

How to Make an Action Better

Marilyn Pease • Mark Whitmeyer^[1]

For two actions in a decision problem, a and b , each of which produces a state-dependent monetary reward, we study how to robustly make action a more attractive. Action \hat{a} improves upon a in this manner if the set of beliefs at which a is preferred to b is a subset of the set of beliefs at which \hat{a} is preferred to b , *irrespective of the risk-averse agent's utility function (in money)*. We provide a full characterization of this relation and discuss applications in politics, bilateral trade, insurance, and information acquisition.

^[1] MP: Kelley School of Business, Indiana University, marpease@iu.edu & MW: Arizona State University, mark.whitmeyer@gmail.com. We thank Arjada Bardhi, Henrique Castro-Pires, Faruk Gul, Ilia Krasikov, Pietro Ortoleva, Sungmin Park, and seminar audiences at ASU and OSU for their comments. We also gratefully acknowledge Konstantin von Beringe and Titus Gao for their careful proofreading of the results. Draft date: November 8, 2024.

1. Introduction

There are many situations in which one agent must decide which option to present to another when there is uncertainty about the outcome of the options. Consider, for instance, a consumer deciding between firm 1 and firm 2's products, without knowing which best suits her needs. Firm 2 produces product b , while firm 1 can choose to produce either product a or product \hat{a} . If the consumer prefers a to b , what are the properties of \hat{a} that guarantee that the consumer also prefers it to b ? What if firm 1 knows neither the consumer's subjective belief nor her utility function?

To elaborate, suppose firm 1 and firm 2 are auto manufacturers. Each firm's car performs differently in different driving conditions, offering one monetary payoff to the consumer when driven in the city and another when driven on the highway. Firm 1 knows that the consumer prefers product a to product b , but knows neither the consumer's belief about her likelihood of city versus highway driving nor her utility function. What properties must an alternative product \hat{a} have for the consumer to also prefer it to b , given that she prefers a ?

The same exercise can be conducted in the context of political elections, where party A must choose between candidates a and \hat{a} to run against party B 's candidate b . When would party A benefit more from running candidate \hat{a} over candidate a , given that b is the opponent? Allowing for a population with wildly heterogeneous beliefs and utility functions, what are the characteristics of \hat{a} that would ensure at least as many people choose it over b as they would a ?

The knee-jerk response to these questions is that it is obvious: when the decision-maker's (DM's)—i.e., consumer's or voter's—utility is known only to be within the class of increasing-in-money utility functions, it must be that \hat{a} corresponds to a first-order stochastic dominance improvement over a . Likewise, when the DM is known also to be risk averse, it must be that \hat{a} corresponds to a second-order stochastic dominance improvement over a . However, these answers do not stand up to scrutiny—in particular, we do not know the lotteries produced by the DM's subjective belief, so we cannot speak directly of dominance of lotteries. State-wise

dominance of a or b is also a promising answer, but it turns out that that is too strong when the agent is known to be risk averse.

In our two main results, Theorems 3.1 and 3.6, we fully characterize these relations in terms of the primitives of the environment—the state-dependent payoffs to actions a , \hat{a} , and b . Theorem 3.1 concerns the case of a risk-averse DM. We show that the DM preferring a to b must imply that she also prefers \hat{a} to b if and only if \hat{a} dominates an action whose payoffs are a convex combination of the payoffs to a and b . Intuitively, if a is preferred to b , then mixing some of a with b is better than b as well. Consequently, the set of beliefs at which \hat{a} is preferred must be at least as large as the set where a was preferred. Then, making \hat{a} better than a mixture of a and b only strengthens its attractiveness versus b .

If we do not assume the DM to be risk averse, our other main result (Theorem 3.6) states that the necessary and sufficient conditions become much more restrictive: \hat{a} dominates a or b (in a state-wise sense). If u is concave, we know from Theorem 3.1 that significant structure on payoff transformations is required. Convexity also imposes structure, but in the opposite manner. Combining these requirements leaves only dominance as necessary (and it is obviously sufficient).

There are a number of natural extensions to our main research question. First, what if we require robust improvements to an action vis-a-vis not just one but multiple alternatives? In §5.1 we argue that such robust improvements correspond to dominance improvements except in special cases. Second, what if we ask for robustness to aggregate risk? We reveal in §5.2 that the conditions for our main theorems remain necessary and sufficient—that is, our definition can be extended—to allow for aggregate risk with no alteration to the results. Third, we discuss briefly modifications to *both* a and b (§5.3) and provide a sufficient condition for \hat{a} to be preferred to b 's replacement, \hat{b} . Fourth, §7 explores information acquisition. Now the question becomes: how must a relate to \hat{a} in a way that guarantees that \hat{a} is chosen more frequently in the DM's problem of binary-choice with flexible information acquisition? With two states, we show that \hat{a} dominating a or b is necessary and sufficient.

We also discuss several examples. Up first is a political example: in a two-party election, suppose a party is contemplating whether to replace its nominee, candidate a . Given that the opponent is candidate b , what are the properties of a candidate that is robustly more appealing versus b than a ? Our first theorem applies directly, so that she must dominate a convex combination of a and b in the sense of our first theorem. Our other two examples posit a status quo trading arrangement—a trade of a risky asset, or a sale of an insurance contract—and asks what contracts must be accepted given that the status quo one is. Again, our main result tells us that they must be convex combinations of sorts (now, with respect to the outside options).

1.1. Related Work

The body of work studying decision-making under uncertainty is sizeable. The work closest to this one is [Pease and Whitmeyer \(2023\)](#). There, we formulate a binary relation between actions: action a is safer than b if the the set of beliefs at which a is preferred to b grows larger, in a set inclusion sense, when we make the DM more risk averse.

[Rothschild and Stiglitz \(1970\)](#) is a seminal work that characterizes (mean-preserving) transformations of lotteries that are preferred by all risk-averse agents. [Aumann and Serrano \(2008\)](#) formulate an “measure of riskiness” of gambles, as do [Foster and Hart \(2009\)](#) (who are subsequently followed up upon by [Bali, Cakici, and Chabi-Yo \(2011\)](#) and [Riedel and Hellmann \(2015\)](#)). Crucially, these indices and measures correspond to inherently stochastic objects—the lotteries at hand. Our conception of an improvement to an action concerns comparisons of state-dependent payoffs, which are themselves non-random objects (they are just real numbers).

Naturally this work is also connected to the broader literature studying actions that are comparatively friendly toward risk. In addition to the aforementioned paper of ours, [Pease and Whitmeyer \(2023\)](#), which, like this one, centers around a decision-maker’s state-dependent payoffs to actions, this literature includes

Hammond III (1974), Lambert and Hey (1979), Karlin and Novikoff (1963), Jewitt (1987), and Jewitt (1989). Notably, they are statements about lotteries, *viz.*, random objects.

Ours is vaguely a comparative statics work—we’re changing an aspect of a decision problem and seeing how it affects a decision-maker’s choice. Our robustness criterion as well as the simplicity of our setting distinguishes our work from the standard pieces, e.g., Milgrom and Shannon (1994), Edlin and Shannon (1998), and Athey (2002). The works involving aggregation (Quah and Strulovici (2012), Choi and Smith (2017), and Kartik, Lee, and Rappoport (2023)) are closer still—as this inherently corresponds to distributional robustness—but none leave as free parameters both the distribution over states and the DM’s utility function, as we do. Special mention is due to Curello and Sinander (2019), who conduct a robust comparative statics exercise in which an analyst, with limited knowledge of an agent’s preferences, predicts the agent’s choice across menus.

In our final section, we explore properties of the new action, \hat{a} , that make it more likely to be selected than a when the alternative is b if information is endogenously acquired. This property is similar in spirit to the observation of Matějka and McKay (2015) that new actions added to a menu may “activate” previously unchosen actions. Muller-Itten, Armenter, and Stangebye (2021) provide a full characterization of this phenomenon. One crucial distinction between our analysis and theirs is that they explore *additions* whereas our modification is a *replacement*.

2. Model

There is a topological space of states, Θ , which is endowed with the Borel σ -algebra, and which we assume to be compact and metrizable. θ denotes a generic element of Θ . We denote the set of all Borel probability measures on Θ by $\Delta \equiv \Delta(\Theta)$. There is also a decision-maker (DM), who is endowed with two actions, a and b . $A = \{a, b\}$ denotes the set of actions, and each action $\tilde{a} \in A$ is a continuous function from the state space to the set of outcomes, $\tilde{a}: \Theta \rightarrow \mathbb{R}$. For convenience, for any $\tilde{a} \in A$, we write $\tilde{a}_\theta \equiv \tilde{a}(\theta)$. Given a probability distribution over states

$\mu \in \Delta$, an action is a (simple) lottery.

We further specify that no action $\tilde{a} \in A$ is weakly dominated by the other. Therefore, we can partition Θ into three sets:

$$\begin{aligned} \mathcal{A} &:= \{\theta \in \Theta: a_\theta > b_\theta\} \quad \mathcal{B} := \{\theta \in \Theta: a_\theta < b_\theta\}, \\ \text{and } \mathcal{C} &:= \{\theta \in \Theta: a_\theta = b_\theta\}. \end{aligned}$$

The DM is an expected-utility maximizer, with a von Neumann-Morgenstern utility function defined on the outcome space $u: \mathbb{R} \rightarrow \mathbb{R}$. We posit that u is strictly increasing, weakly concave, and continuous. On occasion, we will drop the assumption that u is weakly concave, merely requiring that it be strictly increasing (and continuous).

Given A , we are interested in how a modification of a affects the DM's choice of action. That is, we will modify a to some new \hat{a} , in which case the new menu is (\hat{a}, b) , and examine when \hat{a} must be chosen more by the DM than a . To that end, we define the set $P_b(a)$ to be the subset of the probability simplex on which action a is weakly preferred to b ; formally,

$$P_b(a) := \{\mu \in \Delta: \mathbb{E}_\mu u(a_\theta) \geq \mathbb{E}_\mu u(b_\theta)\}.$$

$\bar{P}_b(a)$, in turn, is the subset of the probability simplex on which a is strictly preferred to b :

$$\bar{P}_b(a) := \{\mu \in \Delta: \mathbb{E}_\mu u(a_\theta) > \mathbb{E}_\mu u(b_\theta)\}.$$

When a is transformed to \hat{a} , we define the analogous sets $P_b(\hat{a})$ and $\bar{P}_b(\hat{a})$.

Definition 2.1. *Action \hat{a} is **b-Superior** to action a if, no matter the risk-averse agent's utility function,*

1. $P_b(a) \subseteq P_b(\hat{a})$ and
2. $\bar{P}_b(a) \subseteq \bar{P}_b(\hat{a})$

i.e., the set of beliefs at which the DM prefers a to b is a subset of the beliefs at which the DM prefers \hat{a} to b .

To put differently, action \hat{a} is b -superior to action a if

$$\mathbb{E}_\mu u(a_\theta) \geq \mathbb{E}_\mu u(b_\theta) \Rightarrow \mathbb{E}_\mu u(\hat{a}_\theta) \geq \mathbb{E}_\mu u(b_\theta),$$

and

$$\mathbb{E}_\mu u(a_\theta) > \mathbb{E}_\mu u(b_\theta) \Rightarrow \mathbb{E}_\mu u(\hat{a}_\theta) > \mathbb{E}_\mu u(b_\theta),$$

for any strictly increasing, concave, and continuous u .^[2]

3. Robust Improvements

We now present our main result, characterizing the conditions under which \hat{a} is b -Superior to a . To do so, we first define a **Mixture** of actions a and b to be the action a^λ that yields payoff

$$a_\theta^\lambda := \lambda a_\theta + (1 - \lambda) b_\theta$$

in each state $\theta \in \Theta$ for some $\lambda \in [0, 1]$.

Our main result is, understanding dominance in a state-wise sense,

Theorem 3.1. *Fix a and b . Action \hat{a} is b -superior to action a if and only if $\hat{a}_\theta > b_\theta$ for all $\theta \in \mathcal{A}$ and \hat{a} (weakly) dominates a mixture of a and b .*

Proof. (\Leftarrow) Suppose \hat{a} is a mixture of a and b (as subsequent dominance only makes \hat{a} more enticing). Let L_a be the lottery that pays out a_θ with probability $\mu(\theta)$; L_b , the lottery that pays out b_θ with probability $\mu(\theta)$; and $L_{\hat{a}}$, the lottery that pays out \hat{a}_θ with probability $\mu(\theta)$. Also suppose $\mu \in \Delta$ is such that the DM prefers a to b , i.e., $L_a \succeq L_b$.

By independence (which is implied by expected utility), for any $\lambda \in [0, 1]$,

$$L^\lambda := \lambda L_a + (1 - \lambda) L_b \succeq L_b.$$

Moreover, by assumption there is some $\lambda^* \in [0, 1]$ for which $\lambda^* a_\theta + (1 - \lambda^*) b_\theta = \hat{a}_\theta$ for all $\theta \in \Theta$. This implies that L^{λ^*} is a mean preserving spread of $L_{\hat{a}}$. Consequently, as the DM is risk averse, $L_{\hat{a}} \succeq L^{\lambda^*} \succeq L_b$. In other words, at any belief $\mu(\theta)$ such that a is preferred to b , \hat{a} is also preferred to b , so that \hat{a} is b -superior to a .^[3] ■

^[2] b -superiority imposes that the agent is risk-averse. When we drop this assumption in §3.2, we introduce a new definition for \hat{a} to be preferred more versus b than a (Definition 3.5).

^[3] This sufficiency result persists for preferences other than expected utility (see Appendix B).

The necessity proof of Theorem 3.1 is a little more involved. We prove it in two steps. We say that \hat{a} pairwise-dominates a collection of mixtures of a and b if for any pair $(\theta, \theta') \in \mathcal{A} \times \mathcal{B}$, there exists a $\lambda_{\theta, \theta'} \in [0, 1]$ such that

$$\hat{a}_\theta \geq \lambda_{\theta, \theta'} a_\theta + (1 - \lambda_{\theta, \theta'}) b_\theta \quad \text{and} \quad \hat{a}_{\theta'} \geq \lambda_{\theta, \theta'} a_{\theta'} + (1 - \lambda_{\theta, \theta'}) b_{\theta'}. \quad (1)$$

Then, the first of the two necessity steps is showing that b -superiority implies pairwise dominance.

Lemma 3.2. *\hat{a} is b -superior to a only if i) for any $\theta^\dagger \in \mathcal{C}$, $\hat{a}_{\theta^\dagger} \geq a_{\theta^\dagger}$; ii) for any $\theta \in \mathcal{A}$, $\hat{a}_\theta > b_\theta$; and iii) \hat{a} pairwise-dominates a collection of mixtures of a and b .*

Proof. See Appendix A. ■

Let us consider the meaning of Lemma 3.2. First, it is clear that if \hat{a} is b -superior, then i) and ii) must be true. Next, if iii) is not true and if $\hat{a}_{\theta'} < a_{\theta'} < b_{\theta'}$ and $\hat{a}_\theta > a_\theta > b_\theta$, there are beliefs—assigning zero probability to all states other than the specified θ and θ' —at which a risk-neutral DM will rank the actions $a \succ b \succ \hat{a}$, so \hat{a} is not b -superior. On the other hand, if these two chains of inequalities do hold, although $a \succeq b$ will always imply $\hat{a} \succeq b$ for a risk-neutral DM, we can nevertheless always construct a (concave) utility function such that there exist—again, with positive probability assigned only to states θ and θ' —beliefs that produce $a \succ b \succ \hat{a}$.

Finally, we show that pairwise domination implies domination of a mixture, closing the proof of necessity.

Lemma 3.3. *If \hat{a} pairwise-dominates a collection of mixtures of a and b , $\hat{a}_{\theta^\dagger} \geq a_{\theta^\dagger}$ for all $\theta^\dagger \in \mathcal{C}$, and $\hat{a}_\theta > b_\theta$ for all $\theta \in \mathcal{A}$, then \hat{a} dominates a mixture of a and b .*

Proof. Suppose \hat{a} pairwise-dominates a collection of mixtures of a and b , $\hat{a}_{\theta^\dagger} \geq a_{\theta^\dagger}$ for all $\theta^\dagger \in \mathcal{C}$, and $\hat{a}_\theta > b_\theta$ for all $\theta \in \mathcal{A}$. This implies that for any pair $(\theta, \theta') \in \mathcal{A} \times \mathcal{B}$,

$$\min \left\{ 1, \frac{\hat{a}_\theta - b_\theta}{a_\theta - b_\theta} \right\} \geq \lambda_{\theta, \theta'} \geq \max \left\{ 0, \frac{b_{\theta'} - \hat{a}_{\theta'}}{b_{\theta'} - a_{\theta'}} \right\}.$$

This implies that there exists a $\lambda \in [0, 1]$ for which

$$\inf_{\theta \in \mathcal{A}} \left\{ 1, \frac{\hat{a}_\theta - b_\theta}{a_\theta - b_\theta} \right\} \geq \lambda \geq \sup_{\theta' \in \mathcal{B}} \left\{ 0, \frac{b_{\theta'} - \hat{a}_{\theta'}}{b_{\theta'} - a_{\theta'}} \right\},$$

and so \hat{a} dominates a mixture of a and b . ■

We note that Theorem 3.1 can easily be extended to the case in which a lower-bound (in terms of risk-aversion) for the DM is known. That is, suppose there is some (weakly) risk-averse \bar{u} of which the DM's utility is known to be some strictly increasing (weakly) concave transformation: $u = \phi \circ \bar{u}$. Then, we just redefine a from map $\theta \mapsto a(\theta)$ to $\theta \mapsto \bar{u}(a(\theta))$ and likewise b , before applying the theorem.

3.1. Intuition

To gain intuition for Theorem 3.1, consider the two-state environment, which is easy to visualize. The state is either 0 or 1, and we specify without loss of generality that $a_0 > b_0$ and $b_1 > a_1$. $\mu \in [0, 1]$ denotes the DM's belief that the state is 1.

There is a cutoff belief at which the DM is indifferent between actions. Let $\bar{\mu}_u$ be the belief at which the DM is indifferent between actions a and b with utility u and $\hat{\mu}_u$ be the belief at which the DM is indifferent between \hat{a} and b with the same utility. To see this graphically, let

$$\ell_a(\mu) = a_0(1 - \mu) + a_1\mu,$$

be the expected payoff (in money) of action a as a function of the belief $\mu \in [0, 1]$. In turn, lines

$$\ell_{\hat{a}}(\mu) = \hat{a}_0(1 - \mu) + \hat{a}_1\mu \quad \text{and} \quad \ell_b(\mu) = b_0(1 - \mu) + b_1\mu,$$

are the expected payoffs to actions \hat{a} and b , respectively. In Figure 1, these are the blue (solid), red (dashed), and green (dotted) lines. The top panel of the figure shows the monetary payoffs, or the risk-neutral utility function, while the bottom panel depicts utilities with $u(\cdot) = \sqrt{1 + \cdot} - 3$. Then, \hat{a} being b -superior to a is equivalent to $\bar{\mu}_u \leq \hat{\mu}_u$ for all concave u .

Now we can think about b -superiority in the context of Theorem 3.1. It tells us that \hat{a} must at least weakly dominate a mixture of a and b . For now, just consider an \hat{a} that is a mixture of a and b . The corresponding line, $\ell_{\hat{a}}$, is a counterclockwise rotation of ℓ_a . We know from the theorem that for any belief at which the DM preferred a to b ($\mu \leq \bar{\mu}_u$), she will prefer \hat{a} to b as well. In other words, if a was better than b at μ , then mixing “some” a with b is better than b at μ as well.

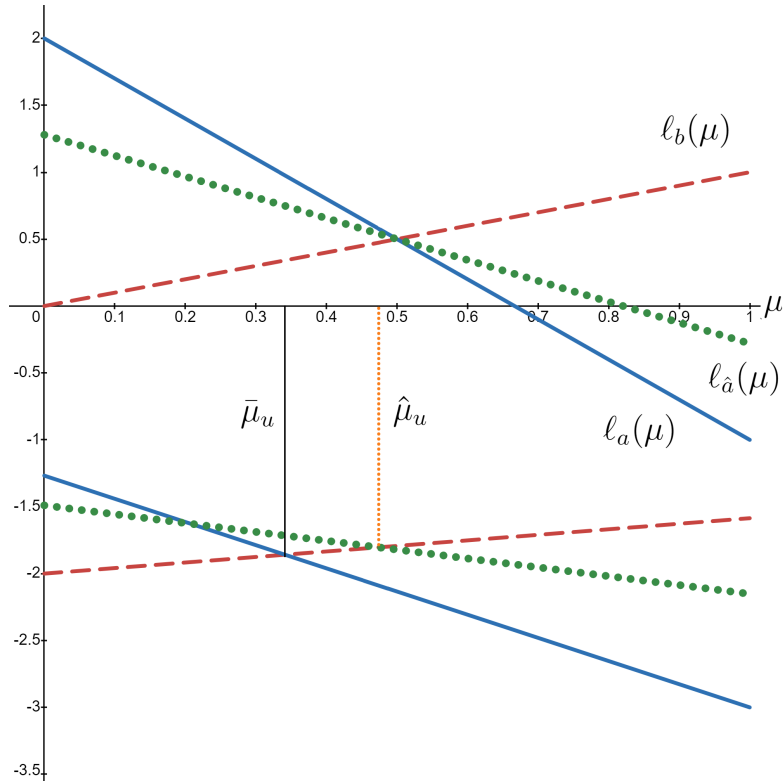


Figure 1: Improving a . Try it yourself (rotate by moving the slider)!

Consequently, the set of beliefs at which \hat{a} is preferred must be *at least* as large as the set where a was preferred. Naturally, making \hat{a} strictly better than a mixture of a and b only strengthens its attractiveness versus b .

Formally, we can write the following result.

Corollary 3.4. *\hat{a} is b -superior to a if and only if $\hat{a}_0 > b_0$ and there exists a $\lambda \in [0, 1]$ such that $\ell_{\hat{a}}$ lies weakly above the line $\lambda \ell_a + (1 - \lambda) \ell_b$.*

3.2. A Broader Class of Utilities

So far we have assumed mild but meaningful structure on the DM's utility function. More is better (u is strictly increasing in the reward), and the DM is risk-averse (u is, at least weakly, concave). Suppose we remove the assumption of risk aversion. We still assume that u is strictly increasing but now allow it to take arbitrary shape.

We then define the equivalent notion to b -superiority without risk aversion.

Definition 3.5. Action \hat{a} is *b-Better* than action a if, no matter the (not necessarily risk-averse) agent's utility function,

1. $P_b(a) \subseteq P_b(\hat{a})$ and
2. $\bar{P}_b(a) \subseteq \bar{P}_b(\hat{a})$

i.e., the set of beliefs at which the DM prefers a to b is a subset of the beliefs at which the DM prefers \hat{a} to b

Put differently, action \hat{a} is b -better than action a if

$$\mathbb{E}_\mu u(a_\theta) \geq \mathbb{E}_\mu u(b_\theta) \Rightarrow \mathbb{E}_\mu u(\hat{a}_\theta) \geq \mathbb{E}_\mu u(b_\theta),$$

and

$$\mathbb{E}_\mu u(a_\theta) > \mathbb{E}_\mu u(b_\theta) \Rightarrow \mathbb{E}_\mu u(\hat{a}_\theta) > \mathbb{E}_\mu u(b_\theta),$$

for any strictly increasing, but not necessarily concave, continuous u .

As we now document, without the assumption of risk aversion, the only transformations that must make b less attractive are unambiguous dominance improvements from a to \hat{a} . This is intuitive; think back to Corollary 3.4. There, the necessary relationship between \hat{a} and a is $\ell_{\hat{a}}$'s dominance of a counterclockwise rotation of ℓ_a . However, the opposite modification is needed when u is convex—which is now permitted—that is, $\ell_{\hat{a}}$ needs to dominate a *clockwise* rotation of ℓ_a . This leaves only dominance improvements.

Theorem 3.6. Action \hat{a} is b -better than action a if and only if for all $\theta \in \mathcal{A}$ $\hat{a}_\theta > b_\theta$ and \hat{a} dominates a or b .

Proof. See Appendix A. ■

4. Applications

We now consider three applications of Theorem 3.1 to solidify intuition and understanding.

4.1. Choosing Which Candidate to Run

There is a unit mass of voters and two candidates a and b running for office. There is an unknown state of the world $\theta \in \Theta$, and candidate a (b) produces monetary benefit a_θ (b_θ) to each voter in every state $\theta \in \Theta$. Neither candidate dominates the other in the sense that there is at least one state of the world in which each provides a strictly higher monetary benefit to the populace than the other.

The party backing candidate a is contemplating whether to replace him with a different candidate \hat{a} . Crucially, it knows neither the risk preferences of the populace (which may be heterogeneous) nor the beliefs of the populace (also possibly heterogeneous). Accordingly, it makes its choice based on a robust criterion, asking what are the candidates who have a better chance of winning versus b , no matter the population's (concave) utilities or beliefs. Who cannot be worse than a versus b ? Theorem 3.1 tells us the answer:

Remark 4.1. *Candidate \hat{a} is no worse than a versus b if and only if she dominates a mixture of a and b .*

4.2. Robustly Optimal Bilateral Trade Modifications

There is a buyer (B) and a seller (S). S possesses an asset that pays out $v_\theta \in \mathbb{R}$ in state θ . The state is contractible and the status quo trade agreement sees transfers of size $\gamma_\theta \in \mathbb{R}$ from B to S in each state $\theta \in \Theta$. We assume that there exist states $\theta', \theta'' \in \Theta$ for which $v_{\theta'} > \gamma_{\theta'}$ and $\gamma_{\theta''} > v_{\theta''}$.

B 's state-dependent payoff from transacting is $v_\theta - \gamma_\theta$, and S 's is γ_θ (budget balance). B 's outside option is the sure-thing 0 and S 's outside option is the asset, with random payoff v_θ in each state $\theta \in \Theta$.

We suppose that the status-quo is acceptable—given the agents' subjective beliefs about uncertainty and their (concave) utility functions, each is willing to participate in the arrangement. What modifications to the arrangement are robustly optimal in that B and S will still remain willing to participate? Theorem 3.1 reveals the answer.

We observe that for all $\theta \in \Theta$ and for any $\lambda \in [0, 1]$,

$$\lambda(v_\theta - \gamma_\theta) = v_\theta - \lambda\gamma_\theta - (1 - \lambda)v_\theta,$$

so that any new transfer that yields the requisite convex combination of B 's status quo payoff from accepting the terms and B 's outside option of 0 equals

$$\lambda\gamma_\theta + (1 - \lambda)v_\theta,$$

which is precisely the convex combination (with the same weight) of S 's status quo payoff and her outside option. This, plus our assumption of budget balance, imply

Remark 4.2. *A new trade agreement, $(\hat{\gamma}_\theta)_{\theta \in \Theta}$ must be acceptable if and only if there exists some $\lambda \in (0, 1]$ such that $\hat{\gamma}_\theta = \lambda\gamma_\theta + (1 - \lambda)v_\theta$ for all $\theta \in \Theta$.*

4.3. Robustly Optimal Insurance Modifications

Now consider the scenario of a risk-neutral insurer and a consumer. The consumer's payoff without insurance is v_θ in each state $\theta \in \Theta$. The status-quo policy pays out (net of the loss) α_θ to the consumer in each state $\theta \in \Theta$. We assume that there exist some $\theta', \theta'' \in \Theta$ for which $\alpha_{\theta'} > v_{\theta'}$ and $\alpha_{\theta''} < v_{\theta''}$.

Remark 4.3. *A new contract, $(\hat{\alpha}_\theta)_{\theta \in \Theta}$ must be acceptable if and only if there exists some $\lambda \in (0, 1]$ such that $\hat{\alpha}_\theta \geq \lambda\alpha_\theta + (1 - \lambda)v_\theta$ for all $\theta \in \Theta$.*

One interesting implication of this remark is that the set of acceptable contracts could very well expose the consumer to more risk. Indeed, suppose that the status-quo policy is risk free, i.e., yields the consumer the same wealth in every state. The consumer's outside option (not buying insurance) is, of course, riskier. In this case, any actuarially-fair contract that the consumer must accept also exposes her to risk as her wealth is (non-trivially) state-dependent.

5. Extensions

In this section, we extend the notion of b -superiority to encompass more than pairwise comparisons. We show that the notion and intuition easily extend to more complicated settings.

5.1. One Versus Many

What if we look for improvements to a that make it more attractive in comparison to multiple alternatives, not just a single one? Consider the scenario from the introduction where the manager of firm 1 is deciding between products a and \hat{a} , but now the firm is competing against a product from firm 2 as well as a product from firm 3. How does the inclusion of firm 3 affect the manager's decision?

For simplicity, we restrict our attention to the two-state environment, $\Theta = \{0, 1\}$. Let A be the DM's finite set of actions, where we assume that none are (strictly) dominated. Indexing the actions by $I = \{1, \dots, m\}$ (with $m \geq 3$), we label the actions so that for all $i \in I$, $a_0^i > a_0^j$ and $a_1^i < a_1^j$ for all $j > i$ ($j \in I$).

There are two special actions, a^1 and a^m . The first is uniquely optimal in state 0 and the second is uniquely optimal in state 1. We term these the two extreme actions. For a specified $a \in A$, we define $B := A \setminus \{a\}$ and

$$P_B(a) := \left\{ \mu \in \Delta : \mathbb{E}_\mu u(a_\theta) \geq \max_{a^j \in B} \mathbb{E}_\mu u(a_\theta^j) \right\}.$$

We define $\bar{P}_B(a)$ to be the analogous set with a strict inequality, and sets $P_B(\hat{a})$ and $\bar{P}_B(\hat{a})$ their compatriots once we've replaced a with \hat{a} .

Definition 5.1. *Action \hat{a} is **B-Superior** to action a if the set of beliefs at which the DM prefers a to the actions other than a , $P_B(a)$, is a subset of the set of beliefs at which the DM prefers \hat{a} to those other actions, $P_B(\hat{a})$; and $\bar{P}_B(a) \subseteq \bar{P}_B(\hat{a})$; no matter the risk-averse agent's utility function.*

We have the following stark result.

Proposition 5.2. *If $a = a^1$ ($a = a^m$), \hat{a} is B-superior to a if and only if a is b-superior to a^2 (a^{m-1}). If a is not one of the extreme actions, \hat{a} is B-superior to a if and only if \hat{a} dominates a .*

The logic behind this proposition is straightforward. For an action to improve versus the action to its right, we need the induced (risk-neutral) line in belief space to lie above some counterclockwise rotation around the indifference point between the two actions. For an action to improve versus the action to its left, we need the

line in belief space to lie above some *clockwise* rotation about the indifference point between the two actions (which is strictly to the left of the other indifference point). A dominance shift is the only way to reconcile these requirements for non-extreme actions.

For concreteness, think about the 3-firm scenario and consider the following example. Let $a_0^1 = 2 > a_0^2 = 1.5 > a_0^3 = 0$ and $a_1^1 = 0 < a_1^2 = 1.5 < a_1^3 = 2$. Then, firm 1 and 3's products are well-matched to the consumer in states 0 and 1 respectively, while firm 2 offers a safe alternative when the state is relatively uncertain. Proposition 5.2 tells us that if either firm 1 or 3 is considering improving its product, the firm must only make it *b*-superior relative to the safe option. If firm 2 wishes to improve its product, however, it must dominate the current payoff of 1.5. In short, if firm 1 is to expand the set of beliefs at which the consumer chooses it, it need only improve relative to 2, but for 2 to expand, it would have to improve relative to both 1 and 3.

5.2. Robustness to Aggregate Risk

What if the DM also has unknown (to the DM) random wealth? We show that Theorems 3.1 and 3.6 are robust to the addition of wealth. Keep in mind the example of firms 1 and 2. How does the consumer having an unknown level of wealth change the firm's choice between products a and \hat{a} ? Formally, in any state θ , the DM's utility from action a is $\mathbb{E}_H u(a_\theta + W)$, where W is a state-independent (finite-mean) random variable, distributed according to $H(\cdot)$.

Abusing notation, we define $P_b(a)$ to be the subset of Δ for which a is preferred to b :

$$P_b(a) := \{\mu \in \Delta : \mathbb{E}_{\mu,H} u(a_\theta + W) \geq \mathbb{E}_{\mu,H} u(b_\theta + W)\}.$$

$\bar{P}_b(a)$ is this set's strict-preference counterpart.

We extend our definitions in the natural way:

Definition 5.3. *Action \hat{a} is **b-** and **Wealth-Superior** to action a if*

1. $P_b(a) \subseteq P_b(\hat{a})$ and

$$2. \bar{P}_b(a) \subseteq \bar{P}_B(\hat{a}),$$

no matter the risk-averse agent's utility function and no matter the agent's state-independent random wealth W .

Definition 5.4. Action \hat{a} is *b- and Wealth-Better* than action a if

1. $P_b(a) \subseteq P_b(\hat{a})$ and
2. $\bar{P}_b(a) \subseteq \bar{P}_B(\hat{a}),$

no matter the (not necessarily risk-averse) agent's utility function and no matter the agent's state-independent random wealth W .

Then,

Corollary 5.5. \hat{a} is *b- and wealth-superior* to a if and only if $\hat{a} > b_\theta$ for all $\theta \in \mathcal{A}$ and \hat{a} (weakly) dominates a mixture of a and b . \hat{a} is *b- and wealth-better* than a if and only if for all $\theta \in \mathcal{A}$ $\hat{a}_\theta > b_\theta$ and \hat{a} dominates a or b .

Proof. Please visit Appendix [A.4](#). ■

Clearly, the addition of wealth does not alter how we should be comparing a to \hat{a} . It is easy to see that \hat{a} dominating a mixture is also necessary and sufficient for \hat{a} being preferred to b whenever a is, for any common state-dependent component of the DM's wealth. Namely, observe that if \hat{a} dominates a mixture of a and b , $\hat{a} + d$ dominates a mixture of $a + d$ and $b + d$ for any $d: \Theta \rightarrow \mathbb{R}$: for all $\theta \in \Theta$, there exists some $\lambda \in [0, 1]$ such that

$$\lambda(a_\theta + d_\theta) + (1 - \lambda)(b_\theta + d_\theta) = \lambda a_\theta + (1 - \lambda)b_\theta + d_\theta \leq \hat{a}_\theta + d_\theta.$$

5.3. Everything Everywhere All at Once

So far, we have focused on altering a single action, leaving the other action (or actions) unaltered. This is reasonable in many situations—for example, if the other action is the DM's outside option, which is unalterable. But what if we change both actions a and b ? Suppose that it is known that firm 2 is changing its product from b to \hat{b} , and firm 1 is considering moving from product a to product \hat{a} but

only wishes to do so if it can at least maintain market share. What type of \hat{a} will achieve this goal?

When both actions change, the basic intuition from the earlier results remains true, though we only provide a partial analysis, primarily to avoid getting bogged down in a tedium of cases. We also restrict attention to the two-state environment ($\Theta = \{0, 1\}$), as it is straightforward to aggregate the conditions (as in [Pease and Whitmeyer \(2023\)](#)) to extend the result to a general state space. Without loss of generality, let $a_0 > b_0$ and $a_1 < b_1$.

Now, we are going from a to \hat{a} and b to some \hat{b} . To that end, we define

$$P_{\hat{b}}(\hat{a}) := \left\{ \mu \in \Delta : \mathbb{E}_{\mu} u(\hat{a}_{\theta}) \geq \mathbb{E}_{\mu} u(\hat{b}_{\theta}) \right\},$$

understanding $\bar{P}_{\hat{b}}(\hat{a})$ to be this set's strict inequality counterpart.

Definition 5.6. *Given a and b , action \hat{a} is \hat{b} -Superior to action a if, no matter the risk-averse agent's utility function,*

1. $P_b(a) \subseteq P_{\hat{b}}(\hat{a})$ and
2. $\bar{P}_b(a) \subseteq \bar{P}_{\hat{b}}(\hat{a})$

i.e., the set of beliefs at which the DM prefers a to b is a subset of the beliefs at which the DM prefers \hat{a} to \hat{b} .

Next, we present a necessary condition for \hat{b} -superiority when $\hat{a}_1 < \hat{b}_1$. First note that $P_b(a) \subseteq P_{\hat{b}}(\hat{a})$ for all risk-averse utility functions, including risk neutrality. Then let $\bar{\mu}$ be the probability of $\theta = 1$ at which the risk-neutral DM is indifferent between a and b , and let $\hat{\mu}$ be the probability at which the DM is indifferent between \hat{a} and \hat{b} . Then, a necessary condition of \hat{b} -superiority is $\hat{\mu} \geq \bar{\mu}$, or

$$\frac{\hat{a}_0 - \hat{b}_0}{\hat{b}_1 - \hat{a}_1} \geq \frac{a_0 - b_0}{b_1 - a_1}. \quad (3)$$

In other words, for \hat{a} to be more preferred, it must be that the “advantage” of \hat{a} relative to \hat{b} is larger than the advantage of a relative to b .

There are many possible sufficient conditions for \hat{a} to be \hat{b} -superior to a , including \hat{a} dominating a or \hat{b} and \hat{b} being dominated by b or \hat{a} . Here is a less obvious sufficient condition.

Proposition 5.7. *If Inequality 3 holds, $a_1 \leq \hat{a}_1 < b_1 \leq \hat{b}_1$, and $\hat{b}_0 \leq b_0 < \hat{a}_0 \leq a_0$, then \hat{a} is \hat{b} -superior to action a .*

Proof. The proof is purely mechanical, so we relegate it to Appendix A.5. ■

Returning to the firm example, consider the conditions of Proposition 5.7. Firm 2 is initially better-matched to the consumer in state 1. By changing its product from b to \hat{b} , firm 2 is improving its already superior product in state 1, sacrificing the consumer's payoff in state 0. If firm 1 wishes to be chosen for at least as many beliefs, it too should concentrate more utility on state 1, even though consumers are relatively ill-matched to its product in that state and it sacrifices the payoff in state 0 ($\hat{a}_0 \leq a_0$). While this may seem counter intuitive, it is because maintaining market share relies on capturing the marginal consumer, not necessarily impressing those with extremely favorable beliefs (e.g., $\mu = 0$) more.

5.4. A Menu of Three

Our conception of superiority does not mean that \hat{a} must be preferred to a , only to b . This begs the following question. What conditions are equivalent to the DM choosing \hat{a} over both a and b ? When will our firm, for example, launch a product that both maintains market share and is preferred to the previous version, a ?

We say that action \hat{a} is (a, b) -superior if

$$\mathbb{E}_\mu u(a_\theta) \geq \mathbb{E}_\mu u(b_\theta) \Rightarrow \mathbb{E}_\mu u(\hat{a}_\theta) \geq \max_{c \in \{a, b\}} \mathbb{E}_\mu u(c_\theta),$$

for any strictly increasing, concave, and continuous u . Then,

Proposition 5.8. *The action \hat{a} is (a, b) -superior if and only if $\hat{a}_\theta \geq a_\theta$ for all $\theta \in \mathcal{A}$ and \hat{a} dominates a mixture of a and b .*

Proof. (\Leftarrow) We want to show that the set of beliefs at which \hat{a} is preferred to both a and b is a superset of $P_b(a)$, the set of beliefs at which a is preferred to b . Observe that the extreme points of $P_b(a)$, extreme $P_b(a)$, are the degenerate distributions $(\delta_\theta)_{\theta \in \mathcal{A} \cup \mathcal{C}}$ and the binary distributions at which the DM is indifferent between a

and b . By our assumptions, for any $\mu' \in \text{extreme } P_b(a)$,

$$\mathbb{E}_{\mu'} u(\hat{a}_\theta) \geq \max_{c \in \{a, b\}} \mathbb{E}_{\mu'} u(c_\theta).$$

By the linearity of the expectation, for any μ in the closed convex hull of extreme $P_b(a)$, the DM prefers \hat{a} to both a and b . By the Krein-Milman theorem, this set is precisely $P_b(a)$, as desired.

(\Rightarrow) This direction is clear. ■

Note that the only change from the conditions of Theorem 3.1 is that now we require $\hat{a}_\theta \geq a_\theta > b_\theta$ rather than $\hat{a}_\theta > b_\theta$ for all $\theta \in \mathcal{A}$. In other words, the new version must be better than a mix of the two previous options and be a payoff improvement for the consumer when they are well-matched with firm 1. This additional condition is clearly necessary. It is also sufficient because the existing conditions already guarantee that \hat{a} is better than b , so what remains is to guarantee that \hat{a} is also better than a .

5.5. A Hidden Alternative

As a final extension of our main model, suppose b is not known; or rather, the DM's state-dependent payoffs to b are not known. This could be the case if the firm does not know the exact payoffs of its rival's product, or of the consumer's outside option relative to its own offering. It is possible, nevertheless, to say something as follows.

Corollary 5.9. *Suppose action b is dominated by some \tilde{b} . Then, \hat{a} is b -superior to a if and only if \hat{a} dominates a mixture of a and \tilde{b} .*

We omit the proof as it is an easy consequence of Theorem 3.1. If we think of b as the DM's "outside option," this corollary tells us that we can formulate a meaningful, robust alteration of action a even when the outside option is not fully specified. All we need is an "upper bound" of sorts on the outside option. Put another way, if the firm knows the worst-case scenario for itself (the best \tilde{b}), then it is still able to introduce a new product which is preferred to any b .

6. A New (Partial) Order

We can also construct a novel partial order over actions in a decision problem, the “ b -superior order.” Fix a state-space Θ , an action b , and a set of actions A , where $b \in A$ and no action $a \in A \setminus \{b\}$ dominates b . Let $\hat{a} \succeq a$ denote that action \hat{a} is b -superior to action a . Then,

Proposition 6.1. \succeq is a partial order on a decision-maker’s set of actions $A \setminus \{b\}$.

Proof. **Reflexivity.** \succeq is obviously reflexive.

Antisymmetry. Suppose $a^1 \succeq a^2$ and $a^2 \succeq a^1$. Then there exist $\lambda_1, \lambda_2 \in [0, 1]$ such that

$$a_\theta^1 \geq \lambda_1 a_\theta^2 + (1 - \lambda_1) b_\theta \quad \forall \theta \in \Theta, \quad (4)$$

and

$$a_\theta^2 \geq \lambda_2 a_\theta^1 + (1 - \lambda_2) b_\theta \quad \forall \theta \in \Theta.$$

Suppose for the sake of contradiction that there exists a $\theta' \in \Theta$ for which $a_{\theta'}^1 > a_{\theta'}^2$. This implies that $\lambda_2 < 1$.

Combining the two inequalities, and doing a little algebra, we get that

$$a_\theta^1 (1 - \lambda_1 \lambda_2) \geq (1 - \lambda_1 \lambda_2) b_\theta \quad \forall \theta \in \Theta,$$

a contradiction as $1 - \lambda_1 \lambda_2 > 0$ and a^1 doesn’t dominate b .

Transitivity. Suppose $a^1 \succeq a^2$ and $a^2 \succeq a^3$. Then there exist $\lambda_1, \lambda_2 \in [0, 1]$ such that Inequality 4 and

$$a_\theta^2 \geq \lambda_2 a_\theta^3 + (1 - \lambda_2) b_\theta \quad \forall \theta \in \Theta$$

hold. Combining these, we get

$$a_\theta^1 \geq \lambda_1 \lambda_2 a_\theta^3 + (1 - \lambda_1 \lambda_2) b_\theta \quad \forall \theta \in \Theta,$$

so $a^1 \succeq a^3$, as desired. ■

We have come up with a partial order that is similar in spirit to that formulated by [Hart \(2011\)](#). Crucially, his concerns lotteries; ours, families of lotteries.

7. Information Acquisition

Up until this point we have focused on a simple decision problem in which a DM has a belief over the states of the world and chooses an optimal action based on this belief. It is easy to see that b -superiority extends to the scenario in which a DM obtains exogenous information before making a decision. That is, \hat{a} is b -superior to a if and only if for any $\mu \in \text{int } \Delta$ and any Bayes-plausible (martingale) distribution over posteriors F , the DM chooses \hat{a} with a higher probability than she chooses a (where, in both cases, b is chosen with the complimentary probability).

There has been a recent explosion of interest in problems with endogenous information acquisition. Suppose we want to know what the properties of \hat{a} are that make it chosen more versus b than a is when the DM flexibly acquires information before taking a choice.

We begin with the binary-state environment ($\Theta = \{0, 1\}$). We maintain our assumptions that neither a nor b dominates the other and that a is uniquely optimal in state 0 and b in state 1. Given a prior $\mu_0 \in (0, 1)$, and defining an agent's value function, in belief μ , as

$$V(\mu) := \max_{\tilde{a} \in \{a, b\}} \mathbb{E}_\mu u(\tilde{a}_\theta),$$

the DM's flexible information acquisition problem is

$$\max_{F \in \mathcal{F}(\mu_0)} \int_{\Delta} V(\mu) dF(\mu) - D(F), \quad (5)$$

where $\mathcal{F}(\mu_0)$ is the set of Bayes-plausible distributions given prior μ_0 and D is a uniformly posterior-separable (UPS) cost functional.^[4] Similarly, the value function when the menu is \hat{a} and b is $\hat{V}(\mu) := \max_{\tilde{a} \in \{\hat{a}, b\}} \mathbb{E}_\mu u(\tilde{a}_\theta)$, and when she has this menu the DM solves

$$\max_{\hat{F} \in \mathcal{F}(\mu_0)} \int_{\Delta} \hat{V}(\mu) d\hat{F}(\mu) - D(\hat{F}). \quad (6)$$

Any solution F^* (\hat{F}^*) to Program 5 (6) produces an optimal choice probability of action a (\hat{a}) by the DM, which we define to be $p_{D, \mu_0, u} \equiv p$ ($\hat{p}_{D, \mu_0, u} \equiv \hat{p}$).

^[4] See, e.g., [Caplin, Dean, and Leahy \(2022\)](#). $D: \Delta^2 \rightarrow \mathbb{R}$ is UPS if $D(F) = \int_{\Delta} c(\mu) dF(\mu) - c(\mu_0)$ for some strictly convex and twice continuously differentiable on $\text{int } \Delta$ function $c: \Delta \rightarrow \mathbb{R}$.

Definition 7.1. *\hat{a} is selected more than a if for any Bayes-plausible D , prior $\mu_0 \in \text{int } \Delta$, and strictly increasing, concave u ; for any optimal p , there exists an optimal choice probability $\hat{p} \geq p$.*

Proposition 7.2. *\hat{a} is selected more than a if and only if \hat{a} dominates a or b .*

Proof. A full proof may be found in Appendix A.6. ■

Most of the work in proving the proposition is in the necessity portion. First consider what happens if \hat{a} is not b -superior to a . Then, we can find a u such that $\hat{p} < \bar{p}$. Take such a u , in which case we can then always find a cost function such that a is selected with probability 1 when the menu is $\{a, b\}$ but when the menu is $\{\hat{a}, b\}$, either the DM learns in a nontrivial manner, and hence selects \hat{a} with probability strictly less than 1, or does not learn but selects b with probability 1.

To understand why \hat{a} must dominate a or b , consider how the DM chooses to learn when there are two states and two actions. The DM chooses an F^* such that there are two posteriors, $\underline{\mu}_a$ and $\bar{\mu}_a \geq \underline{\mu}_a$ such that she chooses a with probability 1 if $\mu_0 \leq \underline{\mu}_a$, b with probability 1 if $\mu_0 \geq \bar{\mu}_a$, and a with probability $p \in (0, 1)$ if $\mu_0 \in (\underline{\mu}_a, \bar{\mu}_a)$.

Next, take \hat{a} to be such that $\ell_{\hat{a}}$ is a counter-clockwise rotation from ℓ_a towards ℓ_b (recall Figure 1). This is equivalent, however, to an increase in learning cost (if payoffs had remained the same) for a risk-neutral DM. A higher marginal cost of information leads to less learning; i.e., $(\underline{\mu}_{\hat{a}}, \bar{\mu}_{\hat{a}}) \subseteq (\underline{\mu}_a, \bar{\mu}_a)$. This means that we can find a prior belief such that $\mu_0 \in (\underline{\mu}_{\hat{a}}, \bar{\mu}_{\hat{a}})$ but $\mu_0 \geq \bar{\mu}_{\hat{a}}$. In other words, action a has a positive probability of being chosen when the menu is $\{a, b\}$, but action \hat{a} will never be chosen when the menu is $\{\hat{a}, b\}$. Hence, b -superiority is clearly not enough to guarantee that \hat{a} is chosen more than a .

7.1. Three or More States

Alas, when there are three states, \hat{a} dominating a no longer implies that \hat{a} is selected more than a , as we now illustrate. Let $\Theta = \{0, 1, 2\}$ with $\mu^2 := \mathbb{P}(2)$ and $\mu^1 := \mathbb{P}(1)$. Let $\hat{\alpha}_2 = 3/2 = \alpha_2 + 1/2$, $\alpha_1 = -1 = \hat{\alpha}_1$ and $\hat{\alpha}_0 = \alpha_0 = 0$; and $\beta_\theta = 0$ for all $\theta \in \Theta$. Evidently, \hat{a} weakly dominates a (and $\hat{a} \neq a$).

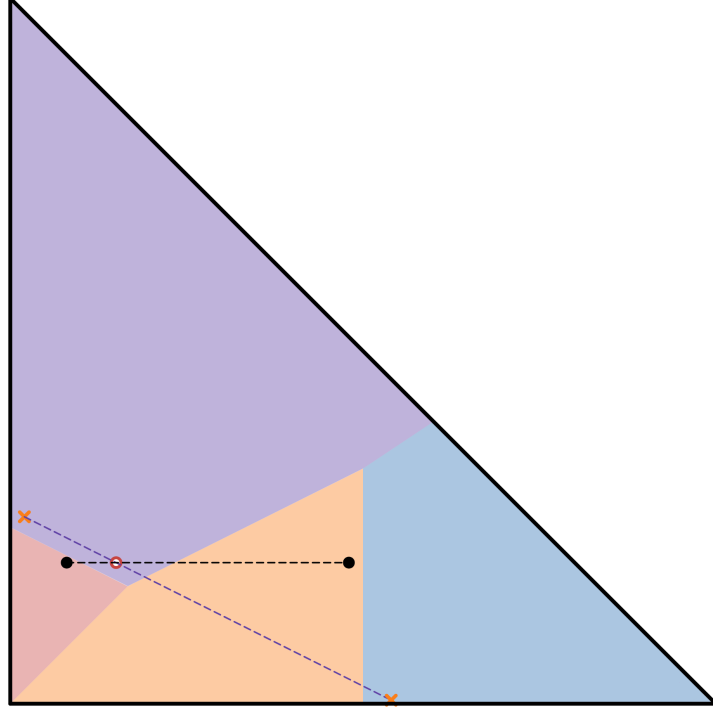


Figure 2: The insufficiency of dominance for three or more states

We have

$$V(\mu^1, \mu^2) = \max \left\{ \mu^2 - \mu^1, 0 \right\}, \quad \text{and} \quad \hat{V}(\mu^1, \mu^2) = \max \left\{ \frac{3}{2}\mu^2 - \mu^1, 0 \right\}.$$

We then define

$$W(\mu^1, \mu^2) := \max \left\{ V, \hat{V} + \frac{1}{4}\mu^1 - \frac{1}{8} \right\},$$

then, appealing to Lemma A.1 in [Whitmeyer \(2023\)](#), conclude that there exists a UPS cost such that the four points

$$\left\{ \left(\frac{2}{25}, \frac{1}{5} \right), \left(\frac{12}{25}, \frac{1}{5} \right), \left(\frac{1}{50}, \frac{53}{200} \right), \left(\frac{27}{50}, \frac{1}{200} \right) \right\}$$

are the support of optimal learning given value function W . Accordingly, we take such a cost function and fix prior $(3/20, 1/5)$ then compute $p = 33/40 > 3/4 = \hat{p}$.^[5]

Figure 2 illustrates this example on the 2-simplex. The four colored polygons are the regions in which V lies above the rotated value function \hat{V} and either a is

^[5] This calculation simply applies Bayes plausibility to the given posteriors. Specifically, $3/20 = 2/25p + 12/25(1 - p)$.

optimal (red) or b is optimal (orange); and the rotated value function \hat{V} lies above V and either \hat{a} is optimal (purple) or b is optimal (blue). The prior is the hollow red dot. The black dots are the support of F^* , and the orange x s are the support of \hat{F}^* .

Leaving the single-dimensional environment corresponding to the two-state simplex engenders this result. What we are doing is taking a value function that in a sense “twist things,” making it so that the learning in the two problems is not along the same line segment. There is enough freedom then to pick an appropriate prior so that $p > \hat{p}$.

References

- Susan Athey. Monotone comparative statics under uncertainty. *The Quarterly Journal of Economics*, 117(1):187–223, 2002.
- Robert J Aumann and Roberto Serrano. An economic index of riskiness. *Journal of Political Economy*, 116(5):810–836, 2008.
- Turan G Bali, Nusret Cakici, and Fousseni Chabi-Yo. A generalized measure of riskiness. *Management science*, 57(8):1406–1423, 2011.
- Andrew Caplin, Mark Dean, and John Leahy. Rationally inattentive behavior: Characterizing and generalizing shannon entropy. *Journal of Political Economy*, 130(6):1676–1715, 2022.
- Michael Choi and Lones Smith. Ordinal aggregation results via karlin’s variation diminishing property. *Journal of Economic theory*, 168:1–11, 2017.
- Gregorio Curello and Ludvig Sinander. The preference lattice. *Mimeo*, 2019.
- Eddie Dekel. An axiomatic characterization of preferences under uncertainty: Weakening the independence axiom. *Journal of Economic theory*, 40(2):304–318, 1986.

- Aaron S Edlin and Chris Shannon. Strict single crossing and the strict spence-mirrlees condition: a comment on monotone comparative statics. *Econometrica*, 66(6):1417–1425, 1998.
- Dean P Foster and Sergiu Hart. An operational measure of riskiness. *Journal of Political Economy*, 117(5):785–814, 2009.
- John S Hammond III. Simplifying the choice between uncertain prospects where preference is nonlinear. *Management Science*, 20(7):1047–1072, 1974.
- Sergiu Hart. Comparing risks by acceptance and rejection. *Journal of Political Economy*, 119(4):617–638, 2011.
- Chew Soo Hong. A generalization of the quasilinear mean with applications to the measurement of income inequality and decision theory resolving the allais paradox. *Econometrica*, 51(4):1065–1092, 1983.
- Dmitri Panov (https://mathoverflow.net/users/943/dmitri_panov). Existence of a strictly convex function interpolating given gradients and values. MathOverflow, 2019. URL <https://mathoverflow.net/q/327966>. URL:<https://mathoverflow.net/q/327966> (version: 2019-04-14).
- Ian Jewitt. Risk aversion and the choice between risky prospects: the preservation of comparative statics results. *The Review of Economic Studies*, 54(1):73–85, 1987.
- Ian Jewitt. Choosing between risky prospects: the characterization of comparative statics results, and location independent risk. *Management Science*, 35(1):60–70, 1989.
- Samuel Karlin and Albert Novikoff. Generalized convex inequalities. *Pacific Journal of Mathematics*, 13(4), 1963.
- Navin Kartik, SangMok Lee, and Daniel Rappoport. Single-crossing differences in convex environments. *Review of Economic Studies*, page rdad103, 2023.

- Peter J Lambert and John D Hey. Attitudes to risk. *Economics Letters*, 2(3): 215–218, 1979.
- Fabio Maccheroni, Massimo Marinacci, Ruodu Wang, and Qinyu Wu. Risk aversion and insurance propensity. *Mimeo*, 2023.
- Filip Matějka and Alisdair McKay. Rational inattention to discrete choices: A new foundation for the multinomial logit model. *American Economic Review*, 105(1): 272–298, 2015.
- Paul Milgrom and Chris Shannon. Monotone comparative statics. *Econometrica*, pages 157–180, 1994.
- Michele Muller-Itten, Roc Armenter, and Zachary Stangebye. Rational inattention via ignorance equivalence. *Mimeo*, 2021.
- Marilyn Pease and Mark Whitmeyer. Safety, in numbers. *Mimeo*, 2023.
- Robert R Phelps. *Convex functions, monotone operators and differentiability*, volume 1364. Springer, 2009.
- Luciano Pomatto, Philipp Strack, and Omer Tamuz. Stochastic dominance under independent noise. *Journal of Political Economy*, 128(5):1877–1900, 2020.
- John K-H Quah and Bruno Strulovici. Aggregating the single crossing property. *Econometrica*, 80(5):2333–2348, 2012.
- Frank Riedel and Tobias Hellmann. The foster–hart measure of riskiness for general gambles. *Theoretical Economics*, 10(1):1–9, 2015.
- Michael Rothschild and Joseph E Stiglitz. Increasing risk: I. a definition. *Journal of Economic Theory*, 2(3):225–243, 1970.
- Mark Whitmeyer. Making information more valuable. *Mimeo*, 2023.

A. Omitted Proofs

A.1. Lemma 3.2 Proof

Proof. The first two conditions are obviously necessary for b -superiority. We assume they hold and suppose for the sake of contraposition that \hat{a} does not pairwise-dominate a collection of mixtures of a and b . This means that there exists at least one pair $(\theta, \theta') \in \mathcal{A} \times \mathcal{B}$ for which there exists no $\lambda_{\theta, \theta'} \in [0, 1]$ such that the two inequalities in Expression 1 hold.

Fix such a pair of θ and θ' . Given utility u , let $\bar{\mu}_u$ solve

$$(1 - \bar{\mu}_u) u(b_\theta) + \bar{\mu}_u u(b_{\theta'}) = (1 - \bar{\mu}_u) u(a_\theta) + \bar{\mu}_u u(a_{\theta'}),$$

and $\hat{\mu}_u$ solve

$$(1 - \hat{\mu}_u) u(b_\theta) + \hat{\mu}_u u(b_{\theta'}) = (1 - \hat{\mu}_u) u(\hat{a}_\theta) + \hat{\mu}_u u(\hat{a}_{\theta'});$$

viz., these are the beliefs on the edge of Δ between the the degenerate probability distributions on states θ and θ' , δ_θ and $\delta_{\theta'}$, at which the DM with utility u is indifferent between a and b and \hat{a} and b , respectively. If the DM is risk-neutral, we drop the subscript and simply write $\bar{\mu}$ and $\hat{\mu}$.

By assumption $\bar{\mu}, \hat{\mu} \in (0, 1)$. Moreover, observe that a necessary condition for b -superiority is $\hat{\mu} \geq \bar{\mu}$, which holds if and only if

$$(b_{\theta'} - a_{\theta'}) (\hat{a}_\theta - b_\theta) \geq (a_\theta - b_\theta) (b_{\theta'} - \hat{a}_{\theta'}). \quad (2)$$

If this inequality does not hold, we are done, as we have produced the right negation. Accordingly, suppose it does hold. This means, given our other assumptions,

$$b_\theta < a_\theta < \hat{a}_\theta, \quad \text{and} \quad \hat{a}_{\theta'} < a_{\theta'} < b_{\theta'}.$$

Our goal is to show that $\bar{\mu}_u > \hat{\mu}_u$ for some strictly-increasing, continuous, and concave u . Let us specify a particular such u . Let

$$u(x) = \begin{cases} x, & \text{if } x \leq a_{\theta'}, \\ \iota x + (1 - \iota) a_{\theta'}, & \text{if } x > a_{\theta'}. \end{cases}$$

for some $\iota \in (0, 1)$, so $\bar{\mu}_u - \hat{\mu}_u$ equals

$$\frac{u(a_\theta) - u(b_\theta)}{u(a_\theta) - u(b_\theta) + \iota(b_{\theta'} - a_{\theta'})} - \frac{u(\hat{a}_\theta) - u(b_\theta)}{u(\hat{a}_\theta) - u(b_\theta) + \iota(b_{\theta'} - a_{\theta'}) + a_{\theta'} - \hat{a}_{\theta'}}.$$

If $b_\theta \geq a_{\theta'}$, this equals

$$\frac{a_\theta - b_\theta}{a_\theta - b_\theta + b_{\theta'} - a_{\theta'}} - 0 > 0,$$

when $\iota = 0$. If $b_\theta < a_{\theta'}$, this equals

$$1 - \frac{u(\hat{a}_\theta) - b_\theta}{u(\hat{a}_\theta) - b_\theta + a_{\theta'} - \hat{a}_{\theta'}} > 0,$$

when $\iota = 0$. ■

A.2. Proposition 3.6 Proof

Proof. (\Leftarrow) This direction is straightforward.

(\Rightarrow) By Theorem 3.1, it is necessary that \hat{a} dominates a mixture of a and b . Thus, it suffices to suppose for the sake of contraposition that \hat{a} dominates a mixture of a and b but dominates neither a nor b . This means that there exists some pair $(\theta, \theta') \in \mathcal{A} \times \mathcal{B}$ for which

$$b_\theta < \hat{a}_\theta < a_\theta, \quad \text{and} \quad a_{\theta'} < \hat{a}_{\theta'} < b_{\theta'}.$$

Our goal is to show that $\bar{\mu}_u > \hat{\mu}_u$ for some strictly-increasing, continuous u . Take

$$u(x) = \begin{cases} \iota x + (1 - \iota)\hat{a}_\theta, & \text{if } x \leq \hat{a}_\theta, \\ x, & \text{if } x > \hat{a}_\theta, \end{cases}$$

for some $\iota \in (0, 1)$. Thus, $\bar{\mu}_u - \hat{\mu}_u$ equals

$$\frac{a_\theta - \hat{a}_\theta - \iota(b_\theta - \hat{a}_\theta)}{a_\theta - \hat{a}_\theta - \iota(b_\theta - \hat{a}_\theta) + u(b_{\theta'}) - u(a_{\theta'})} - \frac{\iota(\hat{a}_\theta - b_\theta)}{\iota(\hat{a}_\theta - b_\theta) + u(b_{\theta'}) - u(\hat{a}_{\theta'})}.$$

If $b_{\theta'} \leq \hat{a}_\theta$, this equals

$$1 - \frac{\hat{a}_\theta - b_\theta}{\hat{a}_\theta - b_\theta + b_{\theta'} - \hat{a}_{\theta'}} > 0,$$

when $\iota = 0$. If $b_{\theta'} > \hat{a}_\theta$, it equals

$$\frac{a_\theta - \hat{a}_\theta}{a_\theta - \hat{a}_\theta + b_{\theta'} - u(a_{\theta'})} - 0 > 0,$$

when $\iota = 0$. ■

A.3. Proposition 5.2 Proof

Proof. (\Leftarrow) Let $a = a^1$ (the proof for $a = a^m$ is identical). Observe that if $\hat{a}_0^1 > a_0^2$, $\hat{a}_0^1 > a_0^j$ for all $j \in \{3, \dots, m\}$. By the monotonicity of u the indifference points between the successive actions are increasing, i.e., the indifference point between \hat{a}^1 and a^2 —if it exists—is (weakly) to the left of that between a^2 and a^3 , which is strictly to the left of that between a^3 and a^4 , etc.^[6] If there is no indifference point between \hat{a}^1 and a^2 , a^2 must be strictly dominated as a result of the transformation. But then it is obvious that the indifference point between \hat{a}^1 and the lowest undominated a^j ($j \in \{3, \dots, m\}$) must be to the right of the indifference point between a^1 and a^2 . We conclude that $P_B(a) \subseteq P_B(\hat{a})$ if $P_b(a) \subseteq P_b(\hat{a})$.

If $a \neq a^1, a^m$ it is immediate that \hat{a} dominating a implies $P_B(a) \subseteq P_B(\hat{a})$.

(\Rightarrow) The first necessity statement is implied by Theorem 3.1. For the second statement, Theorem 3.1 tells us that if $\hat{a} = \hat{a}^i$ is b -superior to both a^{i-1} and a^{i+1} , $\hat{a}_1^i \geq a_1^i$ and $\hat{a}_0^i \geq a_0^i$. ■

A.4. Corollary 5.5 Proof

Proof. (\Rightarrow) Easy: we can just take W to be the degenerate random variable on 0, then appeal to Theorem 3.1.

(\Leftarrow) The exact sufficiency argument for Theorem 3.1 goes through, provided we redefine the lotteries L_a , L_b , and $L_{\hat{a}}$ as follows.

Let Y_a be the random variable yielding a_θ with probability $\mu(\theta)$, and L_a the sum of Y_a and W ; Y_b , the random variable yielding b_θ with probability $\mu(\theta)$, and L_b the sum of Y_b and W ; and $Y_{\hat{a}}$, the random variable yielding \hat{a}_θ with probability $\mu(\theta)$, and $L_{\hat{a}}$ the sum of $Y_{\hat{a}}$ and W .

Suppose $\mu \in \Delta$ and W are such that the DM prefers a to b , i.e., $L_a \succeq L_b$. By independence, for any $\lambda \in [0, 1]$,

$$L^\lambda := \lambda L_a + (1 - \lambda) L_b \succeq L_b.$$

^[6] For any strictly increasing u the slopes of the lines $a_0^i(1 - \mu) + a_1^i\mu$ are strictly increasing in i , which implies this observation.

Moreover, by assumption there is some $\lambda^* \in [0, 1]$ for which $\lambda^* a_\theta + (1 - \lambda^*) b_\theta = \hat{a}_\theta$ for all $\theta \in \Theta$. This implies that $\lambda^* Y_a + (1 - \lambda^*) Y_b$ is a mean preserving spread of $Y_{\hat{a}}$. Moreover, as second-order stochastic dominance is preserved under independent noise,^[7] $L^{\lambda^*} = \lambda^* Y_a + (1 - \lambda^*) Y_b + W$ is a mean preserving spread of $L_{\hat{a}} = Y_{\hat{a}} + W$. Consequently, as the DM is risk averse, $L_{\hat{a}} \succeq L^{\lambda^*} \succeq L_b$. ■

A.5. Proposition 5.7 Proof

Proof. Directly,

$$\begin{aligned}
\hat{\mu}_u &= \frac{u(\hat{a}_0) - u(\hat{b}_0)}{u(\hat{a}_0) - u(\hat{b}_0) + u(\hat{b}_1) - u(\hat{a}_1)} = \frac{\frac{u(\hat{a}_0) - u(\hat{b}_0)}{\hat{a}_0 - \hat{b}_0}}{\frac{u(\hat{a}_0) - u(\hat{b}_0)}{\hat{a}_0 - \hat{b}_0} + \frac{u(\hat{b}_1) - u(\hat{a}_1)}{\hat{a}_0 - \hat{b}_0}} \\
&\geq \frac{\frac{u(a_0) - u(b_0)}{a_0 - b_0}}{\frac{u(a_0) - u(b_0)}{a_0 - b_0} + \frac{u(\hat{b}_1) - u(\hat{a}_1)}{\hat{a}_0 - \hat{b}_0}} \\
&= \frac{u(a_0) - u(b_0)}{u(a_0) - u(b_0) + (a_0 - b_0) \frac{u(\hat{b}_1) - u(\hat{a}_1)}{\hat{a}_0 - \hat{b}_0}} \\
&\geq \frac{u(a_0) - u(b_0)}{u(a_0) - u(b_0) + (b_1 - a_1) \frac{u(\hat{b}_1) - u(\hat{a}_1)}{\hat{b}_1 - \hat{a}_1}} \\
&\geq \frac{u(a_0) - u(b_0)}{u(a_0) - u(b_0) + u(b_1) - u(a_1)} = \bar{\mu}_u,
\end{aligned}$$

where the first and third inequalities follow from the Three-chord lemma (Theorem 1.16 in Phelps (2009)), and the second inequality from Inequality 3. ■

A.6. Proposition 7.2 Proof

Proof. (\Leftarrow) If \hat{a} dominates b , $\hat{p} = 1$ is a solution. Now let \hat{a} not dominate b but dominate a . We denote $\alpha_\theta \equiv u(a_\theta) \leq u(\hat{a}_\theta) \equiv \hat{\alpha}_\theta$ for all $\theta \in \{0, 1\}$. Without loss of generality we normalize $u(b_\theta) = 0$ for all $\theta \in \{0, 1\}$.

Suppose first that μ_0 and c are such that $p = 1$. This is equivalent to $\mu_0 \leq \bar{\mu}_u$ and

$$(\alpha_1 - \alpha_0 - c'(\mu_0)) \mu + c'(\mu_0) \mu_0 + \alpha_0 \geq 0,$$

^[7] See, for instance, the discussion on page 3 of Pomatto, Strack, and Tamuz (2020).

for all $\mu \in [0, 1]$. For any $\mu \in [0, 1]$ this expression is strictly increasing in both α_0 and α_1 , so we have

$$(\hat{\alpha}_1 - \hat{\alpha}_0 - c'(\mu_0))\mu + c'(\mu_0)\mu_0 + \hat{\alpha}_0 \geq 0,$$

for all $\mu \in [0, 1]$, i.e., $\hat{p} = 1$.

Now suppose that μ_0 is such that $p \in (0, 1)$. In this case, the support of F^* , x_L and x_H , solves

$$\alpha_1 - \alpha_0 - c'(x_L) + c'(x_H) = 0,$$

and

$$c'(x_L)x_L - c'(x_H)x_H + c(x_H) + \alpha_0 - c(x_L) = 0.$$

Appealing to the implicit function theorem, we obtain

$$x'_L(\alpha_1) = \frac{x_H}{c''(x_L)(x_H - x_L)} > 0, \quad \text{and} \quad x'_H(\alpha_1) = \frac{x_L}{c''(x_H)(x_H - x_L)} > 0.$$

Likewise,

$$x'_L(\alpha_1) = \frac{1 - x_H}{c''(x_L)(x_H - x_L)} > 0, \quad \text{and} \quad x'_H(\alpha_1) = \frac{1 - x_L}{c''(x_H)(x_H - x_L)} > 0.$$

Evidently, p is strictly increasing in both x_H and x_L , so $\hat{p} \geq p$.

(\Rightarrow) Suppose for the sake of contraposition that \hat{a} dominates neither a nor b . If \hat{a} is dominated (and does not dominate) by a or b , the outcome is trivial. Accordingly, suppose \hat{a} is dominated by neither. There are three possibilities: either i. $\hat{a}_0 > a_0 > b_0$ and $\hat{a}_1 < a_1 < b_1$; or ii. $a_0 > \hat{a}_0 > b_0$ and $b_1 > \hat{a}_1 > a_1$; or iii. $a_0 > b_0 > \hat{a}_0$ and $a_1 < b_1 < \hat{a}_1$.

Case iii is immediate. Let the DM be risk neutral: then there exists a $\mu' \in (0, 1)$, such that for any belief $\mu < \mu'$, it is uniquely optimal for the DM to take action a when her menu is $\{a, b\}$ and b when her menu is $\{\hat{a}, b\}$. Then, one need only pick a sufficiently convex c —that such a convex c can always be found is an implication of Lemma A.1 in Whitmeyer (2023)^[8]—such that for prior μ_0 , no learning is uniquely

^[8] This is not strictly true as Whitmeyer (2023) does not impose that the cost function is twice continuously differentiable, merely strictly convex. However, it is easy to extend that result to smooth functions: see, e.g. (<https://mathoverflow.net/users/943/dmitri-panov>). Alternatively, one could remove the twice-continuously differentiable specification at the expense of not being able to appeal to the implicit function theorem in the sufficiency portion of the proof.

optimal in both problems, in which case $p = 1 > 0 = \hat{p}$.

Case i is also easy. Observe that \hat{a} is not b -superior to a , so there exists a strictly increasing concave u for which $0 < \hat{\mu}_u < \bar{\mu}_u < 1$. We maintain the convention $\alpha_\theta \equiv u(a_\theta)$ and $\hat{\alpha}_\theta \equiv u(\hat{a}_\theta)$ for all $\theta \in \{0, 1\}$ and also introduce the notation $\beta_\theta \equiv u(b_\theta)$ for all $\theta \in \{0, 1\}$.

We tweak the notation

$$\ell_a = \mu\alpha_1 + (1 - \mu)\alpha_0, \ell_{\hat{a}} = \mu\hat{\alpha}_1 + (1 - \mu)\hat{\alpha}_0, \text{ and } \ell_b = \mu\beta_1 + (1 - \mu)\beta_0,$$

and define $W(\mu) := \max\{\ell_a, \ell_{\hat{a}}, \ell_b\}$. We let $\tilde{\mu}$ denote the intersection of ℓ_a and $\ell_{\hat{a}}$, and observe that $0 < \tilde{\mu} < \bar{\mu}$; this holds because $\ell_{\hat{a}}$ has a steeper slope and a strictly larger y -intercept than ℓ_a . Then, Lemma A.1 in [Whitmeyer \(2023\)](#) implies that for any triple $\mu_1 \in (0, \tilde{\mu})$, $\mu_2 \in (\tilde{\mu}, \bar{\mu})$, and $\mu_3 \in (\bar{\mu}, 1)$, there exists a UPS cost such that when the DM's value function is W , any optimal learning has support on the three specified points. Accordingly, for such a cost function, when $\mu_0 = \mu_2$, $p = 1$ and $\hat{p} < 1$.

Finally, case ii: the argument from the previous paragraph allows us to assume that \hat{a} is b -superior to a , or else we are done. Consequently, $\bar{\mu}_u \leq \hat{\mu}_u$ for all permissible u . Fix such a u and normalize payoffs so that $\beta_\theta = 0$ for all $\theta \in \{0, 1\}$ (this is without loss of generality, as u has been fixed). Now take a line $f(\mu) := -\gamma\mu + \delta$, where $0 < \delta < \alpha_0 - \hat{\alpha}_0$, $\gamma > \delta$, and

$$\frac{\delta}{\gamma} > \hat{\mu}_u = \frac{\hat{\alpha}_0}{\hat{\alpha}_0 - \hat{\alpha}_1}.$$

Defining $T(\mu) := \max\{\ell_a, \max\{\ell_{\hat{a}}, 0\} + f, 0\}$, we note that this piecewise-affine curve has three kink points. First, at some $\mu_1 \in (0, 1)$, where ℓ_a and $\ell_{\hat{a}} + f$ intersect. Second, at some $\mu_2 \in (\mu_1, 1)$, where 0 and $\ell_{\hat{a}}$ intersect. Third, at some $\mu_3 \in (\mu_2, 1)$ where f and 0 intersect. Again appealing to Lemma A.1 in [Whitmeyer \(2023\)](#), taking a prior $\mu_0 \in (\mu_2, \mu_3)$ we note the existence of a UPS cost producing $p > 0$ and $\hat{p} = 0$. ■

B. Sufficiency for Other Preferences

The sufficiency portion of Theorem 3.1 continues to hold for a DM with preferences other than expected utility: we say a preference \succ over lotteries satisfies **Betweenness** (Hong (1983), Dekel (1986)) if

$$P \succ Q \quad \Rightarrow \quad P \succ \alpha P + (1 - \alpha)Q \succ Q \text{ for all } \alpha \in (0, 1).$$

A preference over lotteries is **Strongly risk averse** (Maccheroni, Marinacci, Wang, and Wu (2023)) if

$$Q \text{ is a mean-preserving spread of } P \quad \Rightarrow \quad P \succeq Q.$$

Corollary B.1. *Suppose the DM's preferences satisfy betweenness and strong risk aversion. Then, if $\hat{a}_\theta > b_\theta$ for all $\theta \in \mathcal{A}$ and \hat{a} dominates a mixture of a and b , \hat{a} is b -superior to action a .*