upper Γ-semicontinuity, a topological robustness property, demands that such an act be evaluated more negatively because an arbitrarily small perturbation of any state yields a bad consequence. Monotone continuity, a measure-theoretic continuity property, is less conservative because the set of states yielding the bad consequences has small cardinality relative to the full state space.[44] Theorem 5 is a manifestation of the inconsistency between topological and measure-theoretic notions of smallness.

B.2 Axiomatization of uniform robustness

According to Definition 1, an ambiguity set Π is globally robust if for every sequence (π_n) in $\Delta(\Theta)$ that converges to a prior in the closure of Π, we have

$$\liminf_n \left[\langle v, \pi_n \rangle - W_\Pi(v)\right] \geq 0, \qquad v \in B(\Theta).$$

An ambiguity set Π is *uniformly robust* if for every sequence (π_n) in $\Delta(\Theta)$ that converges to a prior in the closure of Π, we have

$$\liminf_n \left(\inf_v \left[\langle v, \pi_n \rangle - W_\Pi(v)\right] \right) \geq 0,$$

where the infimum inside the parentheses is taken over all value functions v in $B(\Theta)$ satisfying $\|v\|_\infty \leq 1$.

We next define upper Γ-semicontinuity analogously to lower Γ-semicontinuity.

Axiom 5 (Upper Γ-semicontinuity)**.** For any bounded sequence (f_n) in $\mathcal{F}_0$ and any $g, h \in \mathcal{F}_0$, if $f_n \succsim h$, for all n, and g is a Γ-limit of (f_n), then $g \succsim h$.

[44] A perfect Polish space must be uncountable.

A relation $\succsim$ on $\mathcal{F}_0$ is Γ-*continuous* if $\succsim$ is upper Γ-semicontinuous and lower Γ-semicontinuous.

Theorem 6 (Uniform robustness)
Let $\succsim = \succsim_{(u,\Pi)}$ *for some nonconstant state-independent utility function* u *and some nonempty subset* Π *of* $\Delta(\Theta)$*. The following are equivalent:*

1. Π *is uniformly robust and tight;*

2. $\succsim$ *is* Γ*-continuous and tight.*

B.3 Continuous moment sets

Here, we give some simple examples of continuous moment sets. Let $\Theta = \mathbf{R}$. Let Π_1 be the set of all priors with mean in the interval $[\underline{\mu}, \bar{\mu}]$ and variance exactly σ^2. Let Π_2 be the set of all priors with mean in the interval $[\underline{\mu}, \bar{\mu}]$ and variance at most $\bar{\sigma}^2$. For each $j = 1, 2$, we have $\Pi_j = M(g, Y_j)$, where $g(\theta) = (\theta, \theta^2)$, and

$$Y_1 = \left\{(y_1, y_2) : \underline{\mu} \leq y_1 \leq \bar{\mu} \text{ and } y_2 = y_1^2 + \sigma^2\right\},$$
$$Y_2 = \left\{(y_1, y_2) : \underline{\mu} \leq y_1 \leq \bar{\mu} \text{ and } 0 \leq y_2 \leq y_1^2 + \bar{\sigma}^2\right\}.$$

In an example with $\sigma^2 > \bar{\sigma}^2$, Figure 4 plots the image $g(\Theta)$ (in blue), the set Y_1 (in purple), and the intersection $Y_2 \cap \operatorname{conv} g(\Theta)$ (shaded orange). Both Π_1 and Π_2 are continuous moment sets: Y_1 is uniformly g-interior, but not star g-interior; Y_2 is star g-interior (relative to the point y_0), but not uniformly g-interior.

B.4 Proof of Theorem 6

We build upon the proof of Theorem 4 (Appendix A.6). Let u be state-independent and nonconstant. Let Π be a nonempty subset of $\Delta(\Theta)$. Let $\succsim = \succsim_{(u,\Pi)}$.

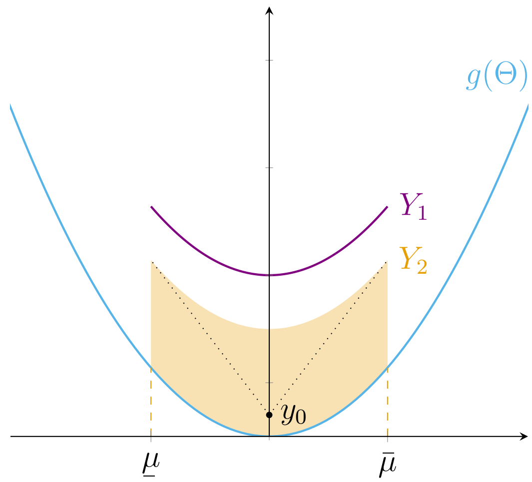

Figure 4. Continuous moment sets

Forward implication Suppose that Π is uniformly robust and tight. We may assume, without loss, that Π is closed.[45] By Theorem 4, it suffices to prove that $\succsim$ is upper Γ-semicontinuous. Fix acts $g, h \in \mathcal{F}_0$ and a bounded sequence (f_n) in $\mathcal{F}_0$. Suppose that $f_n \succsim h$ for all n and that g is a Γ-limit of (f_n). To check that $g \succsim h$, it suffices to prove that $W_\Pi(u \circ g) \geq \liminf_n W_\Pi(u \circ f_n)$.

Let $v_g = u \circ g$. For each n, let $v_n = u \circ f_n$. Since (f_n) is bounded, the sequence (v_n) is bounded. By Lemma 4.ii, $v_g = \Gamma\text{-}\lim_n v_n$. By Lemma 2, $\langle v_g, \cdot \rangle = \Gamma\text{-}\lim_n \langle v_n, \cdot \rangle$. Fix π in Π. There exists a sequence (π_n) in $\Delta(\Theta)$ converging to π such that

$$\langle v_g, \pi \rangle \geq \limsup_n \langle v_n, \pi_n \rangle. \tag{14}$$

For each n, we have

$$\begin{aligned} \langle v_n, \pi_n \rangle &= W_\Pi(v_n) + [\langle v_n, \pi_n \rangle - W_\Pi(v_n)] \\ &\geq W_\Pi(v_n) + \inf_m [\langle v_m, \pi_n \rangle - W_\Pi(v_m)] . \end{aligned}$$

Since (v_n) is bounded, the uniform robustness of Π implies that the limit infi-

[45] Otherwise, replace Π with its closure $\bar{\Pi}$, which is also uniformly robust. Since Π is robust, $W_\Pi = W_{\bar{\Pi}}$, hence $\succsim_{(u,\Pi)} = \succsim_{(u,\bar{\Pi})}$.

mum of the second term is nonnegative. The limit infimum is supperadditive, so we conclude that

$$\liminf_n \langle v_n, \pi_n \rangle \geq \liminf_n W_\Pi(v_n). \tag{15}$$

Combining (14) and (15) shows that $\langle v_g, \pi \rangle \geq \liminf_n W_\Pi(v_n)$. Take the infimum over all π in Π to get $W_\Pi(v_g) \geq \liminf_n W_\Pi(v_n)$, as desired.

Backward implication Suppose that $\succsim$ is Γ-continuous. From the proof of the backward implication in Theorem 4 (Appendix A.6), we know that Π is robust. Suppose for a contradiction that Π is not uniformly robust. Then for some sequence (π_n) converging to a prior π in the closure of Π, there exists $\varepsilon > 0$ and a sequence (v_n) in $B(\Theta)$ with $\sup_n \|v_n\|_\infty \leq 1$ such that for every n,

$$\langle v_n, \pi_n \rangle \leq W_\Pi(v_n) - \varepsilon.$$

After adjusting the sequence (v_n) and the value of ε, we may assume that the sequence (v_n) is in $B_0(\Theta, Z_0)$ for some finite subset Z_0 of $Z = u(\Delta_0(X))$.[46]

By Braides (2002, Proposition 1.42, p. 35), the sequence (v_n) has a Γ-convergent subsequence. After passing to this subsequence, we may assume that (v_n) Γ-converges to some value function v, which must be in $B_0(\Theta, Z_0)$. By Lemma 2, $\Gamma\text{-}\lim_n \langle v_n, \cdot \rangle = \langle v, \cdot \rangle$. Therefore,

$$\begin{aligned} \liminf_n W_\Pi(v_n) - \varepsilon &\geq \liminf_n \langle v_n, \pi_n \rangle \\ &\geq \langle v, \pi \rangle \\ &\geq W_\Pi(v), \end{aligned} \tag{16}$$

where the last inequality follows from the robustness of Π.

Since Z_0 is finite, we can choose a sequence (f_n) in $\mathcal{F}_0$ such that $u \circ f_n = v_n$

[46]Since u is nonconstant, the convex set $Z = u(\Delta_0(X))$ has nonempty interior. After translating and scaling the sequence (v_n) and scaling ε, we may assume that the sequence (v_n) lies in $B_0(\Theta, Z)$. Choose a finite subset Z_0 of Z with mesh $\varepsilon/4$. For each n, replace v_n with the supnorm-closest approximation in $B_0(\Theta, Z_0)$. With this modification, the desired inequality holds with $\varepsilon/2$ in place of ε (which was already scaled above).

for each n and the union $\cup_n\{f_n(\theta) : \theta \in \Theta\}$ is finite and indifference-free. By Lemma 4.i, the sequence (f_n) has a Γ-limit g in $\mathcal{F}_0$. By Lemma 4.ii, $u \circ g = \Gamma\text{-lim}_n v_n = v$. By (16), we can choose a constant act h such that $W_\Pi(v) < u(h) < \liminf_n W_\Pi(v_n)$. Thus, $f_n \succ h$ for all n sufficiently large, but $h \succ g$, contrary to upper Γ-semicontinuity.

B.5 Proof of Lemma 1

We first introduce notation. For any probability measure μ in $\Delta(\Theta)$ and any probability kernel κ, the product $\mu \otimes \kappa$ is the unique measure on the product σ-algebra $\mathcal{B}(\Theta) \otimes \mathcal{B}(\Theta)$ satisfying

$$(\mu \otimes \kappa)(A \times B) = \int_A \kappa(\theta, B)\, \mathrm{d}\mu(\theta), \qquad A, B \in \mathcal{B}(\Theta).$$

Now we turn to the proof. Let d be a bounded, compatible metric on Θ. Let (π_n) be a sequence in $\Delta(\Theta)$ that weakly converges to some prior π in $\Delta(\Theta)$. Let W be the Wasserstein metric induced by d. Since d is bounded, $W(\pi_n, \pi) \to 0$, by Villani (2009, Corollary 6.13, p. 97). For each n, let $\varepsilon_n = \sqrt{W(\pi_n, \pi) + 1/n}$. Thus, $\varepsilon_n \to 0$.

By the definition of the Wasserstein metric, we can choose for each n a probability kernel λ_n such that

$$\pi\lambda_n = \pi_n \quad \text{and} \quad (\pi \otimes \lambda_n)d \leq W(\pi_n, \pi) + 1/n. \tag{17}$$

Let $B(\theta, r)$ denote the closed d-metric ball with center θ and radius r. Let κ_n be the modification of λ_n that fixes any mass that is transported more than distance ε_n. Formally, define the kernel $\kappa_n \colon \Theta \times \mathcal{B}(\Theta) \to [0, 1]$ by

$$\kappa_n(\theta, A) = \lambda_n(\theta, A \cap B(\theta, \varepsilon_n)) + \lambda_n(\theta, \Theta \setminus B(\theta, \varepsilon_n))\delta_\theta(A).$$

The measurability of κ_n follows from Kallenberg (2021, Lemma 3.2.i, p. 56) since for each $\varepsilon > 0$, the set $D(\varepsilon) := \{(\theta, \theta') \in \Theta^2 : d(\theta, \theta') \leq \varepsilon\}$ is closed and hence measurable in $\mathcal{B}(\Theta) \times \mathcal{B}(\Theta)$ by Kallenberg (2021, Lemma 1.2, p. 11).

By construction, $\|\kappa_n\|_d \le \varepsilon_n$. Using (17) and Markov's inequality, we have

$$\begin{aligned}\|\pi_n - \pi\kappa_n\|_{\mathrm{TV}} &= \|\pi\lambda_n - \pi\kappa_n\|_{\mathrm{TV}} \\ &= (\pi \otimes \lambda_n)(\Theta \setminus D(\varepsilon_n)) \\ &\le \varepsilon_n^{-1}(\pi \otimes \lambda_n)d \\ &\le \varepsilon_n^{-1}\left(W(\pi_n, \pi) + 1/n\right) \\ &= \varepsilon_n.\end{aligned}$$

Therefore, $\|\kappa_n\|_d \to 0$ and $\|\pi_n - \pi\kappa_n\|_{\mathrm{TV}} \to 0$.

The desired sequence (κ'_n) can be constructed in the same way: replace λ_n with a kernel λ'_n such that $\pi_n\lambda'_n = \pi$ and $(\pi_n \otimes \lambda'_n)d \le W(\pi_n, \pi) + 1/n$, and complete the proof as before.

B.6 Proof of Lemma 2

For this proof, fix a compatible metric d on Θ that is bounded by 1. Let (v_n) be a bounded sequence in $B(\Theta)$ that Γ-converges to some v in $B(\Theta)$. To show that $\Gamma\text{-}\lim_n\langle v_n, \cdot\rangle = \langle v, \cdot\rangle$, we separately prove the two required properties.

Liminf Fix π in $\Delta(\Theta)$. Let (π_n) be a sequence in $\Delta(\Theta)$ that weakly converges to π. By Lemma 1, there is a sequence (κ_n) of probability kernels such that $\|\kappa_n\|_d \to 0$ and $\|\pi_n - \pi\kappa_n\|_{\mathrm{TV}} \to 0$. Since the sequence (v_n) is bounded, we have

$$\begin{aligned}\liminf_n \langle v_n, \pi_n\rangle &= \liminf_n \langle v_n, \pi\kappa_n\rangle \\ &= \liminf_n \langle \kappa_n v_n, \pi\rangle \\ &\ge \langle \liminf_n \kappa_n v_n, \pi\rangle \\ &\ge \langle v, \pi\rangle,\end{aligned}$$

where the first inequality follows from Fatou's Lemma (which applies because the sequence (v_n) is uniformly bounded below), and the second inequality follows from the pointwise inequality $\liminf_n \kappa_n v_n \geq v$.[47]

Limsup Fix π in $\Delta(\Theta)$. We construct a sequence (π_n) in $\Delta(\Theta)$ that converges to π and satisfies

$$\limsup_n \langle v_n, \pi_n \rangle \leq \langle v, \pi \rangle. \tag{18}$$

Choose strictly positive sequences (δ_j) and (ε_j) that each converge to 0. Since Θ is separable, for each j there exist δ_j-radius balls $B_{j,\ell}$ for $\ell = 1, \ldots, L_j$ such that $\pi(\cup_{\ell=1}^{L_j} B_{j,\ell}) \geq 1 - \varepsilon_j$. To simplify notation below, let $B_{j,L_j+1} = \Theta$. For each $\ell = 1, \ldots, L_j + 1$, choose $\theta_{j,\ell} \in B_{j,\ell}$ such that

$$v(\theta_{j,\ell}) \leq \inf_{\theta \in B_{j,\ell}} v(\theta) + \varepsilon_j.$$

Since $\Gamma\text{-}\lim v_n = v$, we know that for each $\ell = 1, \ldots, L_j + 1$, there exists a sequence $(\theta_{j,\ell}^n)$ such that

$$\lim_n \theta_{j,\ell}^n = \theta_{j,\ell} \quad \text{and} \quad \limsup_n \, v_n(\theta_{j,\ell}^n) \leq v(\theta_{j,\ell}).$$

Therefore, there exists $N_{j,\ell}$ such that for all $n \geq N_{j,\ell}$, we have

$$d(\theta_{j,\ell}^n, \theta_{j,\ell}) \leq \delta_j \quad \text{and} \quad v_n(\theta_{j,\ell}^n) \leq v(\theta_{j,\ell}) + \varepsilon_j.$$

For each j and n, let

$$\pi_j^n = \sum_{\ell=1}^{L_j+1} \pi\left(B_{j,\ell} \setminus \cup_{k=1}^{\ell-1} B_{j,k}\right) \delta(\theta_{j,\ell}^n),$$

where $\delta(\theta)$ denotes the unit mass on θ. Let $N_j = \max_{\ell=1,\ldots,L_j+1} N_{j,\ell}$. If $n \geq N_j$,

[47] To prove this pointwise inequality, fix θ in Θ. For each n, choose θ_n in $\operatorname{supp} \kappa_n(\theta, \cdot)$ such that $v_n(\theta_n) \leq (\kappa_n v_n)(\theta)$. Thus, $d(\theta_n, \theta) \leq \|\kappa_n\|_d \to 0$. Since $\Gamma\text{-}\lim v_n = v$, we conclude that

$$\liminf_n \kappa_n v_n(\theta) \geq \liminf_n v_n(\theta_n) \geq v(\theta).$$

it can be checked that

$$W(\pi_j^n, \pi) \leq 2\delta_j + \varepsilon_j \quad \text{and} \quad \langle v_n, \pi_j^n \rangle \leq \langle v, \pi \rangle + 2\varepsilon_j.$$

For each n, let $\pi_n = \pi_{j(n)}^n$, where $j(n)$ is the largest index j such that $n \geq N_j$. Since $\delta_j \to 0$ and $\varepsilon_j \to 0$, the sequence (π_n) satisfies (18) and converges to π in the Wasserstein metric, and hence weakly, by **?**, Corollary 6.13, p. 97

B.7 Proof of Lemma 3

Let u be state-continuous. Fix f in $\mathcal{F}_0(u)$. Thus, $f = \sum_{j=1}^m x_j A_j$ for some $x_1, \dots, x_m \in \Delta_0(X)$, and some measurable partition $(A_1, \dots, A_m)$ of Θ. For each θ in Θ, let $J(\theta) = \{j : \theta \in \bar{A}_j\}$. Let $g(\theta) = x_{j(\theta)}$, where $j(\theta)$ is the smallest index in $\operatorname{argmin}_{j \in J(\theta)} u(x_j, \theta)$. By construction, $g = \sum_{j=1}^m x_j A'_j$ for some measurable partition $(A'_1, \dots, A'_m)$ of Θ satisfying $A'_j \subset \bar{A}_j$ for each j. Thus, g is in $\mathcal{F}_0(u)$ and g is a graphical limit of f. We claim that $v_g = \operatorname{lsc} v_f$. Fix θ in Θ. There exists a sequence (θ_n) converging to θ such that $f(\theta_n) = g(\theta)$ for all n. Since u is state-continuous, it follows that $v_f(\theta_n) \to v_g(\theta)$. Thus, $v_g \geq \operatorname{lsc} v_f$. For the reverse inequality, note that θ is in the open set $\Theta \setminus \cup_{j \notin J(\theta)} \bar{A}_j$, so for any sequence (θ_n) converging to θ, we know that for all n sufficiently large, θ_n is in $\cup_{j \in J(\theta)} A_j$, hence $v_f(\theta_n) \geq \min_{j \in J(\theta)} u(x_j, \theta_n)$. The right side converges to $v_g(\theta)$ as $n \to \infty$, so we have $\liminf_n v_f(\theta_n) \geq v_g(\theta)$, as desired.

B.8 Proof of Lemma 4

Let $\succsim = \succsim_{(u,\Pi)}$ for some nonconstant, state-independent utility function u and some nonempty subset Π of $\Delta(\Theta)$. Let (f_n) be a sequence in $\mathcal{F}_0$.

i. Suppose that $F \coloneqq \cup_n \{f_n(\theta) : \theta \in \Theta\}$ is finite and indifference-free. Suppose that $(u \circ f_n)$ is Γ-convergent. Let $v = \Gamma\text{-}\lim_n u \circ f_n$. For each θ in Θ, there exists a sequence (θ_n) converging to θ such that $u(f_n(\theta_n)) \to v(\theta)$. Since $\mathcal{U} \coloneqq \cup_n \{u(f_n(\theta)) : \theta \in \Theta\}$ is finite, it follows that $v(\theta)$ is in $\mathcal{U}$. For each θ, let $g(\theta)$ be the unique lottery in F that gives utility $v(\theta)$. By construction, g is in $\mathcal{F}_0$ and $u \circ g = v$. We claim that g is a Γ-limit of (f_n). We check the two required properties.

Since $\Gamma\text{-lim}_n\, u \circ f_n = u \circ g$, there exists a sequence (θ_n) converging to θ such that $u(f_n(\theta_n)) \to u(g(\theta))$. Since $\mathcal{U}$ is finite, there exists m such that for all $n \geq m$, we have $u(f_n(\theta_n)) = u(g(\theta))$ and hence $f_n(\theta_n) = g(\theta)$ (since F is indifference-free).

Let (θ_n) be a sequence converging to θ. Fix h in $\mathcal{F}_0$ with $g(\theta) \succ h$. Since $\Gamma\text{-lim}_n\, u \circ f_n = u \circ g$, we have

$$\liminf_n u(f_n(\theta_n)) \geq u(g(\theta)) > u(h).$$

Thus, there exists m such that for all $n \geq m$, we have $u(f_n(\theta_n)) > u(h)$, hence $f_n(\theta_n) \succ h$.

ii. Let g in $\mathcal{F}_0$ be a Γ-limit of (f_n). We claim that $\Gamma\text{-lim}_n\, u \circ f_n = u \circ g$. Fix θ in Θ. For some sequence (θ_n) converging to θ, there exists m such that $f_n(\theta_n) = g(\theta)$ for all $m \geq n$. Thus, $u(f_n(\theta_n)) \to u(g(\theta))$. Suppose for a contradiction that for some sequence (θ_n) converging to θ, we have

$$\liminf_n u(f_n(\theta_n)) < u(g(\theta)).$$

Then we may select a constant act h such that

$$\liminf_n u(f_n(\theta_n)) < u(h) < u(g(\theta)).$$

Therefore, there exists m such that $h \succ f_n$ for all $n \geq m$, but $g \succ h$, contrary to the fact that g is a Γ-limit of the tail sequence $(f_n)_{n \geq m}$.

B.9 Proof of Lemma 5

For this proof, we introduce some notation. For any measure μ in $\Delta(\Theta)$ and any μ-integrable function $f \colon \Theta \to \mathbf{R}_+$, define the measure $f\mu$ by

$$(f\mu)(A) = \int_A f(\theta)\, \mathrm{d}\mu(\theta), \qquad A \in \mathcal{B}(\Theta).$$

Fix $\varepsilon > 0$. By Prokhorov's theorem (Billingsley, 1999, Theorem 5.2, p. 60), the sequences (π_n) and (π'_n) are both tight. Thus, there exists a compact set K such that for all n we have $\pi_n(K) \geq 1 - \varepsilon$ and $\pi'_n(K) \geq 1 - \varepsilon$.[48] The function H, being continuous, must achieve a maximum over the compact set K. Let $L = 1 + \max_{\theta \in K} H(\theta)$. Let $C = \{\theta \in \Theta : H(\theta) \geq L\}$. The set C is closed and it is disjoint from the compact set K. Therefore, there exists a continuous function $b \colon \Theta \to [0, 1]$ that equals 1 on K and equals 0 on C.[49] For each n, define the nonnegative measure ρ_n by

$$\rho_n = \frac{\langle b, \pi'_n \rangle}{\langle b, \pi_n \rangle} b\pi_n + (1 - b)\pi'_n.$$

By construction, $\rho_n(\Theta) = 1$.

First, we check that ρ_n weakly converges to π. For any bounded, continuous function $f \colon \Theta \to \mathbf{R}$, we have

$$\langle f, \rho_n \rangle = \frac{\langle b, \pi'_n \rangle}{\langle b, \pi_n \rangle} \langle fb, \pi_n \rangle + \langle f(1 - b), \pi'_n \rangle,$$

so

$$\langle f, \rho_n \rangle \to \frac{\langle b, \pi \rangle}{\langle b, \pi \rangle} \langle fb, \pi \rangle + \langle f(1 - b), \pi \rangle = \langle f, \pi \rangle.$$

It remains to check the three properties.

i. We check that $\|\rho_n - \pi_n\|_{\mathrm{TV}} \leq \varepsilon$ for each n. We have

$$\rho_n - \pi_n = \frac{\langle b, \pi'_n - \pi_n \rangle}{\langle b, \pi_n \rangle} b\pi_n + (1 - b)(\pi'_n - \pi_n).$$

We know that $(\rho_n - \pi_n)(\Theta) = 0$ and that $b\pi_n$ is a nonnegative measure.

[48] Since (π_n) is tight, there exists a compact subset K_1 such that $\pi_n(K_1) \geq 1 - \varepsilon$ for all n. Similarly, since (π'_n) is tight, there exists a compact subset K_2 such that $\pi'_n(K_2) \geq 1 - \varepsilon$ for all n. Let $K = K_1 \cup K_2$. The set K is compact and satisfies the desired inequalities.

[49] Here is one construction. The function $\theta \mapsto d(\theta, C)$ is continuous so it achieves its minimum on K. Let $\varepsilon = \min_{\theta \in K} d(\theta, C)$. Since $d(\theta, C) > 0$ for all θ in K, we have $\varepsilon > 0$. Define the function b on Θ by $b(\theta) = (1 - d(\theta, K)/\varepsilon)_+$.

Therefore,

$$\begin{aligned}\|\rho_n - \pi_n\|_{\mathrm{TV}} &\leq \sup_{A \in \mathcal{B}(\Theta)} |\left((1-b)(\pi_n' - \pi_n)\right)(A)| \\ &\leq \max\{\pi_n'(K^c), \pi_n(K^c)\} \\ &\leq \varepsilon,\end{aligned}$$

where the second inequality holds because $1-b$ equals 0 on K.

ii. We check that H is bounded on $\cup_n \operatorname{supp}(\rho_n - \pi_n')$. For each n and any Borel subset A of C, we have $\rho_n(A) = \pi_n'(A)$, so $\operatorname{supp}(\rho_n - \pi_n') \subset \Theta \setminus C$. The function H is bounded above by L on $\Theta \setminus C$.

iii. Let $h \colon \Theta \to \mathbf{R}$ be a continuous function satisfying $|h| \leq H$. By (ii), we know that for each n, the integral $\langle h, \rho_n - \pi_n' \rangle$ is well-defined and finite. For each n, we have

$$\rho_n - \pi_n' = \frac{\langle b, \pi_n' \rangle}{\langle b, \pi_n \rangle} b\pi_n - b\pi_n' = \frac{\langle b, \pi_n' - \pi_n \rangle}{\langle b, \pi_n \rangle} b\pi_n + b\pi_n - b\pi_n'.$$

Therefore,

$$\begin{aligned}|\langle h, \rho_n - \pi_n' \rangle| &\leq \frac{|\langle b, \pi_n' - \pi_n \rangle|}{\langle b, \pi_n \rangle} |\langle hb, \pi_n \rangle| + |\langle hb, \pi_n - \pi_n' \rangle| \\ &\leq \frac{|\langle b, \pi_n' - \pi_n \rangle|}{1 - \varepsilon} L + |\langle hb, \pi_n - \pi_n' \rangle|,\end{aligned} \tag{19}$$

where the second inequality follows from the inequalities $|hb| \leq |Hb| \leq L$ and $\langle b, \pi_n \rangle \geq \pi_n(K) \geq 1 - \varepsilon$. As n tends to ∞, the right side of (19) tends to 0 because (π_n) and (π_n') each converge weakly to π (and the functions b and hb are bounded and continuous).

B.10 Proof of Lemma 6

Fix a prior π in $\Delta(\Theta)$. We first construct a dense π-nullset N. Let Θ_0 consist of all points in Θ with positive π-measure. Since $\pi(\Theta) < \infty$, the set Θ_0 must be countable. For each θ in Θ, the complement $\Theta \setminus \{\theta\}$ is open and dense (since θ cannot be an isolated point because Θ is perfect). By the Baire category theorem, the set $\Theta \setminus \Theta_0 = \cap_{\theta \in \Theta_0} \Theta \setminus \{\theta\}$ is dense as well. Since Θ is separable, there exists a countable subset N of $\Theta \backslash \Theta_0$ that is dense in Θ. Thus, $\pi(N) = 0$.

Let (θ_j) be an enumeration of N. Select a bounded, compatible metric d. For each n, let

$$\pi_n = \sum_{j=1}^{\infty} \pi\left(B(\theta_j, 1/n) \setminus \cup_{k=1}^{j-1} B(\theta_k, 1/n)\right) \delta_{\theta_j}.$$

Using Tonelli's theorem, it can be shown that π_n is countably additive and hence a probability measure. By construction, $W(\pi_n, \pi) \leq 1/n$, so the sequence (π_n) weakly converges to π in the Wasserstein metric, and hence weakly, by Villani (2009, Corollary 6.13, p. 97).