Correlation made simple: Applications to Salience and Regret Theory*

Giacomo Lanzani
MIT

Abstract

In this work, we offer a straightforward axiomatization for decision criterion under uncertainty in which the correlation between the alternatives considered plays a role. Extending to the nondeterministic case the techniques of conjoint measurement developed by Fishburn (1991) we can formally identify Transitivity as the vN-M axiom that has to be relaxed to allow for these richer patterns of behavior that include Regret Theory as a particular case.

To illustrate the advantages of our modeling choice, we provide a simple axiomatization for the Salience Theory model of choice (Bordalo, Gennaioli, and Shleifer, 2013) within our general framework. This approach allows us to single out Ordering as the property that brings Salience Theory outside the Expected Utility.

Finally, since our decision criterion is intransitive, the existence of an optimal choice in a set of three or more alternatives is not guaranteed. However, we prove the existence of optimal randomization over the possible options.

1 Introduction

The goal of this paper is to provide a simple set of axioms for a general class of risk preferences such that the correlation between the alternatives under consideration is relevant for the Decision Maker (henceforth DM). The motivation is two-fold. First, we aim to provide a better understanding of the difference between classical model under which correlation is relevant (see Bell, 1982, Loomes and Sugden, 1982, Fishburn, 1989), and the benchmark model for choice under risk, Expected Utility (henceforth EU). Second, we show that the general framework nest the recent wave of models that highlight the role of correlation (see, e.g., Bordalo, Gennaioli, and Shleifer, 2012, henceforth BGS, and Koszegi and Szeidl, 

---

*I am particularly indebted with Simone Cerreia-Vioglio, Roberto Corrao, Drew Fudenberg, Nicola Gennaioli, and Fabio Maccheroni. I am also grateful to Colin Camerer, Salvatore Greco, Massimo Marinacci, Paulo Natenzon, Pietro Ortoleva, Harry Pei, Antonio Rangel, Aldo Rustichini, Philipp Strack, Peter Wakker, Alexander Woltzky, and Muhamet Yildiz for helpful comments and suggestions. The financial support of Guido Cazzavillan Scholarship is gratefully acknowledged.
2006) and to better understand where do they depart from the classical set of models that is usually identified with Regret Theory.

We accomplish these goals by taking a completely different route than the one followed in the usual axiomatizations of Regret Theory, (see Fishburn 1989, Sugden 1993, and Diecidue and Sundaram 2017). All these papers represent the preferences as a binary relation over acts a la Savage. However, the technical assumptions on the richness of the state space partially obfuscate the comparison with the Subjective Expected Utility model.¹

Instead, we abstract away from the state space formulation, representing the preferences of the DM in the space of lotteries. However, when the correlation between alternatives matters, the use of a binary relation over lotteries is not sufficiently rich as a modelling tool. We elaborate on this with an example. Suppose that we have the two lotteries $p_1 = (10, \frac{1}{3}; 5, \frac{1}{3}; 0, \frac{1}{3})$ and $p_2 = (10, \frac{1}{3}; 4, \frac{1}{3}; 1, \frac{1}{3})$, and consider the following two possible correlation structures:

<table>
<thead>
<tr>
<th>$p$</th>
<th>1</th>
<th>4</th>
<th>10</th>
</tr>
</thead>
<tbody>
<tr>
<td>0</td>
<td>0</td>
<td>1/3</td>
<td>0</td>
</tr>
<tr>
<td>5</td>
<td>0</td>
<td>0</td>
<td>1/3</td>
</tr>
<tr>
<td>10</td>
<td>1/3</td>
<td>0</td>
<td>0</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>$p'$</th>
<th>1</th>
<th>4</th>
<th>10</th>
</tr>
</thead>
<tbody>
<tr>
<td>0</td>
<td>0</td>
<td>0</td>
<td>1/3</td>
</tr>
<tr>
<td>5</td>
<td>1/3</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>10</td>
<td>0</td>
<td>1/3</td>
<td>0</td>
</tr>
</tbody>
</table>

Both the joint distributions $p$ and $p'$ feature $(p_1, p_2)$ as their marginal distributions, but we will see that it is well possible that a Salience sensitive DM strictly prefers $p_1$ under the first correlation structure (driven by the salient realization $(10, 1)$) and $p_2$ under the second correlation structure (driven by the salient realization $(0, 10)$). Therefore, it is not possible to express the DM tastes using a binary relation on lotteries over monetary outcomes, since $p_1$ and $p_2$ cannot be ranked without additional information about their joint distribution. Instead of working with binary relations over alternatives, we apply the concept of preference set, introduced by Fishburn (1991).² Precisely, given a set of possible prizes $X$, tastes are represented by a preference set $P \subseteq \Delta (X \times X)$ with the following interpretation. The DM contemplates a joint distribution $p$ over $X \times X$, and he has to decide if, given the marginals and the correlation structure, he prefers to be paid accordingly to the first component or the second. Then, $p$ belongs to the preference set $P$ if and only if the DM prefers to be paid accordingly to the first component. In our previous example, we have $p \in P$, but $p' \notin P$.

Two reasons motivate our modelling choice. First, as we argued above, and as it will be clear momentarily, this approach abstract away from the ancillary state-space assumptions

¹A notable exception is Diecidue and Sundaram (2017), who improve on the previous representation providing an axiomatization that delivers a continuous regret function on an arbitrary finite state space. However, to deal with acts, they need to use a stronger version of Independence, as well a richness assumption on the space of outcomes.

²Fishburn (1991) introduces the concept of preference sets for intransitive preferences over multiatribute products. To the best of our knowledge, this is the first time preference sets are used to axiomatize preferences under risk.
to obtain a clear comparison with Expected Utility. The second motivation comes from our Salience Theory application. The experimental evidence and the examples used to justify the Salience model in BGS and the subsequent experimental papers consider choices between lotteries, where the only state space is the one defined as the space of all the possible realizations of the two lotteries under scrutiny.\(^3\) Therefore, axioms stated in terms of joint lotteries are both easier to map into the BGS model and can be directly challenged by the existing experimental evidence on the model. Finally, under the alternative state-space formulation, the characterization of the Salience properties postulated by BGS is less clean, and they are demanding in terms of the structural properties of the underlying state space. Among other things, the state space has to be atomless, a property at odds with the use of the small finite set of joint realizations as the state space in BGS.\(^4\)

Finally, if we want to test the theory in the lab, having an axiomatization for the case of choice under risk, instead of one for acts defined over a state space in which probabilities are not specified, allow us to disentangle violation of the axioms at the cornerstone of Regret Theory from failures in formulating a unique, coherent probability measure over the states of the world. Given the prevalence of these failures highlighted by the ambiguity literature, this is a real concern.

We identify the three axioms on the preference set \(P\) that are needed to obtain a representation weaker than the Expected Utility one, and that allows for Regret and Salience sensitivity. These axioms are Completeness, Strong Independence, and Archimedean Continuity and they are one to one with the representation

\[
p \in P \iff \sum_{x,y} \phi(x,y) p(x,y) \geq 0
\]

where \(\phi\) is a skew-symmetric functional. The following result justifies the names we gave to the axioms: if Transitivity is assumed on top of them, the representation reduces to the EU one, and the preferences satisfy the classical vN-M axioms.

After having weakened the axioms to allow for this more general behavior, we axiomatize the additional psychological properties of Salience detection considered by BGS: Ordering, Diminishing Sensitivity, and Reflexivity. We find that Ordering is the property that brings the Salience criterion outside the realm of EU or Prospect Theory, and we characterize it as a submodularity condition on \(\phi\). Instead, Diminishing Sensitivity and Reflexivity combined amount to the usual risk-aversion in gains, risk-loving in loss property featured by Prospect Theory.

We also provide a partial solution to the problem of choice between multiple alternatives. A DM characterized by the nontransitive representation of preferences in (1) may not


\(^4\)On a more technical note, notice that the use of an atomless state space is particularly unsatisfactory since it is a direct consequence of what Fishburn (1991) calls axiom P6*, an axiom that is not necessary for the representation. Therefore, such a richness of the space is not an intrinsic feature of the model, but more the result of a technically convenient assumption.
have an alternative that is weakly preferred to all the others when facing a set of at least three options. However, using the min-max theorem, we prove the existence of optimal randomization over alternatives.

**Related Literature**  Diecidue and Sundaram (2017), improve on the previous representation of Regret models providing an axiomatization that delivers a continuous regret function on an arbitrary finite state space. However, to deal with acts, they need to use a stronger version of Independence, as well a richness assumption on the space of outcomes.

To the best of our knowledge, this paper is the first attempt to axiomatize Salience Theory of choice under risk. Ellis and Masatioglu (2019) proposed an axiomatization of the Salience Theory of consumption (see Bordalo, Gennaioli, and Shleifer, 2013). Apart from the different subject of analysis, our paper also differs from how salience affects choice. In their work, it partitions the space of alternatives in several regions, and options in different regions are evaluated accordingly to a different criterion. In our case, various features of the same choice are evaluated differently according to their salience.

Moreover, their representation relies on the structural assumption of having just two dimensions, and therefore, even if translated in an obvious way to the risk setting, it only applies to binary lotteries. Herweg and Muller (2019) provide a comparison between the Salience and Regret Model, showing that the former can be interpreted as a particular case of the latter. Differently from us, they do not identify the axioms underlying the representation.

**Outline**  The rest of the paper is structured as follows. Section 2 introduces the notation and discusses the mathematical preliminaries. Section 3 describes the weakening of EU that is necessary to capture Salience sensitivity, while Section 4 describe the additional axioms characterizing the Salience model. Section 5 extends the model to choice from nonbinary subsets. Finally, Section 6 gives a Rank-Dependent version of the model. All proofs are in the Appendix.

## 2 Preliminaries

Let $X$ be an arbitrary set of outcomes (or prizes), and denote as $\Delta (X \times X)$ the set of (joint) probability measures over $X \times X$ with finite support (also referred as simple). Recall that the preferences over a set of deterministic alternatives $X$ can be represented as a subset $R$ of the product space $X \times X$, with the interpretation that $x$ is preferred to $y$ if and only if $(x, y) \in R$. In principle, this approach can be extended to risky alternatives in two ways. The standard approach is to represent preferences as a subset $R$ of $\Delta (X) \times \Delta (X)$, again, with the interpretation that $(p_1, p_2) \in R$ if and only if the DM prefers $p_1$ to $p_2$. However, this approach is implicitly assuming that the correlation between two lotteries is irrelevant. In this paper, we follow the other natural extension of the deterministic case, in which the
preferences of the DM are represented by a subset $P$ (called preference set) of the space of simple joint distributions $\Delta(X \times X)$. The interpretation is that the DM faces a joint distribution over outcomes $p$, and he has to decide whether to be paid accordingly to the realization of the first or the second component. Then $p \in P$ if and only if he (weakly) prefers to be paid accordingly to the first component. For every joint distribution $p$, we denote as $p_1 \in \Delta(X)$ and $p_2 \in \Delta(X)$ respectively the first and second marginals of $p$. Formally:

$$p_1(x) = \sum_{y \in X} p(x, y) \text{ and } p_2(y) = \sum_{x \in X} p(x, y).$$

### 2.1 Mapping into experimental implementation

Since preferences sets are not the standard way to describe testable implications of a DM preferences, a roadmap to how test them in an experiment can be helpful. The subject are faced with a finite-support joint distribution $p$ over prizes. The joint distribution can be summarized by the following table:

<table>
<thead>
<tr>
<th>$p$</th>
<th>$y_1$</th>
<th>...</th>
<th>$y_j$</th>
<th>...</th>
<th>$y_m$</th>
</tr>
</thead>
<tbody>
<tr>
<td>$x_1$</td>
<td>$p_{11}$</td>
<td>...</td>
<td>$p_{1j}$</td>
<td>...</td>
<td>$p_{1m}$</td>
</tr>
<tr>
<td>...</td>
<td>...</td>
<td>...</td>
<td>...</td>
<td>...</td>
<td>...</td>
</tr>
<tr>
<td>$x_i$</td>
<td>$p_{i1}$</td>
<td>...</td>
<td>$p_{ij}$</td>
<td>...</td>
<td>$p_{im}$</td>
</tr>
<tr>
<td>...</td>
<td>...</td>
<td>...</td>
<td>...</td>
<td>...</td>
<td>...</td>
</tr>
<tr>
<td>$x_n$</td>
<td>$p_{n1}$</td>
<td>...</td>
<td>$p_{nj}$</td>
<td>...</td>
<td>$p_{nm}$</td>
</tr>
</tbody>
</table>

That is, the experimenter is telling the subject that the outcome $(x_i, y_j)$ realizes with probability $p_{ij}$ and so on. Then, given her knowledge of the correlation structure between the two alternatives, the subject is asked to choose between being paid accordingly to the prizes on the row (the $x$'s) or accordingly to the prizes on the columns (the $y$'s). If he chooses to be paid accordingly to the rows, when the outcome $(x_i, y_j)$ realizes he will be paid $x_i$ regardless of the value of $y_j$.\footnote{Our theory is silent about the information that is revealed to the subject after a joint outcome $(x, y)$ is drawn. One may expect that the behavior may be different whether only the component that is going to be paid out to the DM or the joint realization is revealed.} A joint distribution $p$ is said to belong to the preference set $P$ of the DM if, when faced with such joint distribution, the DM chooses to be paid according to the prizes on the row. Therefore, we can always observe if a joint distribution is in the preference set or not. The typical axioms we impose on preferences sets have the form “if $p \in P$ then some other $p'$ derived from $p$ belongs to $P$ as well,” and therefore can be falsified by showing that a DM chooses to be paid accordingly to the row prizes under $p$ and under the column prizes under $p'$.\footnote{Our theory is silent about the information that is revealed to the subject after a joint outcome $(x, y)$ is drawn. One may expect that the behavior may be different whether only the component that is going to be paid out to the DM or the joint realization is revealed.}
3 General Representation Theorem

To better understand representation (1), it is useful to compare with Expected Utility. Let $p \in \Delta (X \times X)$ be a joint distribution over prizes with marginals $p_1$ and $p_2$. Under EU, there exists a utility function $u$ such that

$$p_1 \succ p_2$$

$$\Leftrightarrow \sum_x u(x) p_1(x) \geq \sum_y u(y) p_2(y)$$

(2)

$$\Leftrightarrow \sum_{x,y} p(x,y) (u(x) - u(y)) \geq 0.$$  

(3)

Given these equivalences, the difference between EU and the representation in (1) can be described in the following way. In principle, when contemplating a joint lottery $p$, two algorithmic procedures can be used to choose according to which component to be paid. The first algorithm is the following:

- Take marginal $p_1$. Consider the utility obtained under each realization. Aggregate these utilities according to the probability measure $p_1$ to get a “score”

$$U(p_1) = \sum_x u(x) p_1(x).$$

Note that this score is independent of $p_2$.

- Follow the same procedure for marginal $p_2$.

- Compare the scores obtained for the two alternatives. Choose to be paid accordingly to the first component if and only if $U(p_1) \geq U(p_2)$.

It is immediate to see that under this procedure, there is no role for correlation between the two marginal distributions. This procedure consists of a Comparison of (Probabilistic) Aggregations, and in the case of EU is given by line (2). Alternatively, one may consider the following procedure:

- Take a possible joint realization $(x, y)$. Compare the two prizes and give a score $\phi(x, y)$, representing a combination of how much $x$ is preferred to $y$ and the attention diverted to that realization, with 0 meaning indifference or zero attention.

- Do the same for every joint realization.

- Aggregate all these local comparisons according to the probability measure $p$ obtaining $\Phi(p) = \sum_{x,y} p(x,y) \phi(x,y)$. 


Choose to be paid accordingly to the first component if and only if $\Phi(p) \geq 0$.

In principle, under this procedure, there is room for correlation. What matters is the joint value assumed by the lotteries under scrutiny. This procedure consists of a (Probabilistic) Aggregation of Comparisons, and it is the kind of reasoning that characterize both Regret and Salience Sensitive DMs, and for EU it corresponds to line (3). The descriptive and normative value of such procedure has been recognized since the pioneering works by Bell (1982), and Loomes and Sugden (1982).

However, the result by vN-M tells us that, under Expected Utility, the two procedures lead to the same conclusion. In the second procedure, the DM will use $\phi(x,y) = u(x) - u(y)$, and it is immediate to see that this separable specification makes correlation irrelevant. The reason is that, for an EU agent, the value of receiving $x$ is $u(x)$ independently of the realization of the counterfactual. Instead, our first step is to provide a set of axioms that characterize the general representation (1) for a (possibly) nonseparable $\phi$.

Given $p \in \Delta(X \times X)$ we define the conjugate distribution as

$$\bar{p}(x,y) = p(y,x).$$

Therefore, the conjugate distribution is just a relabeling of the first and second component into each other.

**Axiom 1 (Completeness)** For all $p \in \Delta(X \times X)$

$$p \not\in P \Rightarrow \bar{p} \in P.$$  

Completeness is a very minimal requirement about the rationality of the DM. If he prefers to be paid accordingly to the second marginal when the joint distribution is $p$, then he (weakly) prefers to be paid accordingly to the first marginal after a relabeling of the two components. Notice that Completeness of the preference set is a weaker assumption that having the standard Completeness assumption assumed on the derived relation $\succeq_P$.

**Lemma 1** If $\succeq_P$ satisfies Completeness, then $P$ satisfies Completeness.

Given this interpretation of the conjugate distribution, given a preference set $P \subseteq \Delta(X \times X)$, the strict preference set is defined as

$$\hat{P} = \{p \in P: \bar{p} \not\in P\}.$$  

In words, a joint distribution $p$ is in the strict preference set if the DM weakly prefers to be paid accordingly to the first component (i.e., $p \in P$) and he does not prefer to be paid accordingly to the second component after a relabeling of the alternatives (i.e., $\bar{p} \not\in P$). We are going to use the concept of strict preference set in our second axiom. This axiom is a generalization to intransitive preferences of the standard principle of reduction for
compound lotteries. If there are two joint distributions $p$ and $q$ such that under each of them the DM prefers to be paid accordingly to the first component, then he prefers to be paid accordingly to the first component even if the joint distribution that is going to be used is $p$ with probability $\lambda$ and $q$ with probability $(1 - \lambda)$. The preference is strict whenever one of the original preferences is.

**Axiom 2 (Strong Independence)** For all $p, q \in P$, $\lambda \in (0, 1)$

$$\lambda p + (1 - \lambda) q \in P$$

and

$$\lambda p + (1 - \lambda) q \in \hat{P}$$

when $q \in \hat{P}$.

Finally, we impose a weak continuity axiom guaranteeing the nonexistence of a joint distribution such that one component is “infinitely preferred” to the other.

**Axiom 3 (Archimedean Continuity)** For all $p \in \hat{P}$, $q \notin P$, there exist $\alpha, \beta \in (0, 1)$ such that

$$\alpha p + (1 - \alpha) q \in \hat{P} \text{ and } \beta p + (1 - \beta) q \notin P.$$  

With these axioms, we are in the position of proving our first representation theorem.$^6$

**Theorem 1** $P \subseteq \Delta (X \times X)$ satisfies Completeness, Strong Independence, and Archimedean Continuity if and only if there exists a skew-symmetric $\phi : X \times X \rightarrow \mathbb{R}$ such that

$$p \in P \iff \sum_{(x,y) \in X \times X} p(x,y) \phi(x,y) \geq 0. \quad (4)$$

Moreover, $\phi$ is unique up to a positive linear transformation.

Notice that a binary relation over marginal distributions induces a preference set, and vice-versa.

**Definition 1** Given a binary relation $\succeq$ on $\Delta (X)$, we define the Preference Set $P_{\succeq}$ as

$$p \in P_{\succeq} \iff p_1 \succeq p_2.$$  

Similarly, a preference set $P$ induces a (possibly incomplete) binary relation on $\Delta (X)$.

---

$^6$The proof of the theorem combines the standard techniques used to prove the vN-M theorem with the ones used to deal with preference sets (see Fishburn 1991), and intransitive preferences over acts (see Fishburn 1989).
Definition 2  Given a preference set $P$, we define the binary relation $\succeq^P$ on $\Delta(X)$ as
\[ p_1 \succeq^P p_2 \iff (\forall q \in \Delta(X \times X) : (q_1, q_2) = (p_1, p_2), q \in P). \]

The requirement imposed in the definition of $\succeq^P$ is strong since for $p_1$ to be preferred to $p_2$ all the joint distributions with those particular marginals have to be in the preference set (i.e., $p_1$ has to be preferred to $p_2$ regardless of their correlation structure). A weaker concept would have replaced “for all $q$” with “for some $q$.” Proposition 2 illustrates the advantage of focusing on this notion.

The names of the axioms used to characterize (4) suggest that, to the celebrated vN-M set of axioms, we are just relaxing Transitivity. The following result shows that this intuition is correct. To do so, we need to translate Transitivity in the language of Preference Sets.

Axiom 4 (Transitivity) For all $p, q, r \in \Delta(X \times X)$, if $p_2 = q_1$, $r_1 = p_1$, and $r_2 = q_2$, then
\[ (p \in P, q \in P) \Rightarrow r \in P. \]

The axiom is a little bit mouthful, but the interpretation is clear. Since $p \in P$, $p_1 = r_1$ is preferred to $p_2 = q_1$ (given the correlation structure described by $p$). Since $q \in P$, $q_1 = p_2$ is preferred to $q_2 = r_2$ (given the correlation structure described by $q$). For Transitivity of the marginals to hold, we then need that $r_1$ is preferred to $r_2$, i.e., $r \in P$.

Proposition 1 If $P$ satisfies Transitivity $\succeq^P$ satisfies Transitivity, $\succeq^P$ satisfies Transitivity and Completeness if and only if $P$ satisfies Transitivity and Completeness.

Following result proves that when Transitivity is imposed on top of the previous axioms, the decision criterion reduces to Expected Utility. Moreover, it characterizes Transitivity as a modularity property for $\phi$.

Proposition 2 If $P$ satisfies Completeness, Strong Independence, and Archimedean Continuity, and $\phi$ represent $P$ as in (4), the following are equivalent:

1. $P$ satisfies Transitivity;
2. $\phi$ is modular:
\[ \phi((x, y) \lor (x', y')) + \phi((x, y) \land (x', y')) = \phi(x, y) + \phi(x', y') \quad \forall x, x', y, y' \in X; \]

\[ (5) \]

\[ ^7 \text{Recall that if } v, v' \in \mathbb{R}^n, \text{ we have } (v \lor v'), (v \land v') \in \mathbb{R}^n \text{ defined as } \\
(v \lor v')_i = \max \{v_i, v'_i\} \text{ and } (v \land v')_i = \min \{v_i, v'_i\}. \]
3. $\succ^P$ satisfies Completeness;

4. $\succ^P$ satisfies Completeness, Transitivity, Continuity, and Independence;

5. $\succ^P$ admits an Expected Utility representation.

Given this special connection between modularity and the EU model, one is left to wonder what happens if $\phi$ is assumed to be either supermodular or submodular. In what follows we will show that precisely a submodular $\phi$ in a subset of his domain is what characterizes the main property of the BGS Salience model.

Indeed, a similar result can be obtained for each axiom separately considered.

**Proposition 3** If $P$ satisfies Transitivity and Completeness, $P$ satisfies Strong Independence if and only if $\succ^P$ satisfies Strong Independence, that is,

$$\forall p_1, p_2, p_3 \in \Delta(X), \alpha \in (0, 1) \quad p_1 \succ^P p_2 \iff \alpha p_1 + (1 - \alpha) p_3 \succeq^P \alpha p_2 + (1 - \alpha) p_3.$$  

### 4 Salience Characterization

In this section, we first describe the two versions of Salience Theory introduced by BGS and their critical properties of Ordering, Diminishing Sensitivity, and Reflexivity, and we show that the Smooth version of the model is a particular case of our general representation. We then propose a behaviorally testable version of the properties, we show that they coincide with the original ones, and we characterize them in terms of the properties of the representing functional.

As a preliminary, since the Salience model is defined for lotteries with monetary outcomes, in this section, we are going to focus on the case where $X = [l, h]$ is a nondegenerate (real) interval. In this setting, it is natural to assume Monotonicity.

**Axiom 5 (Monotonicity)** For all $x, y, z \in X = [l, h]$ and $p \in \Delta(X \times X)$, if $x > y$ and $\alpha \in (0, 1)$, then

$$\alpha \delta_{(y, z)} + (1 - \alpha) p \in P \Rightarrow \alpha \delta_{(x, z)} + (1 - \alpha) p \in \tilde{P}.$$  

Since we do not impose Transitivity, our Monotonicity axiom slightly departs from the usual one: it requires that whenever $x > y$, $x$ is more favorably compared than $y$ to every alternative $z$. Given the representation in (4), Monotonicity is easily characterized in terms of $\phi$.

**Proposition 4** If $P$ admits a representation as in (4), $P$ satisfies Monotonicity if and only if $\phi$ is strictly increasing in the first argument and strictly decreasing in the second argument.
4.1 The BGS Decision Criterion

Salience Theory, as formulated in BGS explains the behavior of a DM that is facing a joint lottery $p \in \Delta (X \times X)$. The main departure of Salience from EU is that expectations are calculated with a *distorted* probability measure, that overweights salient pairs of outcomes. To formalize this idea, BGS introduced the concept of *Salience function*.

**Definition 3** A continuous and bounded function $\sigma : \mathbb{R}^2 \rightarrow \mathbb{R}$ satisfies:

1. **Symmetry** if $\sigma (x, y) = \sigma (y, x)$;
2. **Ordering** if $x' < y', x < y$ and $[x', y'] \subset [x, y]$ imply $\sigma (x', y') < \sigma (x, y)$;
3. **Diminishing sensitivity** if $x > y \geq k > 0$ implies $\sigma (x + k, y) < \sigma (x, y - k)$;
4. **Reflexivity** if $\sigma (x, y) \geq \sigma (w, z) \iff \sigma (-x, -y) \geq \sigma (-w, -z)$.

A Salience function is a function $\sigma : \mathbb{R}^2 \rightarrow \mathbb{R}$ satisfying Symmetry, Ordering, Diminishing Sensitivity, and Reflexivity. We are going to interpret the properties momentarily when we introduce their behaviorally testable counterparts. At this point, it is fundamental to understand a feature of the model: *the salience of a joint realization depends only on its value, not on its probability*. This nature of the distortion is the main difference with Prospect Theory.

4.1.1 Rank-Based Salience Theory

Note that every function $\sigma$ induces a rank on the support of $p$. More precisely, if we let

$$\hat{\sigma}_p (x, y) = \| \{ (x', y') \in \text{supp} : \sigma (x', y') > \sigma (x, y) \} \| \quad \forall (x, y) \in \text{supp} p,$$

we obtain $|\text{supp} p| > \hat{\sigma}_p (x, y) \geq 0$ with $\hat{\sigma}_p (x, y) = 0$ for the most salient pair of outcomes. Given these definitions, we can say when a preference relation admits a Rank-Based Salience Theory representation.

**Definition 4** A preference set $P$ admits a $(\beta, \sigma)$ representation if there exist a continuous and bounded function $\sigma : \mathbb{R}^2 \rightarrow \mathbb{R}$ that satisfies Symmetry, $\beta \in (0, 1]$, and a strictly increasing and concave function $u : \mathbb{R} \rightarrow \mathbb{R}$ such that

$$p \in P \iff \sum_{(x,y) \in \text{supp} \ p} (u (x) - u (y)) \beta^\hat{\sigma}_p (x,y) p (x,y) \geq 0. \quad (6)$$

We say that it admits a Rank-Based Salience Representation if $\sigma$ is a Salience function.

Since $\beta < 1$, and $\hat{\sigma}_p$ is decreasing in the salience of a pair of outcomes, the decision criterion is overweighing the most salient joint realizations. Therefore, this criterion has the advantage of suggesting what the main features of a salience sensitive DM are: he probabilistically aggregate the difference in “hedonic” utilities, with additional weight given to salient pairs of rewards. Notice that EU is included as the particular case in which $\beta = 1$. 11
4.1.2 Smooth Salience Theory

In their paper, BGS recognized that the Rank Dependent version of the model is subject to some issues due to discontinuity, and they suggest the use of a smooth version of the criterion. In their words: “A smooth specification would also address a concern with the current model that states with similar salience may obtain very different weights. This implies that (1) splitting states and slightly altering payoffs could have a large impact on choice, and (2) in choice problems with many states the (slightly) less salient states are effectively ignored.” In what follows, we briefly outline the alternative smooth decision criterion they propose in their Appendix.

**Definition 5** A preference set \( P \) admits a (Smooth) Salience Theory representation if there exist a Salience function \( \sigma : \mathbb{R}^2 \rightarrow \mathbb{R} \) and a strictly increasing and concave function \( u : \mathbb{R} \rightarrow \mathbb{R} \) such that

\[
p \in P \iff \sum_{(x,y) \in \text{supp}} (u(x) - u(y)) \sigma(x,y) p(x,y) \geq 0.
\]

First, we show that the Smooth version of the Salience criterion proposed by BGS is indeed a particular case of our general representation.

**Proposition 5** The (Smooth) Salience Theory model satisfies Completeness, Strong Independence, Archimedean Continuity, and Monotonicity.

At the same time, it is clear that the representation in (4) is much more general than the Salience one, and it allows for behaviors that are at odds with the critical intuition that states, where the alternatives differ more, are overweighted. Rest of this section is devoted to the characterization of Ordering, Diminishing Sensitivity, and Reflexivity in terms both of testable axioms and the properties of the representing \( \phi \).

Here we highlight our strategy. We first propose a testable version of each property, then we show that, under the \((\beta, \sigma)\) representation, these testable versions coincide with the corresponding property postulated by BGS. Finally, motivated by the BGS concern with the discontinuity of the criterion, and the fact that they consider the rank-dependent version only to better single-out the properties of the model, as well as the critique in Kontek (2016), we characterize the functional implication in the smooth case covered by the representation in (4).

---

8See also Kontek (2016), and Dertwinkel-Kalt and Mats Köster (2019) for the reasons to focus on the continuous version of the model.

9In their words, “Our assumption of rank-based discounting buys us analytical tractability ... for simplicity we stick to rank-based discounting”
4.2 The Ordering Axiom

The idea behind the Ordering property proposed by BGS is straightforward. Suppose that outcomes \((x_H, x_L, y_H, y_L)\) are such that

\[ x_H \geq x_L \geq y_H \geq y_L. \]

Then, since \(x_H\) and \(y_L\) differ more than \(x_L\) and \(y_H\), if a joint distribution \(p\) assigns positive probability to both \((x_H, y_L)\) and \((x_L, y_H)\), Ordering implies that the probability of the former outcome will be overweighted with respect to the probability of the latter. However, distortions of probabilities are not observable, and therefore, we cannot directly test the BGS form of Ordering.

Nevertheless, given Proposition 5, we can propose a behavioral (i.e., testable) version of the Axiom. We describe the idea in words before presenting the formal axiom. Suppose that the DM envisions the following joint distribution

\[ p = \left( \frac{1}{4}; (x_H, y_L), \frac{1}{4}; (x_L, y_H), \frac{1}{4}; (y_L, x_L), \frac{1}{4}; (y_H, x_H) \right). \]

There are four possible outcomes, and the two marginal distributions coincide, i.e., \(p_1 = p_2\). Therefore, an Expected Utility maximizer is indifferent between the two components. However, the attention of a Salience sensitive DM is disproportionately drawn to the outcome that has the most significant difference between payoff (in the inclusion sense). Since this outcome is \((x_H, y_L)\), and since it favors the first component, a Salience sensitive DM prefers (at least weakly) to be paid accordingly to the first component. The previous reasoning is crystallized in the Ordering Axiom.

**Axiom 6 (Ordering)** For every

\[ x_H \geq x_L \geq y_H \geq y_L \]

we have that

\[ p = \left( \frac{1}{4}; (x_H, y_L), \frac{1}{4}; (x_L, y_H), \frac{1}{4}; (y_L, x_L), \frac{1}{4}; (y_H, x_H) \right) \in P. \]

It satisfies Strict Ordering if \(p \in \hat{P}\) whenever \(x_H > x_L\) and \(y_H > y_L\).

The next proposition characterizes Ordering in terms of the properties of \(\phi\) and shows that the axiom corresponds to the original property of BGS. It is convenient to define the set of pair of outcomes such that the first component is larger than the second

\[ \oplus := \{(x, y) \in X \times X : x \geq y\} \]

since we are going to consider the restriction \(\phi|_\oplus\) of \(\phi\) on this subset of the domain.
Proposition 6 Suppose that $P$ admits a representation as in in (4). Then it satisfies Ordering (resp. Strict Ordering) if and only if $\phi|_B$ is submodular (resp. strictly submodular).

If $P$ admits a $(\beta, \sigma)$ representation:

1. It satisfies Ordering if $\sigma$ satisfies Ordering, and it satisfies Strict Ordering if $\sigma$ satisfies Ordering and $\beta \in (0, 1)$.

2. The converse is true if when $x > y > z$

$$\text{if } \sigma(x, y) > \sigma(y, z) \Rightarrow (u(x) - u(y)) \geq (u(y) - u(z)).$$

The additional condition imposed in (7) is ruling out a pathological specification of the Salience Model, in which a pair of outcomes that has a negligible difference in hedonic utilities result more salient than another joint realization with a higher difference between hedonic utilities. In other words, it is a comonotonicity requiring that attention is increasing between hedonic experiences. However, notice that is an ordinal restriction that imposes very little structure on the shape of $\sigma$ and $u$. The property was not assumed in the original model by BGS because it is particularly demanding in the context of consumer choice where for example, the difference between two colors of the packaging may be nearly irrelevant in terms of experienced utility, but it may nevertheless capture the attention of the DM. Since in this paper, we focus on the choice between monetary lotteries, it is natural to impose this monotonicity condition.

4.2.1 Ordering and previous models

The combination of Propositions 2 and 6 shows that (Strict) Ordering is the property of the Salience model that cannot be reconciled with EU. More generally, it is immediate to see that Strict Ordering of $P$ implies that the preference relation $\succeq^P$ is incomplete, and therefore it is incompatible with all the decision criterion, like expected Utility and Prospect Theory, that can rank distributions only on the basis of their marginals.

Moreover, Ordering can also be related with the property of Convexity of the Regret Theory model.

Remark 1 Notice that Ordering implies the following property of Convexity that Loomes and Sugden (1987) imposed on Regret Theory:

$$x \geq y \geq z \Rightarrow \phi(x, z) \geq \phi(x, y) + \phi(y, z).$$

Indeed, by inspection of the proof we can see that an even weaker, but more involved condition is needed.
4.3 The Diminishing Sensitivity Axiom

The Diminishing Sensitivity property requires that when two pairs of outcomes have the same absolute difference, the one with the higher relative difference is overweighted. The interpretation is easier for two-outcome lotteries. Suppose that the DM is envisioning the joint probability distribution $p$ that assigns a probability $\frac{1}{2}$ both to $(x, y)$ and $(y + k, x + k)$, with $x > y \geq 0$. Then, the two pairs of outcomes have the same absolute difference, but $(x, y)$ has a higher relative difference. Therefore, $(x, y)$ is overweighted to $(y + k, x + k)$. Since $(x, y)$ favors the first marginal, the DM is going to choose to be paid according to the first component. The previous reasoning is crystallized in the Diminishing Sensitivity axiom.

Axiom 7 (Diminishing Sensitivity) For every $x > y \geq 0$, and $k \in \mathbb{R}_+$

$$p = \left( (x, y), \frac{1}{2}, (y + k, x + k), \frac{1}{2} \right) \in P.$$  

It satisfies Strict Diminishing Sensitivity if $p \in \hat{P}$ whenever $k \in \mathbb{R}_{++}$.

The next proposition shows that our definition of Diminishing Sensitivity corresponds to the original property of BGS, and it characterizes the axiom in terms of the directional derivative of the representing $\phi$. Recall that for a function $\phi$ from $\mathbb{R} \times \mathbb{R}$ to $\mathbb{R}$, $\nabla_{(w,z)} \phi (x, y)$ denotes the derivative in the direction $(w, z)$ computed at $(x, y)$:

$$\nabla_{(w,z)} \phi (x, y) = \lim_{h \to 0} \frac{\phi (x + hw, y + hz) - \phi (x, y)}{h}.$$  

Proposition 7 Suppose that $P$ admits a representation as in (4) with a differentiable $\phi$. Then $P$ satisfies Diminishing Sensitivity if and only if

$$\nabla_{(1,1)} \phi |_{\mathbb{R}} \leq 0.$$  

If instead $P$ admits a $(\beta, \sigma)$ representation, it satisfies Strict Diminishing Sensitivity if $\sigma$ satisfies Diminishing Sensitivity. The converse is true if $u$ is linear and $\beta < 1$.

The assumption of linear $u$ used to obtain the second part of the statement as an if and only if it is maintained in most of the BGS paper. Indeed, we will momentarily see that having both Diminishing Sensitivity and a strictly concave $u$ introduces two different sources of risk aversion and that one of them is redundant.

In particular, it turns out that Diminishing Sensitivity alone is not in contrast with the conventional notion of EU/Prospect Theory. It corresponds to a generalization of the property of risk aversion in gains and risk loving in losses of Prospect Theory to decision criteria that are not necessarily transitive.

Proposition 8 Let $P$ admit an Expected Utility Representation with a strictly increasing utility function. Then $P$ satisfies Diminishing Sensitivity if and only if it satisfies Risk Aversion for gains.
4.4 The Reflexivity Axiom

The last property introduced by BGS is Reflexivity, and it captures the symmetry of distortions to 0.

**Axiom 8 (Reflexivity)** For every \( x, y, w, z \in \mathbb{R}_+ \), \( \alpha \in [0, 1] \)

\[
\alpha \left( (x, y), \frac{1}{2}; (w, z), \frac{1}{2} \right) + (1 - \alpha) \delta_{\max\{x, y, w, z\}} \in P \setminus \hat{P} \iff \\
\alpha \left( (-x, -y), \frac{1}{2}; (-w, -z), \frac{1}{2} \right) + (1 - \alpha) \delta_{\min\{-x, -y, -w, -z\}} \in P \setminus \hat{P}.
\]

**Proposition 9** If \( P \) admits a \((\beta, \sigma)\) representation with linear \( u \), \( \beta < 1 \), and \( \sigma \) satisfying Ordering, it satisfies Reflexivity if and only if \( \sigma \) satisfies Reflexivity.

We did not attach any specific interpretation to the realization of the lotteries, except that they are expressed in monetary units. In particular, they can represent either the total wealth or gains and losses obtained after the realization of some uncertainty. However, the Reflexivity axiom, and the implied role for outcome 0, better suits the latter interpretation. Indeed, we notice that Reflexivity implies a form of the Preference Reversal of risk attitudes featured by Prospect Theory.

**Proposition 10** Suppose that \( P \) has an EU representation with a continuous utility index \( u \) and satisfies Reflexivity. Then \( P \) is risk-averse (resp. risk-loving) for lotteries with values in \([a, b] \subseteq \mathbb{R}_+\) if and only if \( \gtrless \) is risk loving (resp. risk-averse) for lotteries with values in \([-b, -a]\).

4.5 Exponential Parametrization

In this section, we present a flexible and tractable class of preferences that features a representation as in (4) and satisfies the Ordering, Diminishing Sensitivity, and Reflexivity Axioms in a nontrivial way. They are such that

\[
\phi (x, y) = \begin{cases} 
\frac{e^{u(x)} - 1}{e^{u(y)} - 1} & x \geq y \\
1 - \frac{e^{u(y)}}{e^{u(x)}} & x < y
\end{cases}
\]

with

\[
u (x) = \begin{cases} 
v (x) & x \geq 0 \\
v (x) + \log \lambda & x < 0
\end{cases}
\]

and \( v \) strictly increasing, concave on \( \mathbb{R}_+ \) and odd, \( \lambda \in \mathbb{R}_+ \). The shape of the function \( v \) has a key role in determining the relative weights of Ordering and Diminishing Sensitivity, whereas \( \lambda \) can be used to capture the idea of loss aversion. In particular, this parametric class can explain the difference between the two versions of the Allais paradox presented in BGS.
**Example 1** Let \( v(x) = \text{Id}(x), \lambda = 1, \) so that

\[
\phi(x, y) = \begin{cases} 
  e^x - 1 & x \geq y \\
  1 - e^y & x < y.
\end{cases}
\]

Recall that in the Allais paradox the marginal distributions are:

\[
\begin{align*}
  p_1 &= (2500, 0.33; 0, 0.01; z, 0.66) \\
  p_2 &= (2400, 0.34; z, 0.66).
\end{align*}
\]

If we follow BGS and assume that in the standard Allais paradox alternatives are perceived as independent, we have:

\[
p_{z=2400} = ((2500, 2400), 0.33; (0, 2400), 0.01; (2400, 2400), 0.66).
\]

In that case, it is immediate to see that:

\[
\Phi(p_{z=2400}) = 0.33(e^{100} - 1) - 0.01(e^{2400} - 1) < 0.
\]

Similarly,

\[
p_{z=0} = ((2500, 2400), 0.11; (2500, 0), 0.218; (0, 2400), 0.221; (0, 0), 0.442)
\]

moreover, it is immediate to see that

\[
\Phi(p_{z=0}) = 0.11(e^{100} - 1) + 0.218(e^{2500} - 1) - 0.221(e^{2400} - 1) > 0
\]

that is, the results are consistent with the well-documented pattern of choice in the standard Allais paradox. Next, consider the version proposed by BGS, where it is made explicit that the common outcome is paid in the same state of the world by both lotteries. In this case, the joint distribution when \( z = 0 \) remains the same:

\[
q_{z=2400} = ((2500, 2400), 0.33; (0, 2400), 0.01; (2400, 2400), 0.66)
\]

and therefore we still have

\[
\Phi(q_{z=2400}) < 0.
\]

Instead, the joint distribution when \( z = 0 \) is affected:

\[
q_{z=0} = ((2500, 2400), 0.33; (0, 2400), 0.01; (0, 0), 0.66)
\]

and

\[
\Phi(q_{z=0}) = 0.33(e^{100} - 1) - 0.01(e^{2400} - 1) < 0
\]

and this prediction is consistent with the experimental evidence generated in BGS. ▲

Notice that the exponential parametrization with \( u = \text{Id} \) used in the previous example is very stark. Given the exponential structure and the large numbers involved, for the comparison between two alternatives is often sufficient to look at the pair of outcomes where they differ the most. This starkness is not a general feature of this parametric family. It is enough to have a concave or even a linear (with a lower slope) \( u \) and also the other pairs of outcomes start to play a role.
5 Choice from arbitrary sets

An important question left open by the previous analysis involves the choice from a set of more than two alternatives. In general, since the Salience decision criterion is intransitive, it is possible that, given a choice set $A$, no element of $A$ is (weakly) preferred to all the other options. In what follows, we build on a result by Kreweras (1961) to show that when the DM can use a mixed strategy, (i.e., he can choose a probability measure over $A$ using an external randomization device) there exists an optimal randomization over the available alternatives.

It is easier to state the result if we move from the space of lotteries to the space of acts. Formally, we consider a probability space $(S, \mathcal{F}, \mathbb{P})$, and the alternatives under considerations are simple (i.e., they assume a finite number of values) and measurable random variables on that space. Notice that a preference set $P$ (satisfying Completeness, Strong Independence, and Archimedean Continuity) naturally induces a binary relation on the set $B_0(S)$ of such random variables. Indeed, let $\phi$ be the representing skew-symmetric functional whose existence is guaranteed by Theorem 1. Then, notice that two acts $f, g \in B_0(S)$ induce a joint distribution over outcomes $p_{f,g}$ defined as

$$p_{f,g}(x,y) = \mathbb{P}(\{s \in S : f(s) = x, g(s) = y\}).$$

The preference relation on $B_0(S)$ is defined as

$$f \succeq_P g \iff p_{f,g} \in P \iff \sum_{(x,y)} p_{f,g}(x,y) \phi(x,y) \geq 0.$$

Now, suppose that the DM has to choose an alternative in $A \subseteq B_0(S)$. It is easy to see that there is no guarantee $\succeq_P$ is transitive, and therefore there may be no $f \in A$ such that $f \succeq_P g$ for all $g \in A$. However, if the DM is allowed to randomize between the alternatives in $A$, then the existence of a (possibly mixed) optimal alternative is guaranteed.

**Proposition 11** Let $\phi$ represent $P$ as in (4) and $A$ be a finite subset of $B_0(S)$. Then, there exists $\alpha \in \Delta(A)$ such that for all $\beta \in \Delta(A)$

$$\sum_{f \in A} \sum_{g \in A} \alpha(f) \beta(g) \sum_{(x,y)} p_{f,g}(x,y) \phi(x,y) \geq 0.$$

6 Rank Dependent Salience

Although just for technical convenience, most of the analysis in the BGS paper was conducted using the rank-dependent criterion. Even if we have already hinted at the flaws of this approach, it is still possible to give and axiomatize a rank-dependent version that is immune to the critique of Kontek (2016). This approach is the intransitive version of
the dual analysis used by Yaari to axiomatize the standard Rank-Dependent model with linear Bernoulli utility function. To use such dual techniques, we need to move to the space of decumulative distribution functions induced by a lottery. However, since under Salience Theory the correlation between alternatives is crucial, the relevant decumulative distribution is the one of the differences between what is paid by the two components of the lotteries. Since the previous analysis pointed out how Ordering is the key feature of the Salience model, we isolate it by implicitly ruling out Diminishing Sensitivity with the following maintained assumption. In this section, fix also $X = [l, h]$.

**Axiom 9 (Neutrality)** If $p$ and $q$ are such that
\[
p \left( \{ (x, y) : x - y \leq k \} \right) = q \left( \{ (x, y) : x - y \leq k \} \right) \quad \forall k \in X
\]
then $p \in P$ if and only if $q \in P$.

Let DDF be the set of decumulative distribution functions, that is, those functions $F$ that are nonincreasing, right continuous and such that
\[
\lim_{x \to l-h} F(x) = 1 = 1 - \lim_{x \to h-l} F(x).
\]
We denote as $F^{-1}$ the inverse function of $F$. The operation $\boxplus$ in the space of DDF is defined as
\[
\alpha F \boxplus (1 - \alpha) G = \left( \alpha F^{-1} + (1 - \alpha) G^{-1} \right)^{-1} \quad \forall F, G \in \text{DDF}, \alpha \in (0, 1).
\]
Given Neutrality, the statement $F \in P$ will have the unambiguous meaning that every $p$ such that
\[
p \left( \{ (x, y) : x - y \leq k \} \right) = 1 - F(k)
\]
belongs to $P$. The first axiom that we need for our representation is a standard Monotonicity axiom.

**Axiom 10 (DDF Monotonicity)** If $F, G \in \text{DDF}$, $F \in P$, $G \succ_{\text{FOSD}} F$, then $G \in \hat{P}$.

The second assumption is the stronger notion of Continuity axiom imposed in Yaari (1987).

**Axiom 11 (L1 Continuity)** Let $F \in \text{DDF}$, with $F \in \hat{P}$. Then, there exists $\theta > 0$ such that for all $G \in \text{DDF}$, $\| F - G \| < \theta$ implies $G \in \hat{P}$.

For the last axiom we denote as $F^+$ and $F^-$ the negative and positive truncation of $F$. Formally
\[
F^+ (a) = \begin{cases} 
1 & a < 0, \\
F(a) & a \geq 0.
\end{cases}
\]
and
\[
F^- (a) = \begin{cases} 
0 & a > 0, \\
F(a) & a \leq 0.
\end{cases}
\]
Axiom 12 (Dual Independence) Let $F, G \in DDF$. If one of the following holds

- $\tilde{F}^- = \tilde{G}^-$, $\tilde{F}^- = \tilde{G}^-$ and there exists $\tilde{H}, \tilde{H} \in DDF$, $\alpha \in (0, 1)$ such that $\tilde{F}^+ = \alpha F^+ \boxplus (1 - \alpha) \tilde{H}$, $G^+ = \alpha G^+ \boxplus (1 - \alpha) \tilde{H}$ as well as $\tilde{F}^- = \alpha F^- \boxplus (1 - \alpha) \tilde{H}$, $G^- = \alpha G^- \boxplus (1 - \alpha) \tilde{H}$;

- $\tilde{F}^+ = \tilde{G}^+$, $\tilde{F}^+ = \tilde{G}^+$ and there exists $\tilde{H}, \tilde{H} \in DDF$, $\alpha \in (0, 1)$ such that $\tilde{F}^- = \alpha F^- \boxplus (1 - \alpha) \tilde{H}$, $G^- = \alpha G^- \boxplus (1 - \alpha) \tilde{H}$, as well as $\tilde{F}^- = \alpha F^- \boxplus (1 - \alpha) \tilde{H}$, $G^+ = \alpha G^+ \boxplus (1 - \alpha) \tilde{H}$.

then

$$F \in \hat{P}, \ G \notin \hat{P} \implies \gamma(F \notin \hat{P}, \ G \in P).$$

The content of the axiom is independence when taking convex combination only on the outcomes that favor one of the two components. With this, we are ready to give our representation theorem.

Theorem 2 If $P$ satisfies Neutrality, DDF Monotonicity, L1-Continuity and Dual Independence, then there exist continuous nondecreasing $f_-, f_+ \in \mathbb{R}^{[0,1]}$ such that for all $F \in DDF$:

$$F \in P \iff \int_{1-h}^{0} f_-(F(t)) \, dt + \int_{0}^{h-l} f_+(F(t)) \, dt + h - l \geq 0. \quad (9)$$

Notice that when $f_- = \text{Id} = f_+$, with Id denoting the identity function, the decision criterion reduces to choosing the component with the highest expected value.

7 Appendix

Proof of Lemma 1 Let $p \in P$. Since $\succeq^p$ satisfies Completeness at least one between $p_1 \succeq^p p_2$ and $p_2 \succeq^p p_1$ holds, that is, at least one between $p \in P$ and $\tilde{p} \in P$ holds. \hfill \blacksquare

Let $\oplus = \{(x, y) : \delta(x, y) \in P\}$ and $\hat{\oplus} = \{(x, y) : \delta(x, y) \in \hat{P}\}$. Given our maintained Monotonicity assumption, $\oplus$ coincides with the object used in the main text.

Proof of Theorem 1 The necessity of the axioms given the representation is obvious. Next, we prove their sufficiency. We start by establishing some preliminary claims.

Claim 1 If $\text{supp} \subseteq \oplus$, then $p \in P$. 

20
Proof The claim is proved by induction on the size of supp. The claim is clearly true for \(|\text{supp}p| = 1\). Suppose the result holds for all the lotteries with support of size \(n\). Let \(p\) be such that \(|\text{supp}p| = n + 1\). Choose arbitrarily \((x', y') \in \text{supp}p\). Then, if we define \(q\) as

\[
q(x, y) = \begin{cases} 
0 & (x, y) = (x', y') \\
\frac{p(x, y)}{1 - p(x', y')} & \text{otherwise.}
\end{cases}
\]

Since \(|\text{supp}q| = n\) and \(\text{supp}q \subseteq \oplus\), we have \(q \in P\). Moreover,

\[
p(\cdot) = p(x', y') \delta_{(x', y')} (\cdot) + (1 - p(x', y')) q(\cdot)
\]

by Strong Independence, we have \(p \in P\).

Claim 2 Let \(p \in \hat{P}, q \not\in P\), there exists a unique \(\lambda \in (0, 1)\) such that

\[
\lambda p + (1 - \lambda) q \in P\setminus \hat{P}.
\]

Proof We let

\[
A = \left\{ \lambda \in [0, 1] : \lambda p + (1 - \lambda) q \in \hat{P} \right\},
\]

\[
B = \left\{ \lambda \in [0, 1] : \lambda p + (1 - \lambda) q \not\in P \right\}.
\]

By Archimedean Continuity, both \(A\) and \(B\) have a nonempty intersection with \((0, 1)\). Suppose that \(\lambda \in A\) and \(\mu \in (\lambda, 1]\). Then

\[
\mu p + (1 - \mu) q = \frac{\mu - \lambda}{1 - \lambda} p + \frac{1 - \mu}{1 - \lambda} (\lambda p + (1 - \lambda) q)
\]

and Strong Independence implies that \(\mu p + (1 - \mu) q \in \hat{P}\), and \(\mu \in A\).

Suppose instead that \(\lambda \in B\) and \(\mu \in [0, \lambda)\). Then by Completeness

\[
\lambda q + (1 - \lambda) \bar{q} = \lambda p + (1 - \lambda) q \in \hat{P}
\]

and

\[
\mu \bar{p} + (1 - \mu) \bar{q} = \frac{\lambda - \mu}{\lambda} \bar{q} + \frac{\mu}{\lambda} (\lambda \bar{p} + (1 - \lambda) \bar{q}).
\]

Therefore, \(\mu \bar{p} + (1 - \mu) \bar{q} \in \hat{P}\) by Strong Independence, and \(\mu p + (1 - \mu) q \not\in P\). This, in turn, implies that \(\mu \in B\).

Summing up, \(A\) and \(B\) are two intervals in \([0, 1]\) with empty intersection. Suppose by contradiction that \(A \cup B = [0, 1]\). Then, we either have

\[
A = [\lambda^*, 1] \text{ and } B = [0, \lambda^*)
\]
or

\[ A = (\lambda^*, 1) \text{ and } B = [0, \lambda^*]. \]

If we are in the first case, \( \lambda^* p + (1 - \lambda^*) q \in \hat{P}, q \notin \hat{P} \) and Archimedean Continuity imply the existence of a \( \mu \in [0, \lambda^*) \) such that \( \mu p + (1 - \mu) q \in \hat{P} \), a contradiction. Similarly, we can rule out the other case. Therefore, there exists \( \lambda^* \in [0, 1] \setminus (A \cup B) \), that is, \( \lambda^* p + (1 - \lambda^*) q \in P \setminus \hat{P} \).

It only remains to prove uniqueness. Suppose that both \( \lambda^* \) and \( \mu^* \) have the desired property, and let \( \mu^* > \lambda^* \). Then,

\[
\mu^* p + (1 - \mu^*) q = \frac{\mu^* - \lambda^*}{1 - \lambda^*} p + \frac{1 - \mu^*}{1 - \lambda^*} (\lambda^* p + (1 - \lambda^*) q)
\]

and by Strong Independence, \( \mu^* p + (1 - \mu^*) q \in \hat{P} \), a contradiction. \( \square \)

**Claim 3** Let \( x, y, z, w, t, v \in X \), \( \lambda, \mu, \alpha \in (0, 1) \) and \( \delta_{(x,y)}, \delta_{(z,w)}, \delta_{(t,v)} \in \hat{P} \) with

\[
\lambda \delta_{(x,y)} + (1 - \lambda) \delta_{(w,z)} \in P \setminus \hat{P}, \\
\mu \delta_{(z,w)} + (1 - \mu) \delta_{(v,t)} \in P \setminus \hat{P}, \\
\alpha \delta_{(t,v)} + (1 - \alpha) \delta_{(y,x)} \in P \setminus \hat{P}.
\]

Then

\[
\frac{\lambda}{1 - \lambda} \cdot \frac{\mu}{1 - \mu} \cdot \frac{\alpha}{1 - \alpha} = 1.
\]

**Proof** Let

\[
\gamma = \frac{\mu}{\mu + 1 - \lambda}
\]

and

\[ p = \gamma \left( \lambda \delta_{(x,y)} + (1 - \lambda) \delta_{(w,z)} \right) + (1 - \gamma) \left( \mu \delta_{(z,w)} + (1 - \mu) \delta_{(v,t)} \right). \]

By Strong Independence, \( p \in P \setminus \hat{P} \). Since \( \gamma (1 - \lambda) = (1 - \gamma) \mu \), by Completeness we have that

\[
\frac{\gamma (1 - \lambda) \delta_{(w,z)} + (1 - \gamma) \mu \delta_{(z,w)}}{\gamma (1 - \lambda) + (1 - \gamma) \mu} \in P \setminus \hat{P}.
\]

Therefore, applying Strong Independence again, we get

\[
\frac{\gamma \lambda \delta_{(x,y)} + (1 - \gamma) (1 - \mu) \delta_{(v,t)}}{\gamma \lambda + (1 - \gamma) (1 - \mu)} \in P \setminus \hat{P}
\]

and by definition of \( \hat{P} \)

\[
\frac{\gamma \lambda \delta_{(y,x)} + (1 - \gamma) (1 - \mu) \delta_{(t,v)}}{\gamma \lambda + (1 - \gamma) (1 - \mu)} \in P \setminus \hat{P}.
\]
But then Claim 2 gives \(1 - \alpha = \frac{\gamma \lambda}{\gamma \lambda + (1 - \gamma)(1 - \mu)}\) that implies
\[
\alpha \mu \lambda = (1 - \lambda)(1 - \mu)(1 - \alpha)
\]
proving the statement. \(\square\)

Claim 4 If \(\text{supp} \subseteq \oplus\), and \(\text{supp} \cap \ominus \neq \emptyset\) then \(p \in \hat{P}\).

Proof Let \((x', y') \in \text{supp} \cap \ominus\). Let
\[
q(x, y) = \begin{cases} 
0 & (x, y) = (x', y') \\
\frac{p(x, y)}{1 - p(x', y')} & \text{otherwise}.
\end{cases}
\]
By Claim 1 \(q \in P\). Since
\[
p(\cdot) = (1 - p(x', y')) q(\cdot) + (1 - p(x', y')) q(\cdot)
\]
Strong Independence implies that \(p \in \hat{P}\). \(\square\)

Claim 5 If \(s, q \in P \setminus \hat{P}\), then
\[
\lambda r + (1 - \lambda) q \in P \iff \lambda r + (1 - \lambda) s \in P.
\]

Proof By Strong Independence both statements hold if \(r \in P\). If \(r \notin P\), by Completeness \(\bar{r} \in \hat{P}\), and by assumption \(\bar{s}, \bar{q} \in P\). Therefore, by Strong Independence both \(\lambda \bar{r} + (1 - \lambda) \bar{q}\) and \(\lambda \bar{r} + (1 - \lambda) \bar{s}\) are in \(\hat{P}\). But then, neither \(\lambda r + (1 - \lambda) q \in P\) nor \(\lambda r + (1 - \lambda) s \in P\). \(\square\)

If for every \(x, y \in X\), \(\delta_{(x, y)} \in P \setminus \hat{P}\), by Claim 1 every \(p \in \Delta (X \times X)\) is in \(P \setminus \hat{P}\), and the statement of the theorem trivially holds by letting \(\phi(x, y) = 0\) for all \(x, y \in X\). Therefore, fix \((\hat{x}, \hat{y})\) with \(\delta_{(\hat{x}, \hat{y})} \in \hat{P}\) and let \(\phi(\hat{x}, \hat{y})\) be an arbitrary strictly positive real number. Moreover, let \(\phi(x, y) = 0\) for all \(\delta_{(x, y)} \in P \setminus \hat{P}\). If \((x, y) \notin \oplus\), by Claim 2, there exists a unique \(\lambda \in (0, 1)\) with
\[
\lambda \delta_{(\hat{x}, \hat{y})} + (1 - \lambda) \delta_{(x, y)} \in P \setminus \hat{P}.
\]
In this case, let
\[
\phi(x, y) = -\phi(\hat{x}, \hat{y}) \frac{\lambda}{(1 - \lambda)}.
\]
It only remains to define \(\phi\) when \((x, y) \in \oplus\). We set
\[
\phi(x, y) = -\phi(y, x) \quad \forall \ (x, y) \in \oplus.
\]
Given the choice of a particular \((\hat{x}, \hat{y})\), the function \(\phi\) so defined is unique up to a positive linear transformation, since the only degree of freedom is the choice of the (strictly positive)
number \( \phi(\tilde{x}, \tilde{y}) \) and the values assumed by \( \phi \) on the rest of the domain are linear in \( \phi(\tilde{x}, \tilde{y}) \). Suppose that we define \( \tilde{\phi} \) starting from a different \((\tilde{x}, \tilde{y}) \in \hat{\Theta}\). Since we are proving uniqueness only up to a positive linear transformation, we can choose the (strictly positive) value of \( \tilde{\phi}(\tilde{x}, \tilde{y}) \). In particular, set

\[
\tilde{\phi}(\tilde{x}, \tilde{y}) = \phi(\tilde{x}, \tilde{y}) = \phi(\tilde{x}, \tilde{y}) \frac{\mu}{1 - \mu}
\]

where

\[
\mu \delta(\tilde{x}, \tilde{y}) + (1 - \mu) \delta(\tilde{y}, \tilde{x}) \in P \setminus \hat{P}
\]

and consider \((x, y) \notin \hat{\Theta}\). Then, there exist \(\lambda_0, \lambda_1\), such that

\[
\lambda_0 \delta(\tilde{x}, \tilde{y}) + (1 - \lambda_0) \delta(x, y) \in P \setminus \hat{P},
\]

\[
\lambda_1 \delta(\tilde{x}, \tilde{y}) + (1 - \lambda_1) \delta(x, y) \in P \setminus \hat{P}.
\]

Given our definitions,

\[
\phi(x, y) = \tilde{\phi}(x, y) \iff \phi(\tilde{x}, \tilde{y}) = \phi(\tilde{x}, \tilde{y}) \frac{\lambda_0}{1 - \lambda_0} = \phi(\tilde{x}, \tilde{y}) \frac{\lambda_1}{1 - \lambda_1}
\]

\[
\iff \phi(\tilde{x}, \tilde{y}) = \phi(\tilde{x}, \tilde{y}) \frac{\mu}{1 - \mu} \frac{\lambda_1}{\lambda_0} (1 - \lambda_1)
\]

and Claim 3 together with Completeness guarantees that the condition in the last line holds true. Finally, we want to show that

\[
p \in P \iff \sum_{(x, y) \in \text{supp}} p(x, y) \phi(x, y) \geq 0.
\]

We are going to consider three possible cases.

(First Case) Suppose \(\text{supp} \subseteq \hat{\Theta}\), then by Claim 1 \(p \in P\), and by definition of \(\phi\), \(\phi(x, y) \geq 0\) for every \((x, y) \in \text{supp}\).

(Second Case) Suppose \(\text{supp} \subseteq \hat{\Theta}\), and \(\text{supp} \cap \hat{\Theta} \neq \emptyset\). Then by Claim 4 \(\tilde{p} \in \hat{P}\) and \(p \notin P\). By definition of \(\phi\), \(\phi(x, y) \leq 0\) for every \((x, y) \in \text{supp}\), and \(\phi(x, y) < 0\) for some \((x, y) \in \text{supp}\).

(Third Case) Finally, we show that all the other possibilities can be reduced into one of the first two cases. Fix \(t \in X\). Suppose we are not in one of the first two cases, that is, there exists \((x_0, y_0), (x_1, y_1) \in \text{supp}\) with \((x_0, y_0), (y_1, x_1) \in \hat{\Theta}\). Then, by Claim 2 there exists a unique \(\alpha \in \mathbb{R}_+\) such that \(\alpha \delta(\tilde{x}, \tilde{y}) + (1 - \alpha) \delta(x, y) \in P \setminus \hat{P}\), and by definition of the original \(\phi\) and uniqueness up to a positive linear transformation, \(\frac{\alpha}{1 - \alpha} \phi(x_0, y_0) = \phi(y_1, x_1)\). If \(\text{supp} = \{(x_0, y_0), (x_1, y_1)\}\), the proof of Claim 2 immediately guarantees that \(p \in P\) if and only if

\[
\frac{\alpha}{1 - \alpha} \leq \frac{p(x_0, y_0)}{p(x_1, y_1)},
\]

that is, if and only if

\[
p(x_0, y_0) \phi(x_0, y_0) + p(x_1, y_1) \phi(x_1, y_1) \geq 0.
\]

24
Therefore, suppose supp\( \neq \{(x_0, y_0), (x_1, y_1)\}\). If \(\frac{\alpha}{1-\alpha} = \frac{p(x_0, y_0)}{p(x_1, y_1)}\), then Claim 5 guarantees that \(p \in P\) if and only if \(p' \in P\) where\(^{11}\)

\[
p'(x, y) = \begin{cases} 
  p(x, y) & (x, y) \notin \{(x_0, y_0), (x_1, y_1), (t, t)\} \\
  0 & (x, y) \in \{(x_0, y_0), (x_1, y_1)\} \\
  p(t, t) + p(x_0, y_0) + p(x_1, y_1) & (x, y) = (t, t).
\end{cases}
\]

Moreover,

\[p(x_0, y_0) \phi(x_0, y_0) + p(x_1, y_1) \phi(x_1, y_1) = 0 = \phi(t, t) (p(t, t) + p(x_0, y_0) + p(x_1, y_1))\]

so that

\[
\sum_{(x,y) \in \text{supp}} p(x, y) \phi(x, y) \geq 0 \iff \sum_{(x,y) \in \text{supp}'} p'(x, y) \phi(x, y) \geq 0.
\]

If \(\frac{\alpha}{1-\alpha} > \frac{p(x_0, y_0)}{p(x_1, y_1)}\), Claim 5 guarantees that \(p \in P\) if and only if \(p' \in P\) where\(^{12}\)

\[
p'(x, y) = \begin{cases} 
  p(x, y) & (x, y) \notin \{(x_0, y_0), (x_1, y_1), (t, t)\} \\
  0 & (x, y) = (x_0, y_0) \\
  p(x_1, y_1) - \frac{1-\alpha}{\alpha} p(x_0, y_0) & (x, y) = (x_1, y_1) \\
  p(t, t) + p(x_0, y_0) + \frac{1-\alpha}{\alpha} p(x_0, y_0) & (x, y) = (t, t).
\end{cases}
\]

Moreover,

\[
p(x_0, y_0) \phi(x_0, y_0) + p(x_1, y_1) \phi(x_1, y_1) + \phi(t, t) p(t, t)
= -p(x_0, y_0) \frac{1-\alpha}{\alpha} \phi(x_1, y_1) + p(x_1, y_1) \phi(x_1, y_1) + 0
= \left(p(x_1, y_1) - \frac{1-\alpha}{\alpha} p(x_0, y_0)\right) \phi(x_1, y_1) + 0
= p'(x_1, y_1) \phi(x_1, y_1) + \phi(t, t) p'(t, t)
\]

\(^{11}\)In the notation of Claim 5, let \(s = \delta_{(t,t)}\), \(q = \alpha \delta_{(x_0,y_0)} + (1-\alpha) \delta_{(x_1,y_1)}\), \(\lambda = 1 - p(x_0, y_0) - p(x_1, y_1)\), and

\[
r(x, y) = \begin{cases} 
  
  \frac{p(x, y)}{1-p(x_0, y_0)-p(x_1, y_1)} & (x, y) \notin \{(x_0, y_0), (x_1, y_1)\} \\
  0 & \text{otherwise}.
\end{cases}
\]

\(^{12}\)In the notation of Lemma 5, let \(s = \delta_{(t,t)}\), \(q = \alpha \delta_{(x_0,y_0)} + (1-\alpha) \delta_{(x_1,y_1)}\), \(\lambda = 1 - p(x_0, y_0) - \frac{\alpha}{1-\alpha} p(x_0, y_0)\), and

\[
r(x, y) = \begin{cases} 
  \frac{p(x, y)}{1-p(x_0, y_0)-p(x_1, y_1)} & (x, y) \notin \{(x_0, y_0), (x_1, y_1)\} \\
  0 & (x, y) = (x_0, y_0) \\
  \frac{p(x_1, y_1)}{1-p(x_0, y_0)-p(x_1, y_1)} & (x, y) = (x_1, y_1)
\end{cases}
\]
A similar equivalence can be obtained if \( \alpha = X \) and \( p = X \). Let \( \Delta \) joint lotteries with marginals \( \text{supp}^X \). By Completeness, the support is finite, by repeating this procedure a finite number of times, we are going to obtain a \( \tilde{p} \in \Delta (X \times X) \) that falls in one of the first two cases, and such that \( p \in P \iff \tilde{p} \in P \) and completing the proof.

**Proof of Proposition 1** Let \( p_1, p_2, p_3 \in \Delta (X) \) with \( p_1 \succeq^P p_2 \) and \( p_2 \succeq^P p_3 \). Let \( p = p_1 \times p_3, q = p_2 \times p_3 \), and let \( r \) be such that \( r_1 = p_1 \) and \( r_2 = p_3 \). Then, \( p, q \in P \) by definition of \( \succeq^P \) and \( r \in P \) by Transitivity of \( P \). Since \( r \) was chosen arbitrarily among the joint lotteries with marginals \( p_1 \) and \( p_2 \), the result follows.

\(
\succeq^P \text{satisfies Transitivity and Completeness} \Rightarrow P \text{satisfies Transitivity and Completeness}
\)

Let \( p, q, r \in \Delta (X \times X) \), with \( p_2 = q_1 \), \( r_1 = p_1 \), and \( r_2 = q_2 \). Then, Completeness of \( \succeq^P \) implies that \( r_1 = p_1 \succeq^P p_2 = q_1 \succeq^P q_2 = r_2 \), and Transitivity of \( \succeq^P \) implies \( r_1 \succeq^P r_2 \), and the definition of \( \succeq^P \) implies \( r \in P \), that is \( P \) satisfies Transitivity. Moreover, \( P \) satisfies Completeness by Lemma 1.

\( P \text{satisfies Transitivity and Completeness} \Rightarrow \succeq^P \text{satisfies Transitivity and Completeness} \)

That \( \succeq^P \) satisfies Transitivity follows the first part. For Completeness, let \( p_1, p_2 \in \Delta (X) \). Define \( p \) as the product measure \( p = p_1 \times p_2 \in \Delta (X \times X) \). By Completeness of \( P \), either \( p \in P \) or \( \tilde{p} \in P \). If \( p \in P \), let \( r \in \Delta (X \times X) \) be an arbitrary element of \( \Delta (X \times X) \) such that \( r_1 = p_1 \) and \( r_2 = p_2 \), and define \( q = p_2 \times p_2 \). By Completeness, \( q \in P \), and by Transitivity \( \tilde{p} \in P \) and \( q \in P \) together imply that \( r \in P \). Since \( r \) was chosen arbitrarily, \( p_1 \succeq^P p_2 \). Similarly, if \( p \notin P \), by Completeness \( \tilde{p} = p_2 \times p_1 \in P \). But then, let \( r \in \Delta (X \times X) \) be an arbitrary element of \( \Delta (X \times X) \) such that \( r_1 = p_2 \) and \( r_2 = p_1 \), and define \( q = p_1 \times p_1 \). By Completeness, \( q \in P \), and by Transitivity \( \tilde{p} \in P \) and \( q \in P \) together imply that \( r \in P \). Since \( r \) was chosen arbitrarily, \( p_2 \succeq^P p_1 \). Therefore, \( \succeq^P \) satisfies Completeness.

**Proof of Proposition 2** First, notice that by Theorem 1 \( P \) admits a representation as in (4).

2. \( \Rightarrow \) 5. Let \( x_0 \in X \). Define \( u (z) \) as \( \phi (z, x_0) \). Fix a pair \( (z, w) \), with \( z \geq w \). There are three cases:
• $z \geq w \geq x_0$. Apply (5) with $x = z$, $y = x' = x_0$ and $y' = w$. It reads

\[\phi(z, w) + \phi(x_0, x_0) = \phi(z, x_0) + \phi(x_0, w) \Leftrightarrow \]
\[\phi(z, w) = \phi(z, x_0) - \phi(w, x_0) \Leftrightarrow \]
\[\phi(z, w) = u(z) - u(w).\]

• $z \geq x_0 \geq w$. Apply (5) with $x = z$, $y = w$ and $x_0 = y' = x'$. Again, it reads as:

\[\phi(z, x_0) + \phi(x_0, w) = \phi(z, w) + \phi(x_0, x_0) \Leftrightarrow \]
\[\phi(z, w) = u(z) - u(w).\]

• $x_0 \geq z \geq w$. Apply (5) with $x = z$, $y = x' = x_0$ and $y' = w$. Again, it reads as:

\[\phi(x_0, x_0) + \phi(z, w) = \phi(z, x_0) + \phi(x_0, w) \Leftrightarrow \]
\[\phi(z, w) = u(z) - u(w).\]

Proving that $\phi(z, w) = u(z) - u(w)$ whenever $z \geq w$. If $w > z$

\[\phi(z, w) = -\phi(w, z) = -(u(w) - u(z)) = u(z) - u(w)\]

proving that the equality $\phi(z, w) = u(z) - u(w)$ holds for every $z, w \in X$. Therefore, we have $p \in P$ if and only if

\[\sum_{(x,y) \in X \times X} p(x, y) \phi(x, y) \geq 0 \iff \sum_{(x,y) \in X \times X} p(x, y) (u(x) - u(y)) \geq 0 \iff \sum_{x \in X} p_1(x) u(x) \geq \sum_{x \in X} p_2(x) u(x)\]

proving that $P$ admits an EU representation.

5. $\Rightarrow$ 2. Let $P$ admit an EU representation:

\[p \in P \iff \sum_{x \in X} p_1(x) u(x) \geq \sum_{x \in X} p_2(x) u(x),\]

Then

\[p \in P \iff \sum_{(x,y) \in X \times X} p(x, y) (u(x) - u(y)) \geq 0\]

and, if we let $\phi(z, w) = (u(z) - u(w))$, modularity holds: let $x, y, x', y' \in X$

\[\phi((x, y) \lor (x', y')) + \phi((x, y) \land (x', y'))\]
\[= u(x \lor x') - u(y \lor y') + u(x \land x') - u(y \land y')\]
\[= u(x) + u(x') - u(y) - u(y')\]
\[= \phi(x, y) + \phi(x', y').\]
4. $\Leftrightarrow$ 5. It is a version of the vN-M EU theorem, see, e.g., page 399 in Ok (2007).

5. $\Rightarrow$ 1. is evident by the representation, and 4. $\Rightarrow$ 3. holds trivially.

3. $\Rightarrow$ 2. and 1. $\Rightarrow$ 2. are proved by contradiction. Suppose that there exists $x, y, x', z' \in X$ such that

$$\phi ((x, y) \lor (x', y')) + \phi ((x, y) \land (x', y')) > \phi (x, y) + \phi (x', y')$$

with

$$(x \lor x') = x \text{ and } (y \lor y') = y'.$$

Then the inequality reads

$$\phi (x, y') + \phi (x', y) > \phi (x, y) + \phi (x', y'). \quad (10)$$

Choose $(z, w) \in (X \times X)$ and $\alpha \in [0, 1]$ such that

$$\alpha \phi (z, w) + (1 - \alpha) \left( \frac{\phi (x, y') + \phi (x', y)}{2} \right) > 0 > \alpha \phi (z, w) + (1 - \alpha) \left( \frac{\phi (x, y) + \phi (x', y')}{2} \right).$$

The existence of such $(z, w)$ and $\alpha$ is guaranteed by (10). Then

$$\alpha \delta_{(z, w)} + \frac{(1 - \alpha) \delta_{(x, y')}}{2} + \frac{(1 - \alpha) \delta_{(x', y)}}{2} \in \hat{P} \text{ and } \alpha \delta_{(z, w)} + \frac{(1 - \alpha) \delta_{(x, y)}}{2} + \frac{(1 - \alpha) \delta_{(x', y')}}{2} \notin P. \quad (11)$$

We now show that (11) implies that neither Completeness of $\succsim^P$ nor Transitivity of $P$ holds. For the former notice that (11) implies that neither

$$\alpha \delta_{z} + \frac{(1 - \alpha) \delta_{x}}{2} + \frac{(1 - \alpha) \delta_{x'}}{2} \succsim^P \alpha \delta_{w} + \frac{(1 - \alpha) \delta_{y}}{2} + \frac{(1 - \alpha) \delta_{y'}}{2}$$

nor

$$\alpha \delta_{w} + \frac{(1 - \alpha) \delta_{y'}}{2} + \frac{(1 - \alpha) \delta_{y}}{2} \succsim^P \alpha \delta_{z} + \frac{(1 - \alpha) \delta_{x}}{2} + \frac{(1 - \alpha) \delta_{x'}}{2}$$

holds, and $\succsim^P$ does not satisfy Completeness.

As for the latter, let

$$p = \alpha \delta_{(z, z)} + \frac{(1 - \alpha) \delta_{(x, x)}}{2} + \frac{(1 - \alpha) \delta_{(x', x')}}{2},$$

$$q = \alpha \delta_{(z, w)} + \frac{(1 - \alpha) \delta_{(x, y')}}{2} + \frac{(1 - \alpha) \delta_{(x', y)}}{2},$$

$$r = \alpha \delta_{(z, w)} + \frac{(1 - \alpha) \delta_{(x, y)}}{2} + \frac{(1 - \alpha) \delta_{(x', y')}}{2}.$$

Completeness of $P$ implies that $p \in P$, and (11) gives $q \notin P$, $r \notin P$. However, since $p_1 = r_1$, $p_2 = q_2$, and $q_2 = r_2$, Transitivity of $P$ does not hold.
Similar arguments can be used to obtain contradictions for other violations of modularity.

**Proof of Proposition 3** (Only if) Let \( p_1, p_2, p_3 \in \Delta(X), \alpha \in (0,1) \) with \( p_1 \preceq^p p_2 \). Let \( r \in \Delta(X \times X) \) be an arbitrary element of \( \Delta(X \times X) \) such that \( r_1 = \alpha p_1 + (1 - \alpha) p_3 \) and \( r_2 = \alpha p_2 + (1 - \alpha) p_3 \). Define \( p \) as the product measure \( p = p_1 \times p_2 + \Delta(X \times X) \). By definition of \( \preceq^P \), \( p \in P \). Let \( q = p_3 \times p_3 \). Then \( \alpha p + (1 - \alpha) q \in P \) by Strong Independence of \( P \), and since \( \preceq^P \) satisfies Completeness by part 1, \( \alpha p_1 + (1 - \alpha) p_3 \preceq^P \alpha p_2 + (1 - \alpha) p_3 \).

Let \( p_1, p_2, p_3 \in \Delta(X), \alpha \in (0,1) \) with \( \neg (p_1 \preceq^P p_2) \). By part 1, \( \preceq^P \) satisfies Completeness, and therefore this implies that \( p_2 \succ p_1 \). Let \( r \in \Delta(X \times X) \) be an arbitrary element of \( \Delta(X \times X) \) such that \( r_1 = \alpha p_1 + (1 - \alpha) p_3 \) and \( r_2 = \alpha p_2 + (1 - \alpha) p_3 \). Define \( p \) as the product measure \( p = p_1 \times p_2 \in \Delta(X \times X) \). By definition of \( \preceq^P \), \( p \notin P \), and \( \tilde{p} \in \hat{P} \). Let \( q = p_3 \times p_3 \). Then \( \alpha \tilde{p} + (1 - \alpha) q \in \hat{P} \) by Strong Independence of \( P \), and since \( \preceq^P \) satisfies Completeness by part 1, \( \neg (\alpha p_1 + (1 - \alpha) p_3 \preceq^P \alpha p_2 + (1 - \alpha) p_3) \).

(If) Let \( p, q \in P, \alpha \in (0,1) \). Since \( \preceq^P \) satisfies Completeness by part 1, \( p_1 \preceq^P p_2 \) and \( q_1 \preceq^P q_2 \). By Strong Independence of \( \preceq^P \)

\[
\alpha p_1 + (1 - \alpha) q_1 \preceq^P \alpha p_2 + (1 - \alpha) q_1 \quad \text{and} \quad \alpha p_2 + (1 - \alpha) q_1 \preceq^P \alpha p_2 + (1 - \alpha) q_1. \tag{12}
\]
Therefore \((\alpha p_1 + (1 - \alpha) q_1) \times (\alpha p_2 + (1 - \alpha) q_1)\) and \((\alpha p_2 + (1 - \alpha) q_1) \times (\alpha p_2 + (1 - \alpha) q_2)\) are in \( P \). But then, Transitivity implies that \( \alpha p + (1 - \alpha) q \in P \).
Now, suppose that on top of this \( q \in \hat{P} \), then by definition of \( \preceq^P \), \( q \notin P \). By Strong Independence and Completeness of \( \preceq^P \) proved in part 1, this implies that

\[
\neg (\alpha p_2 + (1 - \alpha) q_2 \preceq^P \alpha p_2 + (1 - \alpha) q_1).
\]

Therefore, there exists \( r \) such that \( r_1 = \alpha p_2 + (1 - \alpha) q_2, r_2 = \alpha p_2 + (1 - \alpha) q_1, r \notin P \). Suppose by way of contradiction that \( \alpha p + (1 - \alpha) q \notin \hat{P} \), that is \( \alpha \tilde{p} + (1 - \alpha) \tilde{q} \in P \). But then \((\alpha p_1 + (1 - \alpha) q_1) \times (\alpha p_2 + (1 - \alpha) q_1)\) in \( P \) by (12), and Transitivity implies that \( r \in P \), a contradiction.

**7.1 Salience Characterization**

**Proof of Proposition 4** (If) When \( x > y \), we have \( \phi(x, z) > \phi(y, z) \). Therefore, if \( \alpha \in (0,1) \) and \( p \in \Delta(X) \)

\[
\alpha \delta_{(y,z)} + (1 - \alpha) p \in P
\]

\[
\iff \alpha \phi(x, z) + (1 - \alpha) \sum_{(x', y') \in X \times X} p(x', y') \phi(x', y') \geq 0
\]

\[
\Rightarrow \alpha \phi(x, z) + (1 - \alpha) \sum_{(x', y') \in X \times X} p(x', y') \phi(x', y') > 0
\]

\[
\Rightarrow \alpha \delta_{(x,z)} + (1 - \alpha) p \notin \hat{P}.
\]

29
(Only if) Let \( x_1 > x_2, x_1, x_2, y \in X \). We first prove that \( \phi \) is strictly increasing in the first argument. Given the representation in (4), we have

\[
\frac{\delta_{(x_2,y)}}{2} + \frac{\delta_{(y,x_2)}}{2} \in P.
\]

Then by Monotonicity

\[
\frac{\delta_{(x_1,y)}}{2} + \frac{\delta_{(y,x_2)}}{2} \in \hat{P}
\]

and given the representation in (4) this implies \( \phi (x_1, y) > \phi (x_2, y) \). To see that \( \phi \) is strictly decreasing in the second argument, notice that by skew symmetry:

\[
\phi (x_1, y) > \phi (x_2, y) \Rightarrow -\phi (y, x_1) > -\phi (y, x_2) \Rightarrow \phi (y, x_1) < \phi (y, x_2)
\]

proving the statement.

**Proof of Proposition 5** Given a Smooth Salience representation, let \( \phi (x, y) = \sigma (x, y) (u(x) - u(y)) \).

By the Symmetry axiom for \( \sigma \), we have

\[
\phi (x, y) = \sigma (x, y) (u(x) - u(y)) = \sigma (y, x) (u(x) - u(y)) = -\sigma (y, x) (u(y) - u(x)) = -\phi (y, x)
\]

proving that \( \phi \) is skew-symmetric. Then the Smooth criterion satisfies Completeness, Strong Independence, and Archimedean Continuity by Theorem 1. Since \( u \) is increasing, and \( \sigma \) satisfies Ordering, the decision criterion satisfies Monotonicity by Proposition 4.

**Proof of Proposition 6 First part:**

(Only If) Let \( x, w, y, z \in \mathbb{R} \) be such that

\[
\{(x,y), (w,z), (x,z), (w,y)\} \subseteq \oplus,
\]

that is, \( (x \wedge w) \geq (y \vee z) \). Assume for definiteness that \( x \geq w \) and \( y \geq z \). Then, by the Ordering axiom

\[
\frac{1}{4}\delta_{(x,z)} + \frac{1}{4}\delta_{(w,y)} + \frac{1}{4}\delta_{(z,w)} + \frac{1}{4}\delta_{(y,x)} \in P.
\]

That is

\[
\frac{1}{4}\phi (x, z) + \frac{1}{4}\phi (w, y) + \frac{1}{4}\phi (z, w) + \frac{1}{4}\phi (y, x) \geq 0 \Leftrightarrow \\
\phi (x, y) + \phi (w, z) \leq \phi (x, z) + \phi (w, y).
\]

Now, suppose that in addition \( x > w \) and \( y > z \) and \( P \) satisfies Strict Ordering. Then,

\[
\frac{1}{4}\delta_{(x,z)} + \frac{1}{4}\delta_{(w,y)} + \frac{1}{4}\delta_{(z,w)} + \frac{1}{4}\delta_{(y,x)} \in \hat{P}
\]

30
that is
\[ \phi(x, y) + \phi(w, z) < \phi(x, z) + \phi(w, y) \]
as wanted.

(If) Similarly, let \( x_H \geq x_L \geq y_H \geq y_L \), submodularity on \( \oplus \) implies
\[
\frac{1}{4}\phi(x_H, y_H) + \frac{1}{4}\phi(x_L, y_L) + \frac{1}{4}\phi(y_H, x_H) + \frac{1}{4}\phi(y_L, x_L) \leq 0
\]
that is
\[
\frac{1}{4}\delta(x_H, y_H) + \frac{1}{4}\delta(x_L, y_L) + \frac{1}{4}\delta(y_H, x_H) + \frac{1}{4}\delta(y_L, x_L) \in P.
\]

If in addition \( x_H > x_L \) and \( y_H > y_L \), strict submodularity of \( \phi \) on \( \{(x, y) : x \geq y\} \) implies
\[
\phi(x_H, y_H) + \phi(x_L, y_L) > \phi(x_H, y_L) + \phi(x_L, y_H)
\]
that is
\[
\frac{1}{4}\delta(x_H, y_H) + \frac{1}{4}\delta(x_L, y_L) + \frac{1}{4}\delta(y_H, x_H) + \frac{1}{4}\delta(y_L, x_L) \in P.
\]

Second part:

1. Let \( x_H \geq x_L \geq y_H \geq y_L \). Ordering of \( \sigma \) implies that either
\[
\sigma(x_H, y_L) \geq \sigma(x_L, y_L) \geq \sigma(x_H, y_H) \geq \sigma(x_L, y_H)
\]
or
\[
\sigma(x_H, y_L) \geq \sigma(x_H, y_H) \geq \sigma(x_L, y_L) \geq \sigma(x_L, y_H).
\]

Suppose we are in the first case, and that all the inequalities hold strict. We have to prove that
\[
u(x_H) - u(y_L) + \beta^3(u(x_L) - u(y_H)) - \beta(u(x_H) - u(y_H)) - \beta^2(u(x_L) - u(y_L)) \geq 0
\]
Consider the LHS as a continuous function \( f \) of \( \beta \). Notice that
\[
\begin{align*}
f(1) &= 0 \quad \text{and} \quad f(0) = u(x_H) - u(y_L) \geq 0.
\end{align*}
\]
In addition,
\[
f'(\beta) = 3\beta^2(u(x_L) - u(y_H)) - (u(x_H) - u(y_H)) - 2\beta(u(x_L) - u(y_L))
\]
so that
\[
f'(1) = u(x_L) - u(x_H) - 2(u(y_H) - u(y_L)) \quad \text{and} \quad f'(0) = u(x_H) - u(y_H) < 0.
\]
Therefore, a necessary condition for having a $\beta^* \in (0, 1)$ with $f'(\beta) = 0$ is that there is a $\beta$ such that $f''(\beta) < 0$ but notice that

$$f''(\beta) = 3(u(x_L) - u(y_H)) \geq 0.$$ 

Therefore $f'(\beta) \geq 0$ always, with a strict inequality if $\beta < 1$, so that Strict Ordering holds in that case. To see why ties are not problematic, suppose that for example

$$\sigma(x_H, y_L) = \sigma(x_L, y_L)$$

then, Ordering of $\sigma$ implies $x_H = x_L$, but then

$$\frac{1}{4}\delta(x_{H},y_{H}) + \frac{1}{4}\delta(x_{L},y_{L}) + \frac{1}{4}\delta(y_{H},x_{H}) = \frac{1}{4}\delta(x_{H},y_{H}) + \frac{1}{4}\delta(y_{H},x_{H})$$

and symmetry of $\sigma$ immediately delivers the desired equality. The other case is proved similarly.

2. Consider first the case of two intervals with an extreme in common $x' < y'$, $x < y$ with $[x', y'] \subset [x, y]$, say $x = x'$. Therefore, we are considering the case $x < y' < y$. Suppose, by way of contradiction, that $\sigma(x, y) < \sigma(x, y')$. If also $\sigma(x, y) \leq \sigma(y, y')$, a contradiction with the Ordering Axiom can be immediately obtained. For instance, if $\sigma(x, y) < \sigma(x, y') < \sigma(y, y')$, the $(\beta, \sigma)$ representation implies that

$$\frac{1}{4}\delta(y,x) + \frac{1}{4}\delta(y',y') + \frac{1}{4}\delta(y',y) + \frac{1}{4}\delta(x,y') \leq P \Leftrightarrow$$

$$\frac{1}{4}(u(y) - u(x))\beta^2 - \frac{1}{4}(u(y) - u(y'))\beta - \frac{1}{4}(u(y') - u(x)) \geq 0 \Leftrightarrow$$

$$(u(y) - u(x))\beta^2 - (u(y) - u(y'))\beta - (u(y') - u(x)) \geq 0 \Leftrightarrow$$

$$u(y)(\beta^2 - \beta) - u(x)(\beta^2 - 1) + u(y')(\beta - 1) \geq 0 \Leftrightarrow$$

$$-\beta u(y) + (\beta - 1)u(x) - u(y') \geq 0 \Leftrightarrow$$

$$\beta(u(x) - u(y)) + u(x) - u(y') \geq 0$$

but the last inequality is false since $u$ is strictly increasing. Therefore, let

$$\sigma(y, y') < \sigma(x, y) < \sigma(x, y').$$

Then, consider the lottery

$$\frac{1}{4}\delta(y,x) + \frac{1}{4}\delta(y',y') + \frac{1}{4}\delta(y',y) + \frac{1}{4}\delta(x,y')$$

The case $y = y'$ is completely analogous.
that is in \( P \) by the Ordering Axiom. Then, the \((\beta, \sigma)\) representation and (13) imply that
\[
\begin{align*}
-u(y') + u(x) + \beta (u(y) - u(x)) - \beta^2 (u(y) - u(y')) & \geq 0 \iff \\
\beta (1 - \beta) u(y) - (1 - \beta^2) u(y') + (1 - \beta) u(x) & \geq 0 \iff \\
\beta u(y) - (1 + \beta) u(y') + u(x) & \geq 0 \iff \\
\beta (u(y) - u(y')) & \geq u(y') - u(x)
\end{align*}
\]
a contradiction with the hypothesis. Finally, notice that if the two intervals have no extreme value in common, that is \( x' < y', x < y \) with \( x' > x, y' < y \) we have just proved that
\[
\sigma (x, y) > \sigma (x', y')\quad \text{and} \quad \sigma (x, y') > \sigma (x', y'),
\]
and therefore \( \sigma (x, y) > \sigma (x', y') \).

**Proof of Proposition 7** (If) Let \( x \geq y \geq 0 \), and \( k \in \mathbb{R}_+ \), by the Gradient Theorem (see, e.g., Theorem 10.33 in Rudin 1976), and letting \( \gamma ([x, y], (x + k, y + k)] \) be the straight line between \((x, y)\) and \((x + k, y + k)\), we have that
\[
\phi (x + k, y + k) = \phi (x, y) + \int_{ ((x, y), (x + k, y + k)] } \nabla_{(1,1)} \phi (z, z) \, dz
\]
\[
= \phi (x, y) + \int_{ \gamma ([x, y], (x + k, y + k)] } \nabla_{(1,1)} \phi (z_1, z_2) \, dz \leq \phi (x, y)
\]
so that
\[
\frac{\phi (x, y) + \phi (y + k, x + k)}{2} \geq 0
\]
that is
\[
\left( (x, y), \frac{1}{2}; (y + k, x + k), \frac{1}{2} \right) \in P.
\]
(Only if) It is enough to show that for all \( x \geq y \) and \( k \geq 0 \)
\[
\phi (x + k, y + k) \geq \phi (x, y).
\]
but since Diminishing Sensitivity implies
\[
\frac{\phi (x, y) + \phi (y + k, x + k)}{2} \geq 0
\]
the property immediately follows from Skew Symmetry.

Suppose \( P \) admits a \( \delta - \sigma \) representation and \( \sigma \) satisfies Diminishing Sensitivity. Then, let \( x \geq y \geq 0 \), and \( k \in \mathbb{R}_+ \), \( p = \left( (x, y), \frac{1}{2}; (y + k, x + k), \frac{1}{2} \right) \). Then, Diminishing Sensitivity of \( \sigma \) implies that
\[
\sigma (x, y) \geq \sigma (x + k, y + k)
\]
and therefore
\[ \hat{\sigma}_p(x, y) \leq \hat{\sigma}_p(x + k, y + k). \]
Since \( u \) is concave, \( u(x) - u(y) \geq u(x + k) - u(y + k) \), and therefore
\[ (u(x) - u(y)) \beta^{\hat{\sigma}_p(x, y)} + (u(x + k) - u(y + k)) \beta^{\hat{\sigma}_p(x + k, y + k)} \geq 0. \]
Since \( k \in \mathbb{R}^{++} \) implies \( \hat{\sigma}_p(x, y) < \hat{\sigma}_p(x + k, y + k) \), the previous inequality is strict whenever \( k \in \mathbb{R}^{++} \).

Conversely, let \( P \) admit a \( \delta - \sigma \) representation and satisfy Diminishing Sensitivity. Fix \( x > y \geq k > 0 \). Strict Diminishing Sensitivity of \( P \) implies that
\[ p = \left( (x, y - k), \frac{1}{2}; (y, x + k), \frac{1}{2} \right) \in \hat{P}. \]
Since \( P \) admit a \( \delta - \sigma \) representation with linear \( u \), we have
\[ (x - y + k) \beta^{\hat{\sigma}_p(x, y-k)} + (y - k - x) \beta^{\hat{\sigma}_p(y, x+k)} > 0. \]
Since \( \beta < 1 \), the previous inequality holds if and only if \( \hat{\sigma}_p(x, y - k) > \hat{\sigma}_p(y, x + k) \), that by Symmetry and definition of \( \hat{\sigma}_p \) holds if and only if \( \sigma(x, y - k) > \sigma(x + k, y) \).

**Proof of Proposition 8** Let \( P \) admit an EU representation with a strictly increasing utility function \( u \). That is,
\[ p \in P \iff \sum_{(x, y) \in X \times X} p(x, y) (u(x) - u(y)) \geq 0 \iff \sum_{x \in X} p_1(x) u(x) \geq \sum_{y \in X} p_2(y) u(y). \]
(If) Let \( x \geq y \geq 0 \) and \( k \geq 0 \). Consider the two marginal distributions
\[ q = \left( x, \frac{1}{2}; y + k, \frac{1}{2} \right) \text{ and } q' = \left( x + k, \frac{1}{2}; y, \frac{1}{2} \right) \]
Notice that \( q' \) is a mean-preserving spread of \( q \), since \( q' \) can be obtained by further randomizing each realization \( z \) of \( q \) with the additional random term \( h_z \) with
\[ h_x = \left( k, \frac{(x - y)}{(x - y) + k}; (y - x), \frac{k}{(x - y) + k} \right) \]
and
\[ h_{y+k} = \left( (x - y), \frac{k}{(x - y) + k}; -(y - x), \frac{(x - y)}{(x - y) + k} \right). \]
By Proposition 6.C.1 and 6.D.2. in Mas-Colell, Whinston, and Green,
\[ \sum_{z \in X} q(z) u(z) \geq \sum_{z \in X} q'(z) u(z) \]
34
by rearranging the terms
\[ \frac{1}{2} (u(x) - u(y)) + \frac{1}{2} (u(y + k) - u(x + k)) \geq 0 \]
or
\[ \left( (x, y), \frac{1}{2}, (y + k, x + k), \frac{1}{2} \right) \in P \]
and Diminishing Sensitivity holds.

(Only If) Let \( x_0 \geq y_0 \geq 0 \). By Diminishing Sensitivity
\[ \left( \left( \frac{x_0 + y_0}{2}, y_0 \right), \frac{1}{2}, \left( \frac{x_0 + y_0}{2}, x_0 \right), \frac{1}{2} \right) \in P \]
that is
\[ u \left( \frac{x_0 + y_0}{2} \right) \geq \frac{u(x_0) + u(y_0)}{2} \]
proving the midpoint concavity of \( u \) on the set of positive real numbers. Since \( u \) is strictly increasing, it is measurable. Since the Sierpinski Theorem implies that a midpoint concave and measurable function is concave, the DM is risk-averse on that range. \[ \blacksquare \]

Proof of Proposition 9 Let \( P \) satisfy Reflexivity and admit a \((\beta, \sigma)\) representation with linear \( u \), \( \beta < 1 \), and \( \sigma \) satisfying Ordering. Suppose, by way of contradiction, that there exist \( x, y, w, z \in \mathbb{R}_+ \) such that \( \sigma(x, y) > \sigma(w, z) \) and \( \sigma(-x, -y) < \sigma(-w, -z) \). We consider the case in which \( \left( (x, y), \frac{1}{2}; (w, z), \frac{1}{2} \right) \in \hat{P} \), since the other cases are proved similarly. Denote \( \max \{ x, y, w, z \} \) as \( M \). Notice that
\[ \left( (x, y), \frac{1}{2}; (w, z), \frac{1}{2} \right) \notin P \iff (x - y) + \beta(w - z) < 0. \]
Therefore
\[ \alpha := \frac{M}{M - \frac{\beta(x-y)+\beta^2(w-z)}{2}} \in (0, 1). \]
By construction, and since \( \sigma \) satisfies Ordering,
\[ \alpha \left( (x, y), \frac{1}{2}; (w, z), \frac{1}{2} \right) + (1 - \alpha) \delta(\max\{x,y,w,z\},0) \in P \setminus \hat{P}. \]  \[ (14) \]
However, \( \sigma(-x, -y) < \sigma(-w, -z) \) implies that:
\[ \alpha \left( (x, y), \frac{1}{2}; (w, z), \frac{1}{2} \right) + (1 - \alpha) \delta(0, \min\{-x,-y,-w,-z\}) \in P \setminus \hat{P}. \]
To see this, notice that since $P$ admits a $(\beta, \sigma)$ representation with linear $u$ and $\sigma$ satisfying Ordering, 14 is equivalent to
\[
\begin{align*}
\alpha \frac{\beta (x - y) + \beta^2 (w - z)}{2} + (1 - \alpha) M &= 0 \iff \\
\alpha \frac{-\beta^2 (x - y) - \beta (w - z)}{2} - (1 - \alpha) M &\neq 0 \iff \\
\alpha \left( (-x, -y), \frac{1}{2}; (-w, -z), \frac{1}{2} \right) + (1 - \alpha) \delta_{\{0, \min\{x, y, w, z\}\}} &\in P \setminus \hat{P}.
\end{align*}
\]

Let $P$ satisfy Reflexivity and admit a $(\beta, \sigma)$ representation with linear $u$, $\beta < 1$, and $\sigma$ satisfying Ordering and Reflexivity. Let $x, y, w, z \in \mathbb{R}_+$, $\alpha \in [0, 1]$. Without loss of generality, let $\sigma (x, y) > \sigma (w, z)$. Therefore
\[
\begin{align*}
\alpha \left( (x, y), \frac{1}{2}; (w, z), \frac{1}{2} \right) + (1 - \alpha) \delta_{\{0, \max\{x, y, w, z\}\}} &\in P \setminus \hat{P} \iff \\
\alpha \frac{\beta (x - y) + \beta^2 (w - z)}{2} - (1 - \alpha) \max \{x, y, w, z\} &= 0 \iff \\
\beta \frac{(-x + y) + \beta^2 (-w + z)}{2} - (1 - \alpha) \min \{-x, -y, -w, -z\} &= 0 \iff \\
\alpha \left( (-x, -y), \frac{1}{2}; (-w, -z), \frac{1}{2} \right) + (1 - \alpha) \delta_{\{0, \min\{-x, -y, -w, -z\}\}} &\in P \setminus \hat{P}
\end{align*}
\]

where the last coimplication is due to the fact that $\sigma$ satisfies Ordering and Reflexivity. □

**Proof of Proposition 10** Let $u$ be a vN-M utility index representing $P$ such that $u (0) = 0$, and suppose that $P$ is risk-averse for lotteries with values in $[a, b] \subseteq \mathbb{R}_+$. Let $-b \leq -x \leq -y \leq -a$, $\alpha = \frac{u (x)}{u (x) + u (\frac{x + y}{2}) - u (\frac{y}{2})}$. Since $u$ is concave and increasing, $\alpha \in [0, 1]$. By definition
\[
\alpha u \left( \frac{x + y}{2} \right) - \frac{\alpha u (x)}{2} - \frac{\alpha u (y)}{2} - (1 - \alpha) u (x) = 0
\]

that is
\[
\alpha \left( \left( \frac{x + y}{2}, \frac{x + y}{2}, \frac{1}{2}; \frac{x + y}{2}, y, \frac{1}{2} \right) + (1 - \alpha) \delta_{\{0, \max\{x, y, z + y\}\}} \right) \in P \setminus \hat{P}.
\]

Then, Reflexivity implies that
\[
\alpha \left( \left( \frac{-x + y}{2}, -x, \frac{1}{2}; -x + y, -y, \frac{1}{2} \right) + (1 - \alpha) \delta_{\{0, \min\{-x, -y, z + y\}\}} \right) \in P \setminus \hat{P}
\]

that is
\[
\alpha u \left( \frac{-x + y}{2} \right) - \frac{\alpha u (-x)}{2} - \frac{\alpha u (-y)}{2} - (1 - \alpha) u (-x) = 0
\]

36
Since $u(-x) < 0$, this means that

$$u\left(-\frac{x + y}{2}\right) - \frac{u(-x)}{2} - \frac{u(-y)}{2} < 0 \iff u\left(-\frac{x + y}{2}\right) < \frac{u(-x)}{2} + \frac{u(-y)}{2}.$$ 

Since $-x$ and $-y$ were chosen arbitrarily in $[-b, -a]$, $u$ is midpoint convex in that part of its domain. Since it is also increasing, it is measurable, and by the Sierpinski Theorem it is convex in that part of its domain, proving the statement.

### 7.2 Proof of Theorem 2

To prove the main representation result, we need to use an ancillary binary relation. For every $\varepsilon > 0$, denote as $\Delta^+_{\varepsilon} = \{F \in \text{DDF} : F(0) < 1 - \varepsilon, \lim_{x \to 0^-} F(0) = 1\}$. Define the binary relation $\succ^+_{\varepsilon}$ on $\Delta^+_{\varepsilon}$ in the following way

$$F \succ^+_{\varepsilon} G \iff \left(\exists b \in \mathbb{R}, \delta \in (0, \varepsilon], k > 1 \text{ such that } F^b_{\delta} \in \hat{P}, G^b_{\delta} \notin \hat{P}\right),$$

where

$$F^b_{\delta}(a) = \begin{cases} 1 - \delta & b \leq a < 0, \\ F(ka) & \text{otherwise.} \end{cases}$$

In words, $F^b_{\delta}(a)$ is the DDF obtained by: 1) shifting a weight $\delta$ from the outcomes where the two lotteries coincide to the ones where the second lottery pays $-b$ more, 2) Shifting the weight from the first outcome being larger by $ka$ to the first outcome being larger by $a$. As usual, $\succ^+_{\varepsilon}$ is defined as

$$F \succ^+_{\varepsilon} G \iff \gamma(G \succ^+_{\varepsilon} F).$$

#### Lemma 2 $\succ^+_{\varepsilon}$ is transitive and complete.

**Proof** *(Transitivity)* Let $F \succ^+_{\varepsilon} G$ and $G \succ^+_{\varepsilon} H$. Suppose by way of contradiction that $H \succ^+_{\varepsilon} F$. Then, there exists $b \in \mathbb{R}, \delta \in (0, \varepsilon], k > 1$ such that $H^b_{\delta} \in \hat{P}, F^b_{\delta} \notin \hat{P}$. By $F \succ^+_{\varepsilon} G, G^b_{\delta} \notin \hat{P}$. But then, $H \succ^+_{\varepsilon} G$, a contradiction.

**(Completeness)** Let $F \succ^+_{\varepsilon} G$. Therefore, there exist $b \in \mathbb{R}, \delta \in (0, \varepsilon], k > 1$ such that

$$F^b_{\delta} \in \hat{P} \text{ and } G^b_{\delta} \notin \hat{P}.$$ 

But then, choose arbitrary $b \in \mathbb{R}, \delta \in (0, \varepsilon], k > 1$. By letting $F^b_{\delta} = \tilde{F}, G^b_{\delta} = \tilde{G}, F^b_{\delta} = \tilde{F}$ and $G^b_{\delta} = \tilde{G} - \tilde{H} = \tilde{H} = I_{(-\infty, 0)}$ and $\alpha = \frac{1}{k}$ it follows from Dual Independence that $\gamma\left(F^b_{\delta} \notin \hat{P}, G^b_{\delta} \in \hat{P}\right)$ and therefore $\gamma(G \succ^+_{\varepsilon} F)$, proving the statement.
Lemma 3  If $F,G \in \Delta_+^\delta$ with $F \geq_{\text{FOSD}} G$, then $F \succ_+^\delta G$.

Proof Let $b \in \mathbb{R}$, $\delta \in (0,\varepsilon]$, $k > 1$. Then, $F \geq_{\text{FOSD}} G$ implies that

$$F_{\delta}^{bk} \geq_{\text{FOSD}} G_{\delta}^{bk}.$$  

By DDF Monotonicity, $G_{\delta}^{bk} \in \hat{P}$ would imply that $F_{\delta}^{bk} \in \hat{P}$ as well, and therefore

$$\gamma(G \succ_+^\delta F) \text{ or } F \succ_+^\delta G.$$  

Lemma 4  If $F,G,H,H' \in \Delta_+^\delta$ with $F \succ_+^\delta G$ there exists $\theta > 0$ such that $\|F - H\| < \theta$ and $\|G - H'\| < \theta$ implies $H \succ^+ H'$.

Proof Let $b \in \mathbb{R}$, $\delta \in (0,\varepsilon]$, $k > 1$ be such that

$$F_{\delta}^{bk} \in \hat{P}, \ G_{\delta}^{bk} \notin \hat{P}.$$  

Then, by L1 Continuity and Completeness, there exists a $\tau$ such that

$$\|F' - F_{\delta}^{bk}\| < \tau \text{ implies } F' \in \hat{P},$$  

and

$$\|G' - G_{\delta}^{bk}\| < \tau \text{ implies } G' \notin \hat{P}.$$  

However, since for all $b \in \mathbb{R}$, $\delta \in (0,\varepsilon]$, $k > 1$

$$\|F_{\delta}^{bk} - G_{\delta}^{bk}\| \leq \|F - G\|$$  

and the result follows.

Lemma 5  If $F,G,H \in \Delta_+^\delta$, $\gamma \in [0,1]$ and $F \succ_+^\delta G$ then

$$\gamma F \boxplus (1 - \gamma) H \succ_+^\delta \gamma F \boxplus (1 - \gamma) H.$$  

Proof The claim is obvious for $\gamma = 1$, and it follows from Lemma 2 for $\gamma = 0$. Therefore, let $\gamma \in (0,1)$. Suppose, by way of contradiction, that there exist $b \in \mathbb{R}$, $\delta \in (0,\varepsilon]$, $k > 1$ such that

$$(\gamma F \boxplus (1 - \gamma) H)_{\delta}^{bk} \notin \hat{P}, \ (\gamma G \boxplus (1 - \gamma) H)_{\delta}^{bk} \in \hat{P}.$$  

By L1-Continuity, there exists $\delta^*$ such that

$$F_{\delta^*}^{bk} \in P \setminus \hat{P}.$$  

But then if we apply Dual Independence with $\hat{F} = (\gamma F \boxplus (1 - \gamma) H)_{\delta}^{bk}$, $\hat{G} = (\gamma F \boxplus (1 - \gamma) H)_{\delta}^{bk}$, $\hat{F} = F_{\delta^*}^{bk}$, $\hat{G} = G_{\delta^*}^{bk}$, we obtain

$$G_{\delta^*}^{bk} \in \hat{P},$$  

a contradiction with $F \succ_+^\delta G$.  

38
Lemma 6 There exists a continuous and nondecreasing $f_\varepsilon : [0, 1 - \varepsilon] \rightarrow \mathbb{R}$ such that for all $F, G \in \Delta_+^+$

$$F \succsim^+_\varepsilon G \iff \int_0^{h-l} f_\varepsilon (F(t)) \, dt \geq \int_0^{h-l} f_\varepsilon (G(t)) \, dt.$$ 

Moreover, $f_\varepsilon$ is unique up to a positive affine transformation and

$$(h-l) f_\varepsilon (x) = x f_\varepsilon (1) + (h-l-x) f_\varepsilon (0).$$

**Proof** It immediately follows from Lemmata 2 2 2 2 and Theorem 1 in Yaari 1987.

Notice that if $\varepsilon < \varepsilon$, $\Delta_+^+ \supseteq \Delta_+^+$, and the restriction of $\succsim^+_\varepsilon$ on $\Delta_+^+$ coincides with $\succsim^+_\varepsilon$, and the uniqueness part of Lemma 6 guarantees that $f_\varepsilon$ can be chosen such that it agrees with $f_\varepsilon$ on $[0, 1 - \varepsilon]$. Therefore, $f_+ \varepsilon$ can be unambiguously defined as $f_+ (x) = f_\varepsilon (x)$ for an arbitrary $\varepsilon \leq 1 - x$. A completely symmetric argument allow us to obtain a continuous and nonincreasing $f_-$. Choose $l-h < b < 0 < b < h-l$ such that $(b, \frac{1}{2}; b, \frac{1}{2}) \in \mathcal{P}\setminus\mathcal{P}$, whose existence is guaranteed by DDF Monotonicity and L1-Continuity. Then, since $f_-$ is unique up to a positive affine transformation, we can fix

$$(b-l-h) f_- (1) + f_- \left( \frac{1}{2} \right) b = f_+ \left( \frac{1}{2} \right) b + f_- (0) (h-l).$$

Given this, it Dual Independence guarantees that $f_+ \varepsilon$ and $f_- \varepsilon$ indeed represent the preferences as in (9).

### 7.3 Multiple Alternatives

**Proof of Proposition 11** Consider the problem

$$\max_{\alpha \in \Delta(A)} \min_{\beta \in \Delta(A)} \sum_{f \in A} \sum_{g \in A} \sum_{(x,y)} \alpha (f) \beta (g) p_{f,g} (x,y) \phi (x,y)$$

$$= \min_{\beta \in \Delta(A)} \max_{\alpha \in \Delta(A)} \sum_{f \in A} \sum_{g \in A} \sum_{(x,y)} \alpha (f) \beta (g) p_{f,g} (x,y) \phi (x,y)$$

$$= \min_{\beta \in \Delta(A)} \max_{\alpha \in \Delta(A)} \sum_{f \in A} \sum_{g \in A} \sum_{(x,y)} \alpha (f) \beta (g) \left( - \sum_{(x,y)} p_{g,f} (x,y) \phi (x,y) \right)$$

$$= \min_{\beta \in \Delta(A)} \left( - \max_{\alpha \in \Delta(A)} \left( \sum_{f \in A} \sum_{g \in A} \sum_{(x,y)} \alpha (f) \beta (g) p_{g,f} (x,y) \phi (x,y) \right) \right)$$

$$= - \max_{\beta \in \Delta(A)} \min_{\alpha \in \Delta(A)} \left( \sum_{f \in A} \sum_{g \in A} \sum_{(x,y)} \alpha (f) \beta (g) p_{g,f} (x,y) \phi (x,y) \right)$$

$$= - \max_{\alpha \in \Delta(A)} \min_{\beta \in \Delta(A)} \left( \sum_{f \in A} \sum_{g \in A} \sum_{(x,y)} \alpha (f) \beta (g) p_{g,f} (x,y) \phi (x,y) \right)$$
where the first equality follows from von Neumann min-max Theorem (see, e.g., Sion 1958), the second by Skew Symmetry of $\phi$, and the other equality are obtained with simple algebra. Therefore, $\max_{\alpha \in \Delta(A)} \min_{\beta \in \Delta(A)} \sum_{f \in A} \sum_{g \in A} \sum_{(x,y)} \alpha(f) \beta(g) p_{f,g}(x,y) \phi(x,y) = 0$, that is there exists $\alpha \in \Delta(A)$ such that for all $\beta \in \Delta(A)$

$$\sum_{f \in A} \sum_{g \in A} \sum_{(x,y)} \alpha(f) \beta(g) p_{f,g}(x,y) \phi(x,y) \geq 0.$$
References


