Correlation Made Simple: Applications to Salience and Regret Theory*

Giacomo Lanzani
MIT
December 10, 2020

Abstract

In this work, we offer a straightforward axiomatization for decision criteria under risk in which the correlation between the alternatives considered plays a role. Extending to the nondeterministic case the techniques of conjoint measurement, we can formally identify Transitivity as the vN-M axiom that has to be relaxed to allow for these richer patterns of behavior. To illustrate the advantages of our modeling choice, we provide a simple axiomatization for the Salience Theory model within our general framework. This approach allows us to single out their Ordering property as the feature that brings Salience Theory outside the Expected Utility realm. Finally, even if in a set of three or more alternatives an optimal choice may not exist under our intransitive decision criterion, we prove the existence of an optimal randomization over the available options.

1 Introduction

Correlation between risky alternatives under consideration can play a significant role in decisions. On the one hand, it may be relevant because the Decision Maker (hereafter DM) cares about the received outcome and the counterfactual that he would

---

*I am particularly indebted to Simone Cerreia-Vioglio, Roberto Corrao, Drew Fudenberg, Nicola Gennaioli, and Fabio Maccheroni. I am also grateful to Pierpaolo Battigalli, Colin Camerer, Enrico De Magistris, Pawel Dzewulski, Salvatore Greco, Mats Köster, Massimo Marinacci, Paulo Natenzon, Pietro Ortoleva, Harry Pei, Antonio Rangel, Aldo Rustichini, Marciano Siniscalchi, Philipp Strack, Peter Wakker, Alexander Woltzky, and Muhamet Yildiz for helpful comments and suggestions. The financial support of Guido Cazzavillan Scholarship is gratefully acknowledged.
have received had he chosen differently, a channel emphasized by Regret Theory. On the other hand, the correlation structure can determine the attention and weight that the DM allocates to the various contingencies, as emphasized by Salience Theory. In this paper, we study these possibilities from an axiomatic perspective.

We provide a simple axiomatization for a general class of risk preferences such that the correlation between the alternatives under consideration is relevant for the DM. The motivation is two-fold. First, we show that our general framework nests 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, 2013) as particular cases, so that we can characterize them in terms of additional testable axioms. Then, the study of these axioms allows us to understand better where they depart from the preexisting theories. Second, we use this axiomatization to provide new insights into the difference between classical models for 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).

We accomplish these goals by taking a different route than the one followed in the usual axiomatizations of correlation sensitive preferences (see, e.g., Fishburn 1989, Sugden 1993, and Diecidue and Somaundaram 2017). All these papers represent the preferences as a binary relation over acts a la Savage.

Instead, we abstract away from the state space formulation, and we represent the preferences of the DM in the space of lotteries. In doing so, we face a complication: when the correlation between alternatives matters, the use of a binary relation over lotteries is not sufficiently rich as a modeling 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></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></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. However, 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 $(0, 10)$). These considerations suggest that the classical approach of describing the DM’s tastes using a binary relation over lotteries is not viable since the DM cannot rank $p_1$ and $p_2$ without additional information about their joint distribution. Indeed, applied researchers (e.g., Filiz-Ozbay, and Ozbay, 2007, Braun, and
Muermann, 2004, Smith 1996) have shown that there are a variety of economically important situations such as auctions, insurance and health interventions where the correlation between lotteries impacts choices.

In this paper, we instead apply the preference set concept, introduced by Fishburn (1991). 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 $X \times X$. As we can always represent such a joint distribution in the tabular form used above, we will refer to the first and second marginal respectively as the row and column ones. Given this joint lottery, the DM decides if, given the marginals and the correlation structure, he prefers to be paid accordingly to the realized row or column outcome. Then, we say that $p$ belongs to the preference set $P$ if and only if the DM prefers to be paid accordingly to the row outcome. In our previous example, we have $p \in P$, and $p' \notin P$.

Several reasons motivate our modeling choice. On a theoretical side, as it will be clear momentarily, this approach abstracts away from the ancillary state-space assumptions to obtain a clear comparison with Expected Utility. This abstraction is by no means a technical point. If we want to test the theory, 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 aversion literature, this is a real concern.

The second motivation comes from our Salience Theory application. The experimental evidence and the examples used to justify the 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 joint realizations of the two lotteries under scrutiny. Therefore, axioms stated in terms of joint lotteries are more natural to map into the BGS model, and they can be directly challenged by the existing experimental evidence on the model. Moreover, under the alternative state-space formulation, the characterization of the salience properties postulated by BGS is much more demanding in terms of the underlying state space’s structural properties.\textsuperscript{3}

\textsuperscript{1}Fishburn (1991) introduces the concept of preference sets for intransitive preferences over multiattribute products, and it applies to choices between acts in Fishburn (1990). To the best of our knowledge, this is the first work in which preference sets are used to axiomatize preferences under risk.


\textsuperscript{3}Among other things, the state space has to be atomless, a property at odds with the small
We first identify the three axioms on the preference set $P$ necessary and sufficient to obtain a weaker representation than the Expected Utility one, which allows for Regret and Salience Theory as particular cases. These axioms are Completeness, Strong Independence, and Archimedean Continuity, and they are one to one with the representation

$$ p \in P \Leftrightarrow \sum_{x,y} \phi(x, y) p(x, y) \geq 0 \quad (1) $$

where $\phi$ is a skew-symmetric functional. To justify the names given to the axioms, we show that these properties are mild relaxation of their more familiar counterparts for binary relations and that if Transitivity is assumed on top of them, the representation reduces to the EU one.

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. A payoff of our preference sets setup is that it allows us to state and characterize these properties in an extremely straightforward manner. We find that Ordering is the property that brings Salience Theory outside the EU or Prospect Theory realm. Instead, Diminishing Sensitivity and Reflexivity combined amount to the usual risk-aversion in gains, risk-loving in loss property featured by Prospect Theory. We then explain how to identify the parameters of BGS from the observed pairwise comparisons of a DM.

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 have an alternative that is weakly preferred to all the others when facing a set of at least three options. Suppose now that the agent can randomize over his feasible alternatives with a randomization device independent of the underlying lotteries. In this setting, we prove the existence of optimal randomization using the min-max theorem.

---

4 Of course, Salience Theory also makes some predictions about non-choice behavior that separate it from preexisting models, such as the attention dedicated to each dimension of the alternatives. Thus, the use of additional instruments such as eye-tracking to further investigate is critical.
**Related Literature**  This paper belongs to the literature studying the axiomatization of correlation sensitive models of choice. This literature starts with the classical works of Fishburn (1989), Sugden (1993), and Quiggin (1994). Recently, Diecidue and Somasundaram (2017) significantly improve the Regret model’s previous representation, providing an axiomatization that delivers a continuous regret function on an arbitrary finite state space. Their main conceptual contribution to the pre-existing work is to single out the axioms for the more restrictive version of Regret Theory originally formulated by Loomes and Sugden (1982) and separate the utility and the regret function. In this sense, their work is complementary to ours. In the first part of the paper, we want to axiomatize the more general form of correlation sensitive preferences to characterize not only Regret Theory but also Salience Theory as particular cases of this model.

Fishburn (1990) uses preference sets to provide a different axiomatization of the Skew-symmetric Additive model. On a technical side, the object on which the preferences are defined is different: Fishburn defines the preference sets as subsets of the space of two outcomes acts, whereas we focus on joint distribution over outcomes. This difference naturally leads to a novel set of axioms. Notice that by letting the preference sets being a subset of the multivariate acts, Fishburn (1990) features the general disadvantages discussed above: potential confusion with ambiguity aversion, axioms that are sufficient for the representation but not necessary, more difficult comparison with EU, and a more relevant departure from the version of the model that has been experimentally tested. These disadvantages become even more relevant in our specific Salience Theory application: first, under the act formulation, the additional properties characterizing Salience Theory as a particular case become much more involved. Second, generalizations that build on our axiomatization to combine Salience Theory for consumption and risk (see Köster, 2020) cannot be conciliated with the “structure axiom” needed in Fishburn (1990), therefore limiting the scope of his axiomatization.

To the best of our knowledge, this paper is the first attempt to axiomatize the Salience Theory of choice under risk. Ellis and Masatioglu (2020) proposed an axiomatization of the Salience Theory of consumption (see Bordalo, Gennaioli, and Shleifer, 2013). Apart from the different analysis subjects, our paper also differs from theirs on 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 state
Herweg and Muller (2019) provide a comparison between the Salience and Regret Model, arguing 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 preference sets. Then, in Section 3 we describe the weakening of EU that is necessary to capture sensitivity to correlation, while in Section 4 we provide the additional axioms characterizing Salience Theory. Section 5 extends the model to choice from nonbinary subsets. All the proofs of the results in the main text are in Appendix A. Finally, Appendix B studies the Rank-Dependent version of Salience Theory.

2 Preference sets

Let $X$ be an arbitrary nonempty 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. To better understand the concept of preference set, recall that a binary relation 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 $p_2$. In that case, the subset $R$ can also be described by a binary relation. However, this approach is implicitly assuming that the correlation between the two lotteries is irrelevant.

In this paper, we follow the other natural extension of the deterministic case: we model the preferences of the DM by a subset $P$ (called preference set) of the space of finite support 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 row or the column outcome. Then, we say that $p \in P$ if and only if he (weakly) prefers to be paid accordingly to the row outcome. The fact that the knowledge of the marginal $p_1$ and $p_2$ may be insufficient to determine whether $p \in P$ is the deviation from the standard paradigm of rational choice.
2.1 Experimental Implementation

Since preference sets are not the standard objects used to describe the falsifiable implications of a DM tastes, a roadmap to how to test the axioms involving them in an experiment can be helpful. The subject faces a finite-support joint distribution $p$ over prizes that can be summarized by a table:

$$
\begin{array}{cccc}
    p & y_1 & \cdots & y_j & \cdots & y_m \\
  x_1 & p_{11} & \cdots & p_{1j} & \cdots & p_{1m} \\
  \vdots & \vdots & \ddots & \vdots & \ddots & \vdots \\
  x_i & p_{1i} & \cdots & p_{ij} & \cdots & p_{im} \\
  \vdots & \vdots & \ddots & \vdots & \ddots & \vdots \\
  x_n & p_{n1} & \cdots & p_{nj} & \cdots & p_{nm} \\
\end{array}
$$

That is, the experimenter tells the subject that every pair of outcomes $(x_i, y_j)$ realizes with probability $p_{ij}$. Then, given the correlation structure between the two alternatives, the subject chooses between being paid according to the row prizes (the $x$’s) or the column prizes (the $y$’s). If he chooses to be paid according to the rows (the columns), if outcome $(x_i, y_j)$ realizes he gets $x_i$ ($y_j$) regardless of the value of $y_j$ ($x_i$).\(^5\)

A joint distribution $p$ belongs to the preference set $P$ if, when faced with it, the DM chooses to be paid according to the prizes on the rows. Therefore, we can always observe if a joint distribution is in the preference set or not. The typical axioms we impose on preference sets have the form “if $p \in P$ then $p'$ belongs to $P$,” where $p'$ has some particular relation with $p$. Therefore, such an axiom can be falsified by a DM that chooses to be paid according to the row prizes under $p$ and according to the column prizes under $p'$.

2.2 Preference Sets and Binary Relations

In what follows, we briefly discuss the relation between preference sets and the classical notion of binary relation. For every joint distribution $p$, we denote as $p_1 \in \Delta (X)$ and $p_2 \in \Delta (X)$, respectively, the row and column marginals of $p$. Formally:

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

\(^5\)Our theory is silent about the information revealed to the subject after a joint outcome $(x, y)$ is drawn. One may expect that the behavior may differ, whether only the component that will be paid out to the DM or the joint realization is revealed.
First, notice that a binary relation \( \succeq \) over marginal distributions induces a *unique* preference set \( P_\succeq \) that evaluates joint distribution in coherence with the ranking over marginal distributions induced by \( \succeq \).

**Definition 1** The preference set \( P_\succeq \) induced by a binary relation \( \succeq \) is defined as

\[
p \in P_\succeq \iff p_1 \succeq p_2.
\]

It is easy to see that if \( \succeq \) and \( \succeq' \) are two different binary relations, they induce different preference sets \( P_\succeq \) and \( P_{\succeq'} \). Therefore, it is clear that preference sets allow for more detailed descriptions of the DM tastes than binary relations. Second, a preference set \( P \) induces a (possibly incomplete) binary relation on \( \Delta (X) \).

**Definition 2** The binary relation \( \succeq^P \) induced by a preference set \( P \) is defined 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 \( p \succeq^P p_2 \) requires that 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). Of course, when \( \succeq^P \) is complete, \( \succeq^P \) is sufficient to characterize the DM’s tastes. Unsurprisingly, \( \succeq^P \) may fail to be complete. In this case, the patterns of behavior that can be described using preference sets are much richer than those for binary relations. A weaker concept would have replaced “for all \( q \)” with “for some \( q \).” Proposition 1 witnesses that our way of defining this notion is more fruitful. For the moment, we limit ourselves to observing that the maps between preference sets and binary relations are the inverse of each other.

**Lemma 1** For every binary relation \( \succeq \), the binary relation \( \succeq^P \) coincides with \( \succeq \).

### 2.2.1 Axioms for Binary Relations

Here we collect the definition of some standard axioms for binary relations over marginal distributions in \( \Delta (X) \) that are later referred to in the text. A reader that is mostly interested in the Salience model can skip directly to Section 3.

**Axiom 1 (Completeness)** For all \( p_1, p_2 \in \Delta (X) \), either \( p_1 \succeq p_2 \) or \( p_2 \succeq p_1 \) or both.
Completeness is the standard axiom that requires that the DM can (weakly) rank all the marginal lotteries. Our analysis highlights a new reason why Completeness may fail: comparing some pairs of lotteries may not be performed without knowledge of their correlation structure.

**Axiom 2 (Transitivity)** For all \( p_1, p_2, p_3 \in \Delta(X) \), if \( p_1 \succsim p_2 \) and \( p_2 \succsim p_3 \), then \( p_1 \succsim p_3 \). Moreover, if either \( p_1 \succ p_2 \) or \( p_2 \succ p_3 \), then \( p_1 \succ p_3 \).

Transitivity is the other central tenet of rationality. We next show how this is the key assumption that needs to be dropped to allow for correlation sensitive preferences.

**Axiom 3 (Strong Independence)** For all \( p_1, p_2, p_3 \in \Delta(X) \) and \( \alpha \in (0, 1) \),

\[
p_1 \succsim p_2 \Leftrightarrow \alpha p_1 + (1 - \alpha) p_3 \succsim \alpha p_2 + (1 - \alpha) p_3.
\]

Strong Independence is the axiom that is usually paired to Completeness, Transitivity, and Archimedean Continuity to derive the Expected Utility representation. Since we often work without Transitivity in this paper, we also need to consider an alternative and stronger form of Independence.

**Axiom 4 (Strong B-Independence)** For all \( p_1, p_2, p_3, p_4 \in \Delta(X) \) and \( \alpha \in (0, 1) \),

\[
p_1 \succsim p_2, p_3 \succsim p_4 \Rightarrow \alpha p_1 + (1 - \alpha) p_3 \succsim \alpha p_2 + (1 - \alpha) p_4.
\]

Moreover, if \( p_1 \succ p_2 \), then \( \alpha p_1 + (1 - \alpha) p_3 \succ \alpha p_2 + (1 - \alpha) p_4 \).

Strong B-Independence implies Strong Independence, and the two axioms coincide under Transitivity.

**Lemma 2** If \( \succsim \) satisfies Transitivity and Strong Independence, then \( \succsim \) satisfies Strong B-Independence.

The next axiom is a standard and weak form of continuity imposed on preferences defined over a convex set.

**Axiom 5 (Archimedean Continuity)** For all \( p_1, p_2, p_3 \in \Delta(X) \) such that \( p_1 \succ p_2 \) and \( p_2 \succ p_3 \), there exist \( \alpha, \beta \in (0, 1) \) such that

\[
\alpha p_1 + (1 - \alpha) p_3 \succ p_2 \text{ and } p_2 \succ \beta p_1 + (1 - \beta) p_3.
\]
When dealing with nontransitive preferences, a slightly more demanding version of Archimedean Continuity is needed.

**Axiom 6 (Archimedean B-Continuity)** For all \( p_1, p_2, p_3, p_4 \in \Delta(X) \) such that \( p_1 \succ p_2 \) and \( p_3 \nsubseteq p_4 \), there exist \( \alpha, \beta \in (0,1) \) such that

\[
\alpha p_1 + (1 - \alpha) p_3 \succ \alpha p_2 + (1 - \alpha) p_4 \quad \text{and} \quad \beta p_1 + (1 - \beta) p_3 \nsubseteq \beta p_2 + (1 - \beta) p_4.
\]

Under Completeness, Archimedean B-Continuity is implied by the more standard Sequential Continuity requirement.

**Axiom 7 (Sequential Continuity)** For each pair of sequences \((p_n)_{n \in \mathbb{N}}\) and \((q_n)_{n \in \mathbb{N}}\) in \( \Delta(X) \) such that \((p_n)_{n \in \mathbb{N}} \to p_0\) and \((q_n)_{n \in \mathbb{N}} \to q_0\)

\[
p_n \succeq q_n \quad \text{for all} \quad n \in \mathbb{N} \implies p_0 \succeq q_0.
\]

Archimedean B-Continuity is implied by Sequential Continuity under Completeness.

**Lemma 3** If \( \succeq \) satisfies Sequential Continuity and Completeness, then \( \succeq \) satisfies Archimedean B-Continuity.

For future reference, we state some additional axioms that we use when the prizes set is a subset of the real line. We denote as \( \delta_{(x,y)} \) the joint lottery such that with probability one, the row outcome is \( x \), and the column outcome is \( y \).

**Axiom 8 (Monotone Contribution)** Let \( X \subseteq \mathbb{R} \). For all \( x, y \in X \), \( \alpha \in (0,1) \), \( p_1, p_2 \in \Delta(X) \) if \( x \succ y \) then

\[
\alpha \delta_y + (1 - \alpha) p_1 \succeq p_2 \Rightarrow \alpha \delta_x + (1 - \alpha) p_1 \succ p_2.
\]

Monotone contribution is a generalization of the usual strict monotonicity axiom to preferences that are not necessarily transitive.

**Axiom 9 (MPS Risk Aversion)** Let \( X \subseteq \mathbb{R} \). For all \( p_1, p_2 \) such that \( p_1 \) is a Mean Preserving Spread of \( p_2 \), \( p_2 \succeq p_1 \).
3 General Representation Theorem

To better understand representation (1), it is useful to compare with Expected Utility. Let $p \in \Delta (X \times X)$. Under EU, there exists a utility function $u$ such that

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

(2)

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

(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: (i) 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 $p_2$. (ii) Follow the same procedure for marginal $p_2$. (iii) Compare the scores obtained for the two alternatives. Choose to be paid accordingly to the row outcome if and only if $U (p_1) \geq U (p_2)$. It is immediate to see that there is no role for correlation between the two marginal distributions under this procedure. This procedure consists of a \textit{Comparison of (Probabilistic) Aggregations}, and in the case of EU is given by (2).

Alternatively, one may consider the following procedure: (i) 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. (ii) Do the same for every joint realization. (iii) Aggregate all these comparisons according to the probability measure $p$ obtaining $\Phi (p) = \sum_{x,y} p (x,y) \phi (x,y)$ (iv) Choose to be paid accordingly to the row outcome if and only if $\Phi (p) \geq 0$.

Since what matters is the lotteries’ joint value, there is room for correlation with this procedure. This procedure consists of a (Probabilistic) \textit{Aggregation of Comparisons}, and it is the kind of reasoning that characterizes both Regret and salience sensitive DMs, and for EU it corresponds to line (3). The pioneering works by Bell (1982), and Loomes and Sugden (1982) already recognize the descriptive and normative value of such procedure.
However, the result by vN-M tells us that, under Expected Utility, the two algorithms lead to the same conclusion. In the second procedure, the DM uses $\phi(x, y) = u(x) - u(y)$, 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 $\phi$.

Before going further, a piece of notation is needed. Given $p \in \Delta(X \times X)$ we define its conjugate distribution $\bar{p}$ as

$$\forall (x, y) \in X \times X \quad \bar{p}(x, y) = p(y, x).$$

Therefore, the conjugate distribution is just a relabeling of the row and column outcomes into each other.

**Axiom 10 (Completeness)** For all $p \in \Delta(X \times X)$

$$p \notin P \Rightarrow \bar{p} \in P.$$

Completeness is a very minimal requirement about the rationality of the DM. If he prefers to be paid accordingly to the column marginal when the joint distribution is $p$, he (weakly) prefers to be paid accordingly to the row marginal after a relabeling of rows outcomes into column ones and vice-versa. The next lemma clarifies that Completenessness of the preference set weakens the standard Completeness assumption for binary relations. The example discussed in the introduction shows why it may be a strictly weaker requirement.

**Lemma 4** The following hold true:

1. Let $\succeq$ be a binary relation. If $\succeq$ satisfies Completeness, then $P_\succeq$ satisfies Completeness.

2. Let $P$ be a preference set. If $P_\succeq^P$ satisfies Completeness, then $P$ satisfies Completeness.

That is, the preference set derived from a complete binary relation satisfies Completeness. Moreover, the binary relation induced by a preference set is complete only if the preference set satisfies Completeness.

Given a preference set $P \subseteq \Delta(X \times X)$, the strict preference set is defined as

$$\hat{P} = \{ p \in P : \bar{p} \notin P \}.$$
In words, a joint distribution $p$ is in the strict preference set if the DM weakly prefers to be paid accordingly to the row outcome (i.e., $p \in P$), and he does not prefer to be paid accordingly to the column outcome (i.e., $\bar{p} \notin P$). It is the natural counterpart in the world of preference sets of the asymmetric part of a binary relation. 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 row outcome, it then seems reasonable he prefers to be paid accordingly to the row outcome 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 initial preferences is.

**Axiom 11 (Strong Independence)** For all $p, q \in P$, and all $\lambda \in (0, 1)$

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

Moreover, if $q \in \hat{P}$, then

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

Next lemma makes clear that also this condition is a weakening of the usual axiom for binary relations.

**Lemma 5** If a binary relation $\succeq$ satisfies Strong B-Independence, then $P_\succeq$ satisfies Strong Independence.

The property of the binary relation that implies the Strong Independence of the corresponding preference set is a minor strengthening of the usual Independence axiom. This strengthening is needed because, in general, we are not imposing Transitivity. For transitive binary relations, Strong B-Independence and Strong Independence are equivalent.\(^6\)

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

**Axiom 12 (Archimedean Continuity)** For all $p \in \hat{P}$, $q \notin P$, there exist $\alpha, \beta \in (0, 1)$ such that

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

\(^6\)Indeed, Lemma 8 in the Appendix is a sharper equivalence result between Strong Independence for binary relations and Strong Independence for binary relations under Transitivity.
We can also relate this axiom with the standard Archimedean Axiom for binary relations.

**Lemma 6** Let $\succeq$ be a binary relation. If $\succeq$ satisfies Archimedean B-Continuity, then $P_{\succeq}$ satisfies Archimedean Continuity.

The next theorem provides a representation of the preference sets satisfying these three axioms.\(^7\) Recall that a functional $\phi : X \times X \to \mathbb{R}$ is said to be skew-symmetric if for all $x, y \in \mathbb{R}$, $\phi(x, y) = -\phi(y, x)$.\(^8\)

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

The theorem’s importance stems from the fact that it connects a subset of the most basic EU axioms to a fairly general representation that can consider the role of correlation between the alternatives. Precisely, the names of the axioms used to characterize (4) suggest that, to the celebrated vN-M set of axioms, we are just relaxing Transitivity. Proposition 1 shows that this interpretation is correct. To do so, we need to translate Transitivity in the language of preference sets.

**Axiom 13 (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., $r \in P$. Next result verifies the asserted link between the standard notion of Transitivity and the one we propose.

**Lemma 7** The following hold true:

1. If $\succsim$ satisfies Transitivity, $P\succsim$ satisfies Transitivity.
2. If $P$ satisfies Transitivity, then $\succsim^P$ satisfies Transitivity.
3. $\succsim^P$ satisfies Transitivity and Completeness if and only if $P$ satisfies Transitivity and Completeness.

Interestingly, the following result proves that when Transitivity is imposed on top of the previous axioms, the decision criterion reduces to Expected Utility maximization. This equivalence further justifies the terminology used for Completeness, Strong Independence, and Archimedean Continuity.

**Proposition 1** If $P$ satisfies Completeness, Strong Independence, and Archimedean Continuity, the following are equivalent:

1. $P$ satisfies Transitivity;
2. $\succsim^P$ satisfies Completeness;
3. $\succsim^P$ satisfies Completeness, Transitivity, Archimedean Continuity, and Strong Independence;
4. $\succsim^P$ admits an Expected Utility representation.

The previous proposition can be used to highlight an additional payoff of our Preference Sets approach: we obtain a new set of axioms that are one to one with the Expected Utility Model.

**Theorem 2** The following are equivalent:

1. $\succsim$ satisfy Completeness, Strong B-Independence, and Archimedean B-Continuity;
2. $\succsim$ satisfies Completeness, Transitivity, Archimedean Continuity, and Strong Independence;
3. $\succsim$ admits an Expected Utility representation.
3.1 Monotonicity and Continuity

Since Salience Theory is defined for lotteries with monetary outcomes, from now on, we are going to focus on the case where \( X = [l, h] \subseteq \mathbb{R} \) is a nondegenerate interval endowed with the usual topology. In this richer setting, we discuss how to use Preference Sets to axiomatically describe standard regularity conditions for the representing function, such as Monotonicity and Continuity.

**Axiom 14 (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

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

Since we do not impose Transitivity, our Monotonicity axiom slightly departs from the usual one: it requires that whenever \( x \) is strictly larger than \( y \), \( x \) is more favorably compared than \( y \) to every alternative \( z \).

**Remark 1** A binary relation \( \succeq \) satisfies Monotone Comparisons if and only if \( P_\succeq \) satisfies Monotonicity.

Given the representation in (4), Monotonicity is easily characterized in terms of \( \phi \).

**Remark 2** 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.

Before proceeding with the Salience Theory, a few observations about the connection between FOSD and the Monotonicity axiom in the general model are in order. It is worth noting that the Monotonicity axiom is not enough to guarantee that the preference set \( P \) satisfies First Order Stochastic Dominance, where the latter is defined in our setting as the requirement that

\[
\forall p \in \Delta (X \times X) : p_1 \succeq_{FOSD} p_2, p \in P \tag{5}
\]

with \( p \in \hat{P} \) if \( p_1 \neq p_2 \). However, we can make quite a few predictions about the stochastic monotonicity implications of the decision criterion we are considering. Indeed, the preference set \( P \) satisfies (5) when \( p \) is an independent joint distribution, i.e., \( p = p_1 \times p_2 \).

Finally, this setup also allows for a simple characterization of the continuity properties of \( \phi \).

---

\(^9\)Indeed, the previous proposition guarantees that a preference set that satisfies Completeness, Strong Independence, Archimedean Continuity, and Monotonicity, admits a representing \( \phi \) that satisfies the OPT and I properties of Loomes and Sugden (1987).
Axiom 15 (Continuity in Outcomes) Let \((x_n)_{n \in \mathbb{N}} \to x\). Then, for every \(\alpha \in [0, 1]\), \(y \in X\) and \(p \in \Delta (X \times X)\)

\[
\forall n \in \mathbb{N} \quad \alpha \delta_{(x_n,y)} + (1 - \alpha) p \in P \implies \alpha \delta_{(x,y)} + (1 - \alpha) p \in P
\]

and

\[
\forall n \in \mathbb{N} \quad \alpha \delta_{(y,x_n)} + (1 - \alpha) p \in P \implies \alpha \delta_{(y,x)} + (1 - \alpha) p \in P.
\]

Given Transitivity, Completeness, and Archimedean Continuity, Continuity in Outcomes is one to one with a continuous \(\phi\).

Remark 3 If \(P\) admits a representation as in (4), \(P\) satisfies Continuity in Outcomes if and only if \(\phi\) is continuous in both arguments.

4 Salience Characterization

In this section, we first describe Salience Theory as introduced by BGS and their critical properties of Ordering, Diminishing Sensitivity, and Reflexivity, and we argue why it is a particular case of our general representation. We then propose a behaviorally testable version of the properties, and we show that they coincide with the original ones.

4.1 The BGS Decision Criterion

Salience Theory, as formulated in BGS, explains the behavior of a DM that is facing a joint lottery \(p \in \Delta (X \times X)\). Salience’s main departure from EU Theory 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 function \(\sigma : \mathbb{R}^2 \to \mathbb{R}\) 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. Weak Reflexivity if for all \(x, y, x', y' \in \mathbb{R}_+\) with \(|x - y| = |x' - y'|\),

\[
\sigma (x, y) \geq \sigma (x', y') \iff \sigma (-x, -y) \geq \sigma (-x', -y').
\]
A salience function is a continuous function $\sigma : \mathbb{R}^2 \to \mathbb{R}$ satisfying Symmetry, Ordering, Diminishing Sensitivity, and Weak 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.\(^\text{10}\)

**Definition 4** A preference set $P$ admits a $\sigma$-distorted representation if there exists a continuous function $\sigma : \mathbb{R}^2 \to \mathbb{R}$ that satisfies Symmetry such that

$$p \in P \iff \sum_{(x,y) \in X \times X} (x - y) \sigma(x,y) p(x,y) \geq 0.\quad (6)$$

It admits a (smooth) salience representation if $\sigma$ is a salience function.

We next show that the $\sigma$-distorted representation corresponds amounts to our general representation with the addition of the following axiom.

**Axiom 16 (Continuity at Identity)** Let $x \in X$, and let $(x_n)_{n \in \mathbb{N}}$ be such that $x_n \downarrow x$. Then, for all $k \in \mathbb{R}$ and $\varepsilon \in \mathbb{R}_{++}$ there exists an $m \in \mathbb{N}$ such that for all $n \geq m$

$$(x, x_n), (1 - (x_n - x)); (k, k + \varepsilon), (x_n - x)) \in P.$$  

Then, the characterization of the $\sigma$-distorted representation reads as follow:

**Proposition 2** A preference set $P$ admits a $\sigma$-distorted representation if and only if $P$ satisfies Completeness, Strong Independence, Archimedean Continuity, and Monotonicity, Continuity in Outcomes, and Continuity at Identity.

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. The rest of this section is devoted to the characterization of Ordering, Diminishing Sensitivity, and Weak Reflexivity in terms of testable axioms.

\(^{10}\)In their paper, BGS mainly used the Rank Dependent version of their model. However, they recognized that the Rank Dependent version of the model is subject to some issues due to discontinuity, and they suggest the using 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.” Therefore, in what follows, we stick with the Smooth version. Appendix B analyzes in detail the Rank Dependent version.
4.2 The Ordering Axiom

The idea behind the Ordering property proposed by BGS is straightforward. Fix the outcomes \( x > y \). Then, we can take some \( \alpha, \beta \in (0, 1) \), \( \beta > \alpha \) and consider the two outcomes obtained by mixing \( x \) and \( y \)

\[
x > \beta x + (1 - \beta) y > \alpha x + (1 - \alpha) y > y.
\]

If we consider the two realizations \((x, y)\) and \((\alpha x + (1 - \alpha) y, \beta x + (1 - \beta) y)\) the first pair of outcomes has more widespread values and therefore Ordering implies that its probability will be relatively overweighted. However, distortions of probabilities are not observable, and therefore, we cannot directly test the BGS form of Ordering. Nevertheless, we can propose a behavioral (i.e., testable) version of the property.

Now, if we look at the joint distribution

\[
\left( (x, y), \frac{\beta - \alpha}{1 + \beta - \alpha}; (\alpha x + (1 - \alpha) y, \beta x + (1 - \beta) y), \frac{1}{1 + \beta - \alpha} \right)
\]

the row and column marginals have the same expected value, and they should be indifferent to an expected value maximizer. 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 \((x, y)\), and favors the row component, a salience sensitive DM prefers (at least weakly) to be paid accordingly to the row component. The previous reasoning is crystallized in the Ordering Axiom.

**Axiom 17 (Ordering)** For every \( x, y \in \mathbb{R}, \alpha, \beta \in [0, 1] \) if \( x > y, \beta \geq \alpha, \) and at least one between \( \beta \) and \( \alpha \) is in \((0, 1)\), we have that

\[
\left( (x, y), \frac{\beta - \alpha}{1 + \beta - \alpha}; (\alpha x + (1 - \alpha) y, \beta x + (1 - \beta) y), \frac{1}{1 + \beta - \alpha} \right) \in \hat{P}.
\]

The next proposition shows that the axiom corresponds to the original property of BGS.

**Proposition 3** Let \( P \) admit a \( \sigma \)-distorted representation. Then \( P \) satisfies Ordering if and only if \( \sigma \) satisfies Ordering.
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 highest 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 probability $\frac{1}{2}$ both to $(x, y)$ and $(y + k, x + k)$, with $x > y \geq 0$. 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 row marginal, the DM is going to choose to be paid according to the row outcome. The previous reasoning is crystallized in the Diminishing Sensitivity axiom.

**Axiom 18 (Diminishing Sensitivity)** For every $x > y \geq 0$, and $k \in \mathbb{R}_+$

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

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

The next proposition shows that our behavioral definition of Diminishing Sensitivity corresponds to the original property of BGS.

**Proposition 4** If $P$ admits a $\sigma$-distorted representation, it satisfies Strict Diminishing Sensitivity if and only if $\sigma$ satisfies Diminishing Sensitivity.

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 (cf. also Proposition 7) to decision criteria that are not necessarily transitive.

**Proposition 5** 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.

**Remark 4** In our more general setup, Risk Aversion for gains always implies Diminishing Sensitivity. However, the following example shows that Risk Aversion for gains is a strictly more demanding property. Let the salience function be equal to the leading example in BGS, that is

$$\sigma (x, y) = \frac{x - y}{x + y + 1}.$$
Then $\sigma$ satisfies Ordering and Diminishing Sensitivity, and by Proposition 5 $P$ satisfies Diminishing Sensitivity. The joint distribution $p$ given in the following table is such that the row marginal is a MPS of the row marginal. Therefore risk aversion in the gain domain would prescribe that $p \notin P$:

<table>
<thead>
<tr>
<th>$p$</th>
<th>0</th>
<th>1</th>
<th>2</th>
</tr>
</thead>
<tbody>
<tr>
<td>0</td>
<td>0</td>
<td>1/4</td>
<td>0</td>
</tr>
<tr>
<td>1</td>
<td>0</td>
<td>1/2</td>
<td>0</td>
</tr>
<tr>
<td>2</td>
<td>1/8</td>
<td>0</td>
<td>1/8</td>
</tr>
</tbody>
</table>

However, it is immediate to see that for a DM with salience function given by (7), $p \in P$, the reason being the high salience of the realization $(2,0)$. Therefore, the preference set of such a DM satisfies Diminishing Sensitivity but not Risk Aversion for gains.

### 4.4 The Weak Reflexivity Axiom

The last property introduced by BGS is Weak Reflexivity, and it captures the symmetry around 0 of the distortions. Again, we provide a behaviorally testable counterpart of their axiom.

**Axiom 19 (Weak Reflexivity)** For every $x, y, w, z \in \mathbb{R}_+$, with $x - y = z - w$

\[
\left( (x, y), \frac{1}{2}; (w, z), \frac{1}{2} \right) \in \hat{P} \Leftrightarrow \left( (-y, -x), \frac{1}{2}; (-z, -w), \frac{1}{2} \right) \in \hat{P}.
\]

The axiom is easily seen to be one to one with the corresponding property of the distortion function $\sigma$.

**Proposition 6** If $P$ admits a $\sigma$-distorted representation, $P$ satisfies Weak Reflexivity if and only if $\sigma$ satisfies Weak Reflexivity.

So far, we have not attached any specific interpretation to the lotteries’ realizations, 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 Weak Reflexivity axiom, with the implied role for outcome 0, better suits the latter interpretation. We notice that Weak Reflexivity implies a form of the Preference Reversal of risk attitudes featured by Prospect Theory.
Proposition 7 Suppose that $P$ has an EU representation and satisfies Monotonicity and Weak Reflexivity. Then $P$ is risk-averse (resp. risk-loving) for lotteries with values in $[a,b] \subseteq \mathbb{R}_+$ if and only if $\succeq$ is risk loving (resp. risk-averse) for lotteries with values in $[-b,-a]$.

This result sheds light on the observation made in BGS that Salience Theory can explain the experimental evidence in favor of the fourfold pattern (see, e.g., Bruhin, Fehr-Duda, and Epper 2010 and Kahneman 2011). Diminishing Sensitivity only induces the differences between the gain and loss domains, whereas it is its combination with Ordering and Weak Reflexivity that induces the differences within the two domains.

4.5 Complete Characterization of the Salience Model

We will now put the pieces of the analysis togethere, and we will provide a complete characterization of the Salience model. To do so, we will need a last technical axiom.

Axiom 20 (Continuity at Identity) Let $x \in X$. Then, for every $(x_n)_{n \in \mathbb{N}}$ be such that $x_n \downarrow x$, and for every $k \in \mathbb{R}$ and $\varepsilon \in \mathbb{R}_{++}$ there exists an $m \in \mathbb{N}$ such that for all $n \geq m$

$($(x,x_n),(1-(x_n-x)),(k+\varepsilon,k),(x_n-x)) \in P$.

Moreover, for every $(x_n)_{n \in \mathbb{N}}$ be such that $x_n \uparrow x$, and for every $k \in \mathbb{R}$ and $\varepsilon \in \mathbb{R}_{++}$ there exists an $m \in \mathbb{N}$ such that for all $n \geq m$

$($(x_n,x),(1-(x_n-x)),(k+\varepsilon,k),(x_n-x)) \in P$.

With this, we have a complete characterization of the Salience model.

Theorem 3 A preference set $P$ admits a (smooth) salience representation if and only if $P$ satisfies Completeness, Strong Independence, Archimedean Continuity, Monotonicity, Continuity in Outcomes, Continuity at Identity, Ordering, Diminishing Sensitivity and Weak Reflexivity.

Relation with Regret Theory This behavioral characterization allows us to immediately compare Salience Theory with Regret Theory. Indeed, recall that Loomes and Sugden (1987) say that a preference set $P$ admits a Regret Theory representation if $P$ admits a representation as in (4), it satisfies Monotonicity, and the representing $\phi$ satisfies Convexity:

$\phi(x,y) > \phi(x,z) + \phi(z,y)$ for all $x > z > y$. 

22
Corollary 4 If a preference set $P$ satisfies Completeness, Strong Independence, Archimedean Continuity, and Ordering, the representing $\phi$ satisfies Convexity. Therefore, Salience Theory is a particular case of Regret Theory.

4.6 Identification of the Salience Function

Another advantage of our use of preference sets is that in light of Theorem 1, we can directly test Salience Theory by first constructing a candidate salience function $\sigma$ and then checking whether it satisfies the properties imposed by BGS.\(^{11}\) As a preliminary observation, it is immediate to see from (6) that if the preferences set $P$ admits a smooth salience representation with salience function $\sigma$, they also admit a smooth salience representation with salience function $\lambda \sigma$ whenever $\lambda \in \mathbb{R}_+$. Therefore, to eliminate this degree of freedom, we set $\sigma(1,0) = 1$.

Now, for every $(x,y) \in \mathbb{R}^2$ with $y > x$, if the preference set $P$ admits a smooth salience representation, by Proposition 2 $P$ satisfies Completeness, Archimedean Continuity, and Monotonicity, and therefore by Theorem 1 there exists an $\alpha_{x,y} \in (0,1)$ such that

$$\alpha_{x,y} \delta_{(x,y)} + (1 - \alpha_{x,y}) \delta_{(1,0)} \in P \setminus \hat{P}.$$ 

Therefore, we can define

$$\sigma(x,y) = \frac{(1 - \alpha_{x,y})}{\alpha_{x,y}(y-x)}.$$

It is immediate to check that this is the only possible value for $\sigma$. We can use this procedure and the fact that by symmetry $\sigma(x,y) = \sigma(y,x)$ for those $(x,y) \in \mathbb{R}^2$ with $x > y$ to construct the candidate salience function. At this point, checking Salience Theory only boils down to verifying that $\sigma$ satisfies the Ordering, Diminishing Sensitivity, and Weak Reflexivity properties postulated by BGS.\(^{12}\)

\(^{11}\)Since identification is not the focus of this paper, we restrict ourselves to the main case considered in BGS, instead of the more general version

$$\sum_{(x,y) \in X \times X} (u(x) - u(y)) \sigma(x,y) p(x,y).$$

Measurement of $\sigma$ under that more general specification may be performed using the concept of standard sequence of outcomes introduced by Bleichrodt, Cillo, and Diecidue (2010).

\(^{12}\)Clearly, deriving the entire Salience function involves an infinite number of comparisons. An implementable test involves checking the properties on a finite grid of values.
5 Choice from arbitrary sets

We now turn to an important question left open by the previous analysis: how the DM chooses from a set $A$ 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. Fortunately, we can 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), an optimal randomization over the available alternatives exists.

It is easier to state the result if we move from the space of lotteries to acts. Formally, we consider a probability space $(S, \mathcal{F}, \mathbb{P})$, and the alternatives under consideration are random variables that assume a finite number of values in that space. We denote as $B_0(S)$ the space of such random variables. Starting from a preference set $P$ that satisfies Completeness, Strong Independence, and Archimedean Continuity, we can naturally derive a binary preference relation on $B_0(S)$. 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 $p_{f,g} \in \Delta (X \times X)$ 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

$$f \succeq g \iff p_{f,g} \in P \iff \sum_{(x,y)} p_{f,g}(x,y) \phi(x,y) \geq 0. \tag{8}$$

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 $\succeq$ is transitive, and therefore there may be no $f \in A$ such that $f \succeq g$ for all $g \in A$. However, in some situations, the DM may have the possibility to randomize over the alternatives with a randomization device (with finite support) that is independent of the state of the world $s \in S$.\footnote{If the randomization technology is not independent of the realization of the state $s \in S$, the DM may use the realization of the random device to infer something about the payoffs paid by the act, changing the problem of choosing a randomization to an information acquisition problem.} Given that this randomization has no additional correlation, it makes good sense to extend the binary relation $\succeq$ to $\Delta (A)$ in a linear fashion. In particular, given two randomizations over alternatives $\alpha, \beta \in \Delta (A)$:

$$\alpha \succeq \beta \iff \sum_{f \in A} \sum_{g \in A} \alpha(f) \beta(g) \sum_{(x,y)} p_{f,g}(x,y) \phi(x,y) \geq 0.$$
The next result guarantees the existence of a (possibly randomized) optimal alternative.

**Proposition 8** Let $A$ be a finite subset of $B_0(S)$. Then, there exists $\alpha \in \Delta(A)$ such that $\alpha \succeq \beta$ for all $\beta \in \Delta(A)$.

In other words, even if $\succeq$ is an incomplete binary relation over $\Delta(A)$, there exists a $\succeq$-maximum in $\Delta(A)$.

**Example 1 (The effect of salience on random choice)** Suppose that the probability space is composed of three equally likely states: $S = \{s_1, s_2, s_3\}$ and $\mathbb{P}(\{s_1\}) = \mathbb{P}(\{s_2\}) = \mathbb{P}(\{s_3\}) = \frac{1}{3}$. The preference set $P$ admits a salience representation with Salience function $\sigma(x, y) = |x - y|$. First, suppose that the DM faces a choice set $A = \{f, g, h\}$ composed by three symmetric acts:

<table>
<thead>
<tr>
<th></th>
<th>$s_1$</th>
<th>$s_2$</th>
<th>$s_3$</th>
</tr>
</thead>
<tbody>
<tr>
<td>$f$</td>
<td>3</td>
<td>2</td>
<td>1</td>
</tr>
<tr>
<td>$g$</td>
<td>1</td>
<td>3</td>
<td>2</td>
</tr>
<tr>
<td>$h$</td>
<td>2</td>
<td>1</td>
<td>3</td>
</tr>
</tbody>
</table>

Here, choosing deterministically a single act is not $\succeq$-optimal, since for such a salience sensitive DM $f \succeq g$, $g \succeq h$ and $h \succeq f$. However, it is easy to see that the unique $\succeq$-optimal randomization sees the DM randomizing uniformly over the three acts. Next, suppose that the DM faces the choice set $A' = \{f, g, h, l\}$ where

<table>
<thead>
<tr>
<th></th>
<th>$s_1$</th>
<th>$s_2$</th>
<th>$s_3$</th>
</tr>
</thead>
<tbody>
<tr>
<td>$l$</td>
<td>5</td>
<td>0</td>
<td>0</td>
</tr>
</tbody>
</table>

Notice that for this salience sensitive DM $f \succeq l$, $h \succeq l$ and $l \succeq h$, since when $l$ is compared to $g$, state $s_1$ results sufficiently salient to tilt the comparison in favor of $l$. It is easy to see that when faced with the choice set $A'$, a uniform randomization over $\{f, g, h\}$ is no longer optimal for the agent, and that the unique optimal randomization is $(f, \frac{1}{2}; g, 0; h, \frac{1}{6}; l, \frac{1}{3})$.

\[\text{\textsuperscript{14}}\text{A different approach to the choice between multiple alternatives would have been to study the top-cycle of } \succ^p.\]
6 Conclusion

This work provides a simple axiomatic characterization of preferences over risky choices affected by the correlation between the alternatives considered. In particular, the early work on Skew Symmetric Additive representations for preferences over acts suggested that SEU’s essential relaxation that introduces a role for correlation involves Transitivity. We proved that, when the joint distribution is included in the decision environment’s description, this point can be formalized in the much simpler setting of preferences between lotteries. This setting, in turn, allows a cleaner axiomatic comparison of these theories with EU.

We then apply this framework to provide an axiomatization of the salience model by Bordalo, Gennaioli, and Shleifer within the realm of these correlation sensitive preferences. More precisely, we map one-on-one the functional properties of Ordering, Diminishing Sensitivity, and Weak Reflexivity to testable counterparts. This exercise shows how Ordering is the property that cannot be reconciled with the EU or Prospect Theory model, whereas Diminishing Sensitivity paired with Weak Reflexivity corresponds to the usual risk-averse in gain, risk-loving in loss feature of Prospect Theory.

Finally, we provide an answer to the issue of choice between multiple alternatives. Even if there may not exist a maximal alternative, by extending the decision criterion to independent randomizations over (possibly correlated) distributions, we can show the existence of an optimal randomization. This leads naturally to a Stochastic Choice counterpart of the deterministic model.

A Proofs

Proof of Lemma 1 Let \( p_1 \succeq p_2 \). Then, if \( q \in \Delta (X \times X) \) and \( (q_1, q_2) = (p_1, p_2) \), \( q \in P_\succeq \) by definition of \( P_\succeq \). However, since \( q \) was an arbitrary joint lottery with marginals \( p_1 \) and \( p_2 \), by definition of \( \succeq_P \), we have \( p_1 \succeq_P p_2 \).

Let \( p_1 \succeq_P p_2 \). Then, by definition of \( \succeq_P \), \( p_1 \times p_2 \in P_\succeq \). But by definition of \( P_\succeq \), this means that \( p_1 \succeq p_2 \).

Proof of Lemma 2 Let \( p_1, p_2, p_3, p_4 \in \Delta (X) \) and \( \alpha \in (0, 1) \) with \( p_1 \succeq p_2 \) (resp. \( p_1 \succ p_2 \)), \( p_3 \succeq p_4 \). Then by Strong Independence \( \alpha p_1 + (1 - \alpha) p_3 \succeq \alpha p_2 + (1 - \alpha) p_3 \) (resp. \( \alpha p_1 + (1 - \alpha) p_3 \succ \alpha p_2 + (1 - \alpha) p_3 \)) and \( \alpha p_2 + (1 - \alpha) p_3 \succeq \alpha p_2 + (1 - \alpha) p_4 \). Therefore, by Transitivity \( \alpha p_1 + (1 - \alpha) p_3 \succeq \alpha p_2 + (1 - \alpha) p_4 \) (resp. \( \alpha p_1 + (1 - \alpha) p_3 \succ \alpha p_2 + (1 - \alpha) p_4 \)).
Proof of Lemma 3 Let \( p_1, p_2, p_3, p_4 \in \Delta (X) \) be such that \( p_1 \succ p_2 \) and \( p_3 \npreceq p_4 \). We first show that there exists \( \alpha \in (0, 1) \) such that

\[
\alpha p_1 + (1 - \alpha) p_3 \succ \alpha p_2 + (1 - \alpha) p_4.
\]

Define \( r_n = \left(1 - \frac{1}{n}\right) p_1 + \frac{1}{n} p_3 \) and \( q_n = \left(1 - \frac{1}{n}\right) p_2 + \frac{1}{n} p_4 \). If \( r_n \succ q_n \) for some \( n \in \mathbb{N} \), the result follows by setting \( \alpha = 1 - \frac{1}{n} \). Otherwise, by Completeness of \( \succ \), we have \( q_n \succeq r_n \), but by Sequential Continuity this implies that \( \lim_n q_n = p_2 \succeq \lim_n r_n = p_1 \), a contradiction.

The existence of \( \beta \in (0, 1) \) such that

\[
\beta p_1 + (1 - \beta) p_3 \npreceq \beta p_2 + (1 - \beta) p_4
\]

follows from the first part and noticing that under Completeness \( p_3 \npreceq p_4 \implies p_1 \succ p_3 \) and \( \beta p_1 + (1 - \beta) p_3 \npreceq \beta p_2 + (1 - \beta) p_4 \implies \beta p_2 + (1 - \beta) p_4 \succ \beta p_1 + (1 - \beta) p_3 \).

\[\blacklozenge\]

Proof of Lemma 4 (1) Let \( p \in \Delta (X \times X) \). Since \( \succ \) satisfies Completeness, at least one between \( p_1 \succ p_2 \) and \( p_2 \succ p_1 \) holds. By definition of \( P_{\succ} \) this implies that at least one between \( p \in P \) and \( \bar{p} \in P \) holds.

(2) Let \( p \in \Delta (X \times X) \). Since \( \succ^P \) satisfies Completeness at least one between \( p_1 \succeq^P p_2 \) and \( p_2 \succeq^P p_1 \) holds, and this implies that at least one between \( p \in P \) and \( \bar{p} \in P \) holds.

\[\blacklozenge\]

Proof of Lemma 5 (1) Let \( p, q \in P_{\succ} \) (resp. \( q \in \hat{P}_{\succ} \)) and \( \lambda \in (0, 1) \). By definition of \( P_{\succ} \), \( p_1 \succ p_2 \) and \( q_1 \succeq q_2 \) (resp. \( q_1 \succ q_2 \)). Since \( \succ \) satisfies Strong B-Independence, \( \lambda p_1 + (1 - \lambda) q_1 \succeq \lambda p_2 + (1 - \lambda) q_2 \) (resp. \( \lambda p_1 + (1 - \lambda) q_1 \succ \lambda p_2 + (1 - \lambda) q_2 \)), and by definition of \( P_{\succ} \), we have \( \lambda p + (1 - \lambda) q \in P_{\succ} \) (resp. \( \lambda p + (1 - \lambda) q \in \hat{P}_{\succ} \)).

\[\blacklozenge\]

Proof of Lemma 6 Let \( p \in \hat{P}_{\succ} \) and \( q \notin P_{\succ} \). By definition of \( P_{\succ} \), this means that \( p_1 \succeq p_2 \) and \( q_1 \npreceq q_2 \). But then, by Archimedean B-Continuity, there exists \( \alpha \in (0, 1) \) and \( \beta \in (0, 1) \) such that \( \alpha p_1 + (1 - \alpha) q_1 \succ \alpha p_2 + (1 - \alpha) q_2 \) and \( \beta p_1 + (1 - \beta) q_1 \npreceq \beta p_2 + (1 - \beta) q_2 \). By definition of \( P_{\succ} \), this means that \( \alpha p + (1 - \alpha) q \in P_{\succ} \) and \( \beta p + (1 - \beta) q \notin P_{\succ} \).

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

Proof of Theorem 1 (Necessity of the axioms) Completeness is necessary since the skew symmetry of \( \phi \) guarantees that

\[
\sum_{(x,y) \in X \times X} p(x,y) \phi(x,y) = - \sum_{(x,y) \in X \times X} \bar{p}(x,y) \phi(x,y)
\]

27
so that the expectation of $\phi$ under at least one between $p$ and $\bar{p}$ is weakly larger than 0. For Strong Independence, let $p, q \in P$ (resp. $p \in P$ and $q \in \bar{P}$) and $\lambda \in (0, 1)$. Then

$$
\sum_{(x,y) \in X \times X} (\lambda p + (1 - \lambda) q) (x, y) \phi(x, y)
= \lambda \sum_{(x,y) \in X \times X} p(x, y) \phi(x, y) + (1 - \lambda) \sum_{(x,y) \in X \times X} q(x, y) \phi(x, y) \geq (\text{resp.} >)0.
$$

For Archimedean Continuity, let $p \in \bar{P}$, $q \notin P$. If we define

$$
K := \sum_{(x,y) \in X \times X} p(x, y) \phi(x, y) > 0 > \sum_{(x,y) \in X \times X} q(x, y) \phi(x, y) =: k,
$$

then any $\alpha > \frac{k}{K - k}$ and $\beta < \frac{K}{K - k}$ is easily seen to satisfy the requirements.

(Sufficiency of the axioms) We start by establishing some preliminary claims to streamline the argument.

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

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

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

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

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

and by Strong Independence, we have $p \in P$. \qed

Claim 2 Let $p \in \bar{P}$, $q \notin P$, there exists a unique $\lambda \in (0, 1)$ such that

$$
\lambda p + (1 - \lambda) q \in P \setminus \bar{P}.
$$
Proof We let

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

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 \( \mu p + (1 - \mu) q \in \hat{P} \). This, in turn, implies that \( \mu \in A \).

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

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

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 \( \mu p + (1 - \mu) q \notin P \). 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

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

or

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

In the first case, \( \lambda^* p + (1 - \lambda^*) q \in \hat{P} \), \( q \notin P \) 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. \( \Box \)
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

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

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

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

Suppose by way of contradiction that

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

Then Completeness implies that

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

and by Strong Independence, \( \bar{p} \in \hat{P} \). But this leads to the contradiction \( p \notin P \).

Similarly, suppose by contradiction that

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

Then, Strong Independence implies that \( p \in \hat{P} \), another contradiction. Therefore, we have that

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

30
and by definition of $\hat{P}$

$$\frac{\gamma \lambda \delta_{(y,x)} + (1 - \gamma) (1 - \mu) \delta_{(t,w)}}{\gamma \lambda + (1 - \gamma) (1 - \mu)} \in P \setminus \hat{P}. $$

Thus 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. \hfill \Box

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

**Proof** If $p = \delta_{(x,y)}$ for some $(x, y)$ the result holds by definition of $\hat{P}$. Therefore, suppose that $p$ is supported at least on two joint outcome realizations, let $(x', y') \in \text{supp} \cap \oplus$, and define

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

By Claim 1, $q \in P$. Since

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

Strong Independence implies that $p \in \hat{P}$. \hfill \Box

**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 $\lambda r + (1 - \lambda) s \in P$. \hfill \Box

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, by Completeness, we can assume that there exists $(\hat{x}, \hat{y})$ 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

$$\lambda \delta_{(x,y)} + (1 - \lambda) \delta_{(x,y)} \in P \setminus \hat{P}.$$ 

In this case, let

$$\phi(x, y) = -\phi(x, \hat{y}) \frac{\lambda}{(1 - \lambda)}.$$ 

It only remains to define $\phi$ when $(x, y) \in \hat{\oplus}$. We set

$$\phi(x, y) = -\phi(y, x) \quad \forall (x, y) \in \hat{\oplus}.$$ 

Given the choice of a particular $(x, \hat{y})$, 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 $\phi(x, \hat{y})$ and the values assumed by $\phi$ on the rest of the domain are linear in $\phi(x, \hat{y})$. Suppose that we define $\tilde{\phi}$ starting from a different $(\tilde{x}, \tilde{y}) \in \hat{\oplus}$. 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

$$\tilde{\phi}(\tilde{x}, \tilde{y}) = \phi(\tilde{x}, \tilde{y}) = \phi(\hat{x}, \hat{y}) \frac{\mu}{(1 - \mu)}$$

where

$$\mu \delta_{(x,\tilde{y})} + (1 - \mu) \delta_{(\tilde{y},x)} \in P \setminus \hat{P}$$

and consider $(x, y) \notin \oplus$. Then, by Claim 2 there exist $\lambda_0, \lambda_1$, such that

$$\lambda_0 \delta_{(\hat{x},\tilde{y})} + (1 - \lambda_0) \delta_{(x,y)} \in P \setminus \hat{P},$$

$$\lambda_1 \delta_{(x,\tilde{y})} + (1 - \lambda_1) \delta_{(x,y)} \in P \setminus \hat{P}.$$ 

Given our definitions,

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

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

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

We are going to consider three possible cases.
(First Case) Suppose $\text{supp}p \subseteq \hat{\Phi}$, then by Claim 1, $p \in P$, and by definition of $\phi$, $\phi(x, y) \geq 0$ for every $(x, y) \in \text{supp}p$.

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

(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 $(x_0, y_0), (x_1, y_1) \in \text{supp}p$ with $(x_0, y_0), (y_1, x_1) \in \hat{\Phi}$. Then, by Claim 2 there exists a unique $\alpha \in \mathbb{R}_+$ such that $\alpha \delta_{(x_0, y_0)} + (1 - \alpha) \delta_{(x_1, y_1)} \in P \setminus \hat{P}$.

By Claim 3 and uniqueness up to a positive linear transformation, $\frac{\alpha}{1 - \alpha} \phi(x_0, y_0) = \phi(y_1, x_1)$. If $\text{supp}p = \{(x_0, y_0), (x_1, y_1)\}$, 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

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

Therefore, suppose $\text{supp}p \neq \{(x_0, y_0), (x_1, y_1)\}$. 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\(^{15}\)

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

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

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

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\(^{16}\)

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

\(^{15}\)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

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

\(^{16}\)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) -
Moreover,
\[ p(x_0, y_0) \phi(x_0, y_0) + p(x_1, y_1) \phi(x_1, y_1) + \phi(t, t) p(t, t) = -p(x_0, y_0) \frac{1 - \alpha}{\alpha} \phi(x_1, y_1) + p(x_1, y_1) \phi(x_1, y_1) + 0 \]
\[ = \left( p(x_1, y_1) - \frac{1 - \alpha}{\alpha} p(x_0, y_0) \right) \phi(x_1, y_1) + 0 \]
\[ = p'(x_1, y_1) \phi(x_1, y_1) + \phi(t, t) p'(t, t) \]
so that
\[ \sum_{(x, y) \in \text{supp} p} p(x, y) \phi(x, y) \geq 0 \iff \sum_{(x, y) \in \text{supp} p'} p'(x, y) \phi(x, y) \geq 0. \]

A similar equivalence can be obtained if \( \frac{\alpha}{1 - \alpha} < \frac{p(x_0, y_0)}{p(x_1, y_1)} \). In every instance, the resulting \( p' \) has strictly fewer elements in the support that do not belong to \( P \setminus \hat{P} \) than the original \( p \). Since the support is finite, by repeating this procedure a finite number of times, we are going to obtain a \( \hat{p} \in \Delta(X \times X) \) that falls in one of the first two cases, and such that \( p \in P \Leftrightarrow \hat{p} \in P \) and
\[ \sum_{(x, y) \in \text{supp} p} p(x, y) \phi(x, y) \geq 0 \iff \sum_{(x, y) \in \text{supp} \hat{p}} \hat{p}(x, y) \phi(x, y) \geq 0 \]
concluding the proof.

For every two binary relations \( \succeq \) and \( \succeq \), we simply write \( p' \succeq p'' \succeq p''' \) to mean \( p' \succeq p'' \) and \( p'' \succeq p''' \).

Proof of Lemma 7 (1) Let \( p, q, r \in \Delta(X \times X) \), with \( p_2 = q_1, r_1 = p_1, \) and \( r_2 = q_2, \)
and \( p \in P_{\succeq}, q \in P_{\succeq}. \) By definition of \( P_{\succeq} \), we have \( r_1 = p_1 \succeq p_2 = q_1 \) and \( q_1 \succeq q_2 = r_2. \)
Since \( \succeq \) is Transitive, this implies that \( r_1 \succeq r_2, \) and by definition of \( P_{\succeq}, \) we have \( r \in P_{\succeq}. \)

(2) 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 \( p = p_1 \times p_2, q = p_2 \times p_3, \) 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
\[
\frac{\alpha}{1 - \alpha} p(x_0, y_0),
\]

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

34
r ∈ P by Transitivity of P. Since r was chosen arbitrarily among the joint lotteries with marginals p₁ and p₃, p₁ ≽ P p₂, and the result follows.

(3) ($\succeq^p$ satisfies Transitivity and Completeness $\Rightarrow$ P satisfies Transitivity and Completeness) Let p, q, r ∈ Δ(X × X), with p₂ = q₁, r₁ = p₁, and r₂ = q₂, and p ∈ P₂, q ∈ P₂. Then, Completeness of $\succeq^p$ implies that p₁ ≽ P p₂ ≽ P q₂, and Transitivity of $\succeq^p$ implies r₁ ≽ P r₂, and the definition of $\succeq^p$ implies r ∈ P, that is P satisfies Transitivity. Moreover, P satisfies Completeness by Lemma 4.

(P satisfies Transitivity and Completeness $\Rightarrow$ $\succeq^p$ satisfies Transitivity and Completeness) That $\succeq^p$ satisfies Transitivity follows from the first part. For Completeness, let p₁, p₂ ∈ Δ(X). Define p as the product measure p = p₁ × p₂ ∈ Δ(X × X). By Completeness of P, either p ∈ P or p $\bar{\in}$ P. If p ∈ P, let r ∈ Δ(X × X) be an arbitrary element of Δ(X × X) such that r₁ = p₁ and r₂ = p₂, and define q = p₂ × p₂. By Completeness, q ∈ P, and by Transitivity p ∈ P and q ∈ P together imply that r ∈ P. Since r was chosen arbitrarily among the joint lotteries with marginals p₁ and p₂, p₁ ≽ P p₂. Suppose p = p₂ × p₁ ∈ P. Let r ∈ Δ(X × X) be an arbitrary element of Δ(X × X) such that r₁ = p₂ and r₂ = p₁, and define q = p₁ × p₁. By Completeness, q ∈ P, and by Transitivity p ∈ P and q ∈ P together imply that r ∈ P. Since r was chosen arbitrarily among the joint lotteries with marginals p₂ and p₁, p₂ ≽ P p₁. Therefore, $\succeq^p$ satisfies Completeness.

Given a preference set P that admits a representation as in (4), we define the binary relation $\succeq$ over outcomes as

$$\forall x, y \in X \quad x \succeq y \iff \delta(x, y) \in P.$$ 

We will be interested in whether $\phi$ is modular with respect to this binary relation, i.e.,

$$\forall x, x', y, y' \in X \quad \phi((x, y) \lor (x', y')) + \phi((x, y) \land (x', y')) = \phi(x, y) + \phi(x', y'). \quad (9)$$

Notice that since positive linear transformations preserve modularity, it does not matter which representing $\phi$ we consider.

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

(The representing $\phi$ satisfies (9)) $\Rightarrow$ 4. Let $x_0 \in X$. Define $u(z)$ as $\phi(z, x_0)$. Fix a pair $(z, w)$, with $z \succeq w$. There are three cases:

- $z \succeq w \succeq x_0$. Apply (9) with $x = z$, $y = x'$ and $y' = w$. It reads

$$\phi(z, w) + \phi(x_0, x_0) = \phi(z, x_0) + \phi(x_0, w) \iff$$
$$\phi(z, w) = \phi(z, x_0) - \phi(w, x_0) \iff$$
$$\phi(z, w) = u(z) - u(w)$$

35
where the first coimplication follows from the skew symmetry of $\phi$.

- $z \geq x_0 \geq w$. Apply (9) with $x = z$, $y = w$ and $x_0 = y' = x'$. Again, it reads as:
  \[
  \phi(z, x_0) + \phi(x_0, w) = \phi(z, w) + \phi(x_0, x_0) \iff \\
  \phi(z, w) = u(z) - u(w)
  \]
  where the coimplication follows from the skew symmetry of $\phi$ and the definition of $u$.

- $x_0 \geq z \geq w$. Apply (9) with $x = z$, $y = x' = x_0$ and $y' = w$. Again, it reads as:
  \[
  \phi(x_0, x_0) + \phi(z, w) = \phi(z, x_0) + \phi(x_0, w) \iff \\
  \phi(z, w) = u(z) - u(w)
  \]
  where the coimplication follows from the skew symmetry of $\phi$ and the definition of $u$.

This proves that $\phi(z, w) = u(z) - u(w)$ whenever $z \geq w$. If $w > z$, by skew-symmetry of $\phi$

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

$4 \Rightarrow$ (The representing $\phi$ satisfies (9)). 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 define $\phi(z, w) = (u(z) - u(w))$, modularity holds: let $x, y, x', y' \in X$

$$
\phi ((x, y) \lor (x', y')) + \phi ((x, y) \land (x', y')) \\
= u(x \lor x') - u(y \lor y') + u(x \land x') - u(y \land y') \\
= u(x) + u(x') - u(y) - u(y') \\
= \phi (x, y) + \phi (x', y').
$$

3 $\Leftrightarrow$ 4 is a version of the vN-M EU theorem, see, e.g., page 399 in Ok (2007).

4 $\Rightarrow$ 1 is straightforward given the representation,

4 $\Rightarrow$ 2 holds trivially.

2 $\Rightarrow$(The representing $\phi$ satisfies (9)) and 1 $\Rightarrow$(The representing $\phi$ satisfies (9)) 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

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

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

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

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

$$
\alpha \delta_z + \frac{(1 - \alpha) \delta_x}{2} + \frac{(1 - \alpha) \delta_{x'}}{2} \succeq^P \alpha \delta_w + \frac{(1 - \alpha) \delta_{y'}}{2} + \frac{(1 - \alpha) \delta_y}{2}
$$

nor

$$
\alpha \delta_w + \frac{(1 - \alpha) \delta_{y'}}{2} + \frac{(1 - \alpha) \delta_y}{2} \succeq^P \alpha \delta_z + \frac{(1 - \alpha) \delta_x}{2} + \frac{(1 - \alpha) \delta_{x'}}{2}
$$

holds, and $\succeq^P$ does not satisfy Completeness.
As for the latter, let

\[
\begin{align*}
p &= \alpha \delta_{(z,z)} + \frac{(1 - \alpha) \delta_{(x,x)}}{2} + \frac{(1 - \alpha) \delta_{(x',x')}}{2}, \\
q &= \alpha \delta_{(z,w)} + \frac{(1 - \alpha) \delta_{(x,y')}}{2} + \frac{(1 - \alpha) \delta_{(x',y')}}{2}, \\
r &= \alpha \delta_{(z,w)} + \frac{(1 - \alpha) \delta_{(x,y)}}{2} + \frac{(1 - \alpha) \delta_{(x',y')}}{2}.
\end{align*}
\]

Completeness of \( P \) implies that \( p \in P \), and (11) gives \( q \in P, r \notin P \). However, since \( p_1 = r_1, p_2 = q_1, \) 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 Theorem 2** It immediately follows by combining Lemmata 1, 5, 4, 6 and Proposition 1 and the vN-M Expected Utility Theorem.

Finally, the next lemma highlights that Strong Independence for preference sets weakens the standard notion for binary relations.

**Lemma 8** If \( P \) satisfies Transitivity and Completeness, then \( P \) satisfies Strong B-Independence if and only if \( \succcurlyeq^P \) satisfies Strong Independence.

**Proof** (Only if) First, observe that \( \succcurlyeq^P \) satisfies Completeness by part 3 of Lemma 7. Let \( p_1, p_2, p_3 \in \Delta(X), \alpha \in (0,1) \). Suppose first that \( p_1 \succcurlyeq^P p_2 \). Define \( p \) as the product measure \( p = p_1 \times p_2 \in \Delta(X \times X) \). By definition of \( \succcurlyeq^P \), \( p \in P \). Let \( q = p_3 \times p_3 \). Then \( \alpha p + (1 - \alpha) q \in P \) by Strong Independence of \( P \), and since \( \succcurlyeq^P \) satisfies Completeness by part 3 of Lemma 7, \( \alpha p_1 + (1 - \alpha) p_3 \succcurlyeq^P \alpha p_2 + (1 - \alpha) p_3 \). Now, suppose that it is not true that \( p_1 \succcurlyeq^P p_2 \). Since \( \succcurlyeq^P \) satisfies Completeness, this means that \( p_2 \succ^P p_1 \). By Completeness of \( P \) and the definition of \( \succcurlyeq^P \) this implies that there exists \( r \in P \) with \( r_1 = p_2 \) and \( r_2 = p_1 \). Let \( q = p_3 \times p_3 \). Then \( \alpha r + (1 - \alpha) q \in P \) by Strong Independence of \( P \), and since \( \succcurlyeq^P \) satisfies Completeness, \( \alpha p_1 + (1 - \alpha) p_3 \succcurlyeq^P \alpha p_2 + (1 - \alpha) p_3 \) showing that \( \succcurlyeq^P \) satisfies Strong Independence. Then, Transitivity of \( \succcurlyeq^P \) (guaranteed by Lemma 7) and Lemma 2 implies that \( \succcurlyeq^P \) satisfies Strong B-Independence.

(If) Let \( p, q \in P, \alpha \in (0,1) \). Since \( \succcurlyeq^P \) satisfies Completeness by Lemma 4, \( p_1 \succcurlyeq^P p_2 \) and and \( q_1 \succcurlyeq^P q_2 \). By Strong Independence of \( \succcurlyeq^P \)

\[
\alpha p_1 + (1 - \alpha) q_1 \succcurlyeq^P \alpha p_2 + (1 - \alpha) q_1 \quad \text{and} \quad \alpha p_2 + (1 - \alpha) q_1 \succcurlyeq^P \alpha p_2 + (1 - \alpha) q_2. \tag{12}
\]

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 \(\geq^P\), \(\neg (q_2 \geq^P q_1)\). By Strong Independence and Completeness of \(\geq^P\), this implies that
\[
\neg \left( \alpha p_2 + (1 - \alpha) q_2 \geq^P \alpha p_2 + (1 - \alpha) q_1 \right).
\]
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 \(\alpha \tilde{p} + (1 - \alpha) \tilde{q} \notin \hat{P}\). But since \((\alpha p_1 + (1 - \alpha) q_1) \times (\alpha p_2 + (1 - \alpha) q_1) \in P\) by (12), and Transitivity then implies that \(r \in P\), a contradiction.

**Proof of Remark 1** (Monotone Comparisons implies Monotonicity) Let \(x, y, z \in X\), \(\alpha \in (0, 1)\) and \(p \in \Delta (X \times X)\), with \(x > y\), \(\alpha \delta_{(y,z)} + (1 - \alpha) p \in P_{\geq}\). By definition of \(P_{\geq}\), this implies
\[
\alpha \delta_y + (1 - \alpha) p_1 \geq \alpha \delta_z + (1 - \alpha) p_2.
\]
By Monotone Comparisons, this implies that
\[
\alpha \delta_x + (1 - \alpha) p_1 \succ \alpha \delta_z + (1 - \alpha) p_2
\]
and by definition of \(P_{\geq}\)
\[
\alpha \delta_{(x,z)} + (1 - \alpha) p \in \hat{P}_{\geq}.
\]
(Monotonicity implies Monotone Comparisons) Let \(x, y, z \in X\), \(\alpha \in (0, 1)\) and \(p \in \Delta (X \times X)\), with \(x > y\),
\[
\alpha \delta_y + (1 - \alpha) p_1 \geq \alpha \delta_z + (1 - \alpha) p_2.
\]
By definition of \(P_{\geq}\), this implies \(\alpha \delta_{(y,z)} + (1 - \alpha) p \in P_{\geq}\). By Monotonicity, this implies that \(\alpha \delta_{(x,z)} + (1 - \alpha) p \in \hat{P}_{\geq}\). But then, by definition of \(\hat{P}_{\geq}\), \(\alpha \delta_x + (1 - \alpha) p_1 \succ \alpha \delta_z + (1 - \alpha) p_2\).

**Proof of Remark 2** (If) Let \(x, y, z \in X\) with \(x > y\). By assumption, we have \(\phi (x, z) > \phi (y, z)\). Therefore, for every \(\alpha \in (0, 1)\), \(z \in X\) and \(p \in \Delta (X)\)
\[
\alpha \delta_{(y,z)} + (1 - \alpha) p \in P
\]
\[
\iff \alpha \phi (y, z) + (1 - \alpha) \sum_{(x', y') \in X \times X} p (x', y') \phi (x', y') \geq 0
\]
\[
\implies \alpha \phi (x, z) + (1 - \alpha) \sum_{(x', y') \in X \times X} p (x', y') \phi (x', y') > 0
\]
\[
\implies \alpha \delta_{(x,z)} + (1 - \alpha) p \in \hat{P}.
\]
(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:

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

concluding the proof.

**Proof of Remark 3** (Only if) If \( \phi \) is always equal to 0 the claim is obvious. Therefore, suppose there exists \( z, w \in X \) with \( \phi(z, w) > 0 \). We first prove continuity in the first argument. Then continuity in the second argument follows from skew-symmetry. Let \( (x_n)_{n \in \mathbb{N}} \to x \), and suppose that there exists \( y \in X \) such that \( \phi(x_n, y) \to \phi(x, y) \). There are two cases: there exists an infinite subsequence of \( (x_n)_{k \in \mathbb{N}} \) and an \( \varepsilon > 0 \) such that \( \phi(x_n, y) \geq \phi(x, y) + \varepsilon \) for all \( k \in \mathbb{N} \). If \( \phi(x, y) \geq -\varepsilon \) notice that

\[
\forall k \in \mathbb{N} \quad \frac{\phi(z, w)}{\phi(x, y) + \varepsilon + \phi(z, w)} \phi(x_n, y) + \frac{\phi(x, y) + \varepsilon}{\phi(x, y) + \varepsilon + \phi(z, w)} \phi(w, z) \geq 0
\]

\[
\iff \forall k \in \mathbb{N} \quad \frac{\phi(z, w)}{\phi(x, y) + \varepsilon + \phi(z, w)} \delta_{(x_n,y)} + \frac{\phi(x, y) + \varepsilon}{\phi(x, y) + \varepsilon + \phi(z, w)} \delta_{(w,z)} \in P
\]

\[
\implies \frac{\phi(z, w)}{\phi(x, y) + \varepsilon + \phi(z, w)} \delta_{(x,y)} + \frac{\phi(x, y) + \varepsilon}{\phi(x, y) + \varepsilon + \phi(z, w)} \delta_{(w,z)} \in P
\]

\[
\iff \phi(x, y) \geq \phi(x, y) + \varepsilon
\]

a contradiction. If \( \phi(x, y) < -\varepsilon \) notice that

\[
\forall k \in \mathbb{N} \quad \frac{\phi(w, z)}{\phi(x, y) + \varepsilon + \phi(w, z)} \phi(x_n, y) + \frac{\phi(x, y) + \varepsilon}{\phi(x, y) + \varepsilon + \phi(w, z)} \phi(z, w) \geq 0
\]

\[
\iff \forall k \in \mathbb{N} \quad \frac{\phi(w, z)}{\phi(x, y) + \varepsilon + \phi(w, z)} \delta_{(x_n,y)} + \frac{\phi(x, y) + \varepsilon}{\phi(x, y) + \varepsilon + \phi(w, z)} \delta_{(z,w)} \in P
\]

\[
\implies \frac{\phi(w, z)}{\phi(x, y) + \varepsilon + \phi(w, z)} \delta_{(x,y)} + \frac{\phi(x, y) + \varepsilon}{\phi(x, y) + \varepsilon + \phi(w, z)} \delta_{(z,w)} \in P
\]

\[
\iff \phi(x, y) \geq \phi(x, y) + \varepsilon
\]
a contradiction. The case in which there exists an infinite subsequence of \((x_{n_k})_{k \in \mathbb{N}}\) and an \(\varepsilon > 0\) such that \(\phi(x_{n_k}, y) \leq \phi(x, y) - \varepsilon\) for all \(k \in \mathbb{N}\) is proved along the same lines.

(If) Trivial.

A.1 Salience Characterization

Proof of Proposition 3 Let \(P\) admit a representation as in (4), let \(x, y \in \mathbb{R}\), \(\alpha, \beta \in [0, 1]\) with \(x > y\) and \(\beta > \alpha\) with at least one between \(\alpha\) and \(\beta\) in \((0, 1)\). We have

\[
\left( (x, y), \frac{\beta - \alpha}{1 + \beta - \alpha}; (\alpha x + (1 - \alpha) y, \beta x + (1 - \beta) y), \frac{1}{1 + \beta - \alpha} \right) \in \hat{P}
\]

if and only if

\[
\frac{\beta - \alpha}{1 + \beta - \alpha} \phi(x, y) + \frac{1}{1 + \beta - \alpha} \phi(\alpha x + (1 - \alpha) y, \beta x + (1 - \beta) y) > 0
\]

that by skew symmetry of \(\phi\) is equivalent to

\[
\phi(x, y) (\beta - \alpha) > \phi(\beta x + (1 - \beta) y, \alpha x + (1 - \alpha) y).
\]

(13)

Now, let \(P\) admit a \(\sigma\)-distorted representation. We show that Ordering of \(P\) implies Ordering of \(\sigma\), the other direction being trivial.

We first show that if \(x \geq z > w \geq y\) with \([x, y] \supset [w, z]\), then \(\sigma(x, y) > \sigma(w, z)\). Define \(\alpha = \frac{w-y}{x-y}\) and \(\beta := \frac{z-y}{x-y}\) and notice that \(0 \leq \alpha < \beta \leq 1\) with at least one of the two inequalities being strict. Therefore, (13) implies that

\[
(\beta - \alpha) (x - y) \sigma(x, y) > (\beta - \alpha) (x - y) \sigma(\alpha x + (1 - \alpha) y, \beta x + (1 - \beta) y)
\]

\[
= (\beta - \alpha) (x - y) \sigma(w, z),
\]

and \(\sigma(x, y) > \sigma(w, z)\).

Next, let \(z = w\), with \(x \geq w \geq y\) and at least one of the two inequalities strict, say \(x > w \geq y\). Suppose by way of contradiction that \(\sigma(w, w) \geq \sigma(x, y)\). By the previous part of the proof, \(\sigma\left(\frac{x+w}{2}, y\right) < \sigma(x, y) \leq \sigma(w, w)\). However, by continuity of \(\sigma\), there exists an \(\varepsilon < x-w\) with \(\sigma(w+\varepsilon, w) > \sigma\left(\frac{x+w}{2}, y\right)\). But this is a contradiction with what proved in the previous paragraph.

Proof of Proposition 4 Let \(P\) admit a \(\sigma\)-distorted representation and satisfy Strict Diminishing Sensitivity. Fix \(x > y \geq k > 0\), we have that

\[
p = \left( (x, y - k), \frac{1}{2}; (y, x + k), \frac{1}{2} \right) \in \hat{P}.
\]

41
Given the $\sigma$-distorted representation, this reads as

$$(x - y + k) \sigma (x, y - k) + (y - k - x) \sigma (y, x + k) > 0.$$  

The previous inequality holds if and only if $\sigma (x, y - k) > \sigma (x + k, y)$. The converse is even more immediate. $\blacksquare$

**Proof of Proposition 5** (If) Let $x \geq y \geq 0$ and $k \geq 0$. Consider the two marginal distributions

$$q = \left( x, \frac{1}{2}; y + k, \frac{1}{2} \right) \quad \text{and} \quad q' = \left( x + k, \frac{1}{2}; y, \frac{1}{2} \right).$$

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 $h_z$ defined as

$$h_x = \left( k, \frac{x - y}{(x - y) + k}; (y - x), \frac{k}{(x - y) + k} \right)$$

and

$$h_{y+k} = \left( (x - y), \frac{k}{(x - y) + k}; -k, \frac{(x - y)}{(x - y) + k} \right).$$

Therefore, it is well known that:

$$\sum_{z \in X} q(z) u(z) \geq \sum_{z \in X} q'(z) u(z).$$

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

$$u \left( \frac{x_0 + y_0}{2} \right) \geq \frac{u(x_0) + u(y_0)}{2}$$

42
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.

**Proof of Proposition 6** Let $x, y, w, z \in \mathbb{R}_+$, with $x - y = z - w > 0$. Under a $\sigma$-distorted representation

$$
\left( (x, y), \frac{1}{2}; (w, z), \frac{1}{2} \right) \in \hat{P} \iff \left( (-y, -x), \frac{1}{2}; (-z, -w), \frac{1}{2} \right) \in \hat{P}
$$

is tantamount to

$$(x - y) \sigma (x, y) > (w - z) \sigma (w, z) \iff (x - y) \sigma (-x, -y) < (w - z) \sigma (-w, -z)$$

that is equivalent to

$$\sigma (x, y) > \sigma (w, z) \iff \sigma (-x, -y) > \sigma (-w, -z).$$

The case in which $x - y = z - w < 0$ is completely analogous.

**Proof of Proposition 7** We will prove only the case in which $P$ is risk-averse in $[a, b]$ the other case being completely analogous. 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 $-b \leq -x \leq -y \leq -a$, since $u$ is concave on $[a, b]$, we have

$$u(x) - u \left( \frac{x + y}{2} \right) \leq u \left( \frac{x + y}{2} \right) - u(y)$$

that is

$$\left( \left( x, \frac{x + y}{2} \right), \frac{1}{2}; \left( x, \frac{x + y}{2}, y \right), \frac{1}{2} \right) \notin \hat{P}.$$

By Weak Reflexivity, this means that

$$\left( \left( -\frac{x + y}{2}, -x \right), \frac{1}{2}; \left( -y, -\frac{x + y}{2} \right), \frac{1}{2} \right) \notin \hat{P}$$

or

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

This shows that $u$ is mid-point convex on $[-a, -b]$. 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.
Proof of Theorem 3 (Only If) Given a smooth salience representation, let \( \phi(x, y) = \sigma(x, y) (u(x) - u(y)) \). By the Symmetry axiom for \( \sigma \), we have
\[
\phi(x, y) = \sigma(x, y) (x - y) = \sigma(y, x) (y - x) = -\sigma(y, x) (y - x) = -\phi(y, x)
\]
proving that \( \phi \) is skew-symmetric. Then \( P \) satisfies Completeness, Strong Independence, and Archimedean Continuity by Theorem 1. Since \( \sigma \) satisfies Ordering, \( P \) satisfies Monotonicity by Proposition 2. It satisfies Continuity in Outcomes by Proposition 3 and since \( \phi(y, x) \) is the product of two jointly continuous functions. Finally, let \( x \in X \), \( (x_n)_{n \in \mathbb{N}} \) be such that \( x_n \downarrow x \), \( k \in \mathbb{R} \) and \( \varepsilon \in \mathbb{R}_{++} \). Then
\[
((x, x_n), (1 - (x_n - x)); (k + \varepsilon, k), (x_n - x)) \in P
\]
where the last inequality holds true by continuity of \( \sigma \), proving that \( P \) satisfies the first condition of Continuity at Identity. An analogous argument establishes the second part.

(If) Since \( \phi \) satisfies Completeness, Strong Independence, and Archimedean Continuity by Theorem 1 it admits the representation
\[
p \in P \iff \sum_{(x, y) \in X \times X} p(x, y) \phi(x, y) \geq 0.
\]
Define \( \sigma \) by
\[
\sigma(x, y) = \frac{\phi(x, y)}{x - y} \quad \forall x \neq y
\]
and \( \sigma(x, x) = 0 \) for all \( x \in X \). It is immediate to see that
\[
p \in P \iff \sum_{(x, y) \in X \times X} (x - y) \sigma(x, y) p(x, y) \geq 0.
\]
We know check that \( \sigma \) satisfies all the required properties. First, \( \sigma \) satisfies Symmetry since \( \phi \) is skew symmetric. Moreover since \( \phi \) is continuous by Proposition 3 \( \sigma \) is continuous at every \( (x, y) \) such that \( x \neq y \). We know show that it is continuous at each \( (x, x) \in \mathbb{R} \times \mathbb{R} \). We show that \( x_n \downarrow x \) implies \( \sigma(x_n, x) \to 0 \), the proof for
the case in which $x_n \uparrow x$ is completely analogous. By Continuity at Identity, for all $k \in \mathbb{R}$ and $\varepsilon \in \mathbb{R}_{++}$, there exists an $m \in \mathbb{N}$ such that for all $n \geq m$,

$$(x, x_n), (1 - (x_n - x)), (k + \varepsilon, k), (x_n - x) \in P$$

$$\Leftrightarrow \phi(x, x_n)(1 - (x_n - x)) \leq \phi(k + \varepsilon, k)(x_n - x)$$

$$\Leftrightarrow \sigma(x, x_n)(x - x_n)(1 - (x_n - x)) \leq \sigma(k + \varepsilon, k)\varepsilon(x_n - x)$$

$$\Leftrightarrow \sigma(x, x_n)(1 - (x_n - x)) \leq \sigma(k + \varepsilon, k)\varepsilon.$$

However, by Continuity of Outcomes we can make the RHS $\sigma(k + \varepsilon, k)\varepsilon = \phi(x, y)$ arbitrarily small by choosing a sufficiently small $\varepsilon$, proving that $\sigma(x, x_n)(1 - (x_n - x))$ is converging to 0. With this, Propositions 3,4, and 6 guarantee that $\sigma$ satisfies respectively Ordering, Diminishing Sensitivity, and Weak Reflexivity. \hfill \blacksquare

**Proof of Corollary 4** Let $x > z > y$. Then, there exists $\alpha \in (0, 1)$ with $\lambda x + (1 - \lambda)y = z$. Applying Ordering with $\alpha = \lambda$ and $\beta = 1$ we get

$$\phi(x, y)(1 - \alpha) > \phi(x, z).$$

Applying Ordering with $\beta = \lambda$ and $\alpha = 0$ we get

$$\phi(x, y)\alpha > \phi(z, y).$$

By summing the two inequalities, we get the desired result. \hfill \blacksquare

**A.2 Multiple Alternatives**

**Proof of Proposition 8** We have that

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

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

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

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

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

45
where the first equality follows from the 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, 
\[
\max_{a \in \Delta(A)} \min_{b \in \Delta(A)} \sum_{f \in A} \sum_{g \in A} \sum_{(x,y)} \alpha(f) \beta(g) p_{f,g}(x,y) \phi(x,y) = 0,
\]
that is there exists \( \alpha \in \Delta(A) \) such that for all \( \beta \in \Delta(A) \)
\[
\sum_{f \in A} \sum_{g \in A} \sum_{(x,y)} \alpha(f) \beta(g) p_{f,g}(x,y) \phi(x,y) \geq 0.
\]

B Analysis of the Rank-Based version

In this section, we provide a detailed analysis of the alternative Rank-Dependent Salience Theory proposed in BGS. First, note that every function \( \sigma : X \times X \rightarrow \mathbb{R} \) induces a rank on the support of \( p \). More precisely, if we let
\[
\hat{\sigma}_p(x,y) = \left| \{ (x',y') \in \text{supp} p : \sigma(x',y') > \sigma(x,y) \} \right| \quad \forall (x,y) \in \text{supp} p,
\]
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 5** A preference set \( P \) admits a \((\beta, \sigma)\) representation if there exist a continuous 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
\[
p \in P \Leftrightarrow \sum_{(x,y) \in \text{supp} p} (u(x) - u(y)) \beta^{\hat{\sigma}_p(x,y)} p(x,y) \geq 0. \tag{14}
\]
We say that it admits a Rank-Based Salience Representation if \( \sigma \) is a salience function.

Since \( \beta \leq 1 \), and \( \hat{\sigma}_p \) is decreasing in the salience of a pair of outcomes, the decision criterion is overweighting 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 aggregates 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 \( \beta = 1 \).
B.1 Weakness

Rank-Based Salience Theory captures the idea that the distortion in the evaluation of an event depends only on his relative salience. Hence, small perturbations in the amount paid by an act in a state can result in a dramatic variation in his evaluation. As outlined above, this decision criterion is intransitive, and it does not satisfy the weaker axiom of Transitive Consistency.\footnote{For an in-depth analysis of Transitive Consistency, see Cerreia- Vioglio and Ok (2018), and Nishimura and Ok (2018). For examples of intransitive binary relations satisfying this axiom, see Cerreia- Vioglio, Giarlotta, Greco, Maccheroni and Marinacci (2020).} For a joint distribution $p$ define the conditional row distribution of $p$ given $y$ as

$$
\hat{p}_y(x) = \frac{p(x,y)}{\sum p(x,y)} \quad \forall y \in \text{supp}_2.
$$

Axiom 21 (Transitive Consistency) Let $p,q$ be such that $p_2 = q_2$ and for all $y \in \text{supp}_2$

$$
\hat{p}_y \succeq_{FOSD} q_y
$$

then $q \in P$ implies $p \in P$.

Transitive Consistency is a minimum rationality requirement imposed on an intransitive DM. The underlying idea is that, under the joint distribution $p$, the row marginal has been improved conditional on every possible realization of the column marginal. This implies that $p_1 \succeq_{FOSD} q_1$, and Transitive Consistency is satisfied both by Regret Theory and the Smooth Salience Theory.

The following example illustrates the possible Transitive Inconsistencies of Rank-Based Salience Theory.

Example 2 Let $p$ and $q$ be

\begin{tabular}{ccc}
$p$ & 5 & 11.5 \\
7 & 1/3 & 0 \\
9 & 0 & 2/3 \\
\end{tabular}

\begin{tabular}{ccc}
$q$ & 5 & 11.5 \\
7 & 1/3 & 0 \\
8.8 & 0 & 1/3 \\
9 & 0 & 1/3 \\
\end{tabular}

Suppose that we use the main example of salience function proposed in BGS

$$
\sigma(x,y) = \frac{|x-y|}{|x|+|y|}
$$

17 For an in-depth analysis of Transitive Consistency, see Cerreia-Vioglio and Ok (2018), and Nishimura and Ok (2018). For examples of intransitive binary relations satisfying this axiom, see Cerreia-Vioglio, Giarlotta, Greco, Maccheroni and Marinacci (2020).
we set $\beta = 1/2$, and we let the utility be linear. We obtain

$$\sigma(7, 5) > \sigma(9, 11.5).$$

Therefore $p \notin P$ since

$$\beta^1 \frac{1}{3}[7 - 5] + \beta^2 \frac{2}{3}[9 - 11.5] = \frac{1}{2} \cdot \frac{1}{3} \cdot 2 - \frac{1}{4} \cdot \frac{2}{3} \cdot 2.5 < 0.$$ 

On the other hand, $q \in P$ since

$$\sigma(7, 5) > \sigma(8.8, 11.5) > \sigma(9, 11.5).$$

and

$$\beta^1 \frac{1}{3}[7 - 5] + \beta^2 \frac{1}{3}[8.8 - 11.5] + \beta^3 \frac{1}{3}[9 - 11.5] = \frac{1}{2} \cdot \frac{1}{3} \cdot 2 - \frac{1}{4} \cdot \frac{1}{3} \cdot 2.7 - \frac{1}{8} \cdot \frac{1}{3} \cdot 2.5 > 0.$$ 

B.2 Analysis

Given these issues we have just highlighted, as well as the critique in Kontek (2016), we follow the following approach to the Rank-Based Salience model. We propose testable versions of the Ordering, Diminishing Sensitivity, and Reflexivity properties, and then we show that, under the $(\beta, \sigma)$ representation, these testable versions coincide with the corresponding property postulated by BGS. Finally, we characterize the functional implication in the smooth case covered by the representation in (4).

First, observe that we can characterize Transitivity as a modularity property of the representing functional.

**Proposition 9** If $P$ satisfies Completeness, Strong Independence, and Archimedean Continuity, then $P$ satisfies Transitivity if and only if any $\phi$ representing $P$ as in (4) is modular:

$$\forall x, x', y, y' \in X \quad \phi ((x, y) \lor (x', y')) + \phi ((x, y) \land (x', y')) = \phi (x, y) + \phi (x', y').$$

**Proof** It is an immediate consequence of the proof of Proposition 1. ■

Given this special connection between modularity and the EU model, a natural question is what happens if $\phi$ is assumed to be either supermodular or submodular.
The idea behind the Rank-Based 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 \((x_L, y_H)\), Rank-Based 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 Rank-Based Ordering.

Nevertheless, given Proposition 2, 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

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

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 the row outcome \(x_H\) is larger than the column outcome \(y_L\), a salience sensitive DM prefers (at least weakly) to be paid accordingly to the row outcome. The previous reasoning is crystallized in the Rank-Based Ordering Axiom.

**Axiom 22 (Rank-Based Ordering)** For every

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

we have that

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

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

The next proposition characterizes Rank-Based Ordering in terms of the properties of \(\phi\) and shows that the axiom corresponds to the original property of BGS.
Proposition 10 Suppose that $P$ admits a representation as in in (4). Then it satisfies Rank-Based Ordering (resp. Strict Rank-Based Ordering) if and only if $\phi|_\oplus$ is submodular (resp. strictly submodular).

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

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

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

   
   \[
   \sigma (x, y) > \sigma (y, z) \Rightarrow (u (x) - u (y)) \geq (u (y) - u (z)).
   \]  

Proof 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, $(x \wedge w) \geq (y \lor z)$. Assume for definiteness that $x \geq w$ and $y \geq z$. Then, by the Rank-Based Ordering axiom

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

That is

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

\[
\phi (x, y) + \phi (w, z) \leq \phi (x, z) + \phi (w, y).
\]

Now, suppose that in addition $x > w$ and $y > z$ and $P$ satisfies Strict Rank-Based Ordering. Then,

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

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

\[
\frac{1}{4} \phi (x_H, y_H) + \frac{1}{4} \phi (x_L, y_L) \leq \phi (x_H, y_L) + \phi (x_L, y_H) \iff
\]

\[
\frac{1}{4} \phi (x_H, y_H) + \frac{1}{4} \phi (x_L, y_L) + \frac{1}{4} \phi (y_H, x_H) + \frac{1}{4} \phi (y_L, x_L) \leq 0
\]

50
that is
\[
\frac{1}{4} \delta(x_H, y_L) + \frac{1}{4} \delta(x_L, y_H) + \frac{1}{4} \delta(y_L, x_L) + \frac{1}{4} \delta(y_H, x_H) \in P.
\]

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
\[
\frac{1}{4} \delta(x_H, y_L) + \frac{1}{4} \delta(x_L, y_H) + \frac{1}{4} \delta(y_L, x_L) + \frac{1}{4} \delta(y_H, x_H) \in \hat{P}.
\]

Second part:

1. Let \(x_H \geq x_L \geq y_H \geq y_L\). Ordering of \(\sigma\) implies that either
\[
\sigma(x_H, y_L) \geq \sigma(x_H, y_H) \geq \sigma(x_L, y_L) \geq \sigma(x_L, y_H)
\]
(16)
or
\[
\sigma(x_H, y_L) \geq \sigma(x_L, y_L) \geq \sigma(x_H, y_H) \geq \sigma(x_L, y_H).
\]
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
\[
f(1) = 0\) and \(f(0) = u(x_H) - u(y_L) \geq 0.
\]
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)) < 0\) and \(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 \(\hat{\beta}\) such that \(f'''(\hat{\beta}) < 0\). However,
\[
f'''(\beta) = 3(u(x_L) - u(y_H)) \geq 0.
\]
Therefore $f(\beta) \geq 0$ always, with a strict inequality if $\beta < 1$, so that Strict Rank-Based Ordering holds in that case. To see why equalities in equation (16) 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
\[ \frac{1}{4} \delta(x_H, y_H) + \frac{1}{4} \delta(x_L, y_L) + \frac{1}{4} \delta(y_L, x_L) = \frac{1}{4} \delta(x_H, y_H) + \frac{1}{4} \delta(y_L, x_H) + \frac{1}{4} \delta(y_H, x_H) \]
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 Rank-Based Ordering Axiom can be immediately obtained. For instance, if $\sigma(x, y) < \sigma(x, y') < \sigma(y, y')$, the $(\beta, \sigma)$ representation implies that
\[
\frac{1}{4} (u(y) - u(x)) \beta^2 - \frac{1}{4} (u(y) - u(y')) \beta - \frac{1}{4} (u(y') - u(x)) \geq 0 \Leftrightarrow \\
\frac{1}{4} (u(y) - u(x))(\beta^2 - \beta) - u(x)(\beta^2 - 1) + u(y') (\beta - 1) \geq 0 \Leftrightarrow \\
-\beta u(y) + (\beta + 1) u(x) - u(y') \geq 0 \Leftrightarrow \\
\beta (u(x) - u(y)) + u(x) - u(y') \geq 0
\]
but the last inequality is false since $u$ is strictly increasing. Therefore, let
\[ \sigma(y, y') < \sigma(x, y) < \sigma(x, y') \] (17)
Then, consider the lottery
\[ \frac{1}{4} \delta(y, x) + \frac{1}{4} \delta(y', y) + \frac{1}{4} \delta(y', y') + \frac{1}{4} \delta(x, y') \]
\[ ^{18} \text{The case } y = y' \text{ is completely analogous.} \]
that is in $P$ by the Rank-Based Ordering Axiom. Then, the $(\beta, \sigma)$ representation and (17) imply that

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

da 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

$$
\sigma(x, y) > \sigma(x, y') \text{ and } \sigma(x, y') > \sigma(x', y'),
$$

and therefore $\sigma(x, y) > \sigma(x', y')$. \hfill \qed

The additional condition imposed in (15) 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 $u$.\textsuperscript{19} 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. However, 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.

\textbf{B.2.2 Diminishing Sensitivity and the Rank-Based Criterion}

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 $(x, y)$:

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

\textsuperscript{19}Indeed, by inspection of the proof, we can see that an even weaker, but more involved condition is needed.
Proposition 11 Suppose that $P$ admits a representation as in (4) with a continuously differentiable $\phi$. Then $P$ satisfies Diminishing Sensitivity if and only if

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

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 $\beta < 1$.

Proof (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

$$\phi (x + k, y + k) = \phi (x, y) + \int_{\gamma [(x, y), (x + k, y + k)]} \nabla_{(1,1)}\phi (z_1, z_2) \, dz \leq \phi (x, y)$$

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 $(\beta, \sigma)$ representation and $\sigma$ satisfies Diminishing Sensitivity. 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 $(\beta, \sigma)$ representation and satisfy Diminishing Sensitivity. Fix $x > y \geq k > 0$. Strict Diminishing Sensitivity of $P$ implies that

$$p = \left(\left(\frac{1}{2}, \frac{1}{2}\right); \frac{1}{2}, \frac{1}{2}\right) \in \hat{P}.$$ 

Since $P$ admits a $(\beta, \sigma)$ representation with linear $u$, we have

$$(x - y + k) \beta^\phi(x, y-k) + (y - k - x) \beta^\phi(y, x+k) > 0.$$ 

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 $\sigma(x, y - k) > \sigma(x + k, y)$.

The assumption of linear $u$ used to obtain the "if and only if" statement is maintained in most of the BGS paper.

### B.3 Exponential Parametrization

In this section, we present a flexible and tractable class of preferences that features a representation as in (4) and satisfies the Rank-Based Ordering, Diminishing Sensitivity, and Weak 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, $\lambda \in (0, 1]$. The shape of the function $v$ has a key role in determining the relative weights of Rank-Based Ordering and Diminishing Sensitivity, whereas $\lambda$ can be used to capture loss aversion. In particular, this parametric class can explain the difference between the two versions of the common-consequence Allais paradox presented in BGS: the paradox disappears if it is made explicit that the common consequence is realized in a common contingency.\(^{20}\)

---

\(^{20}\)The subsequent experiment by Esponda and Vespa (2019) shows that the paradox does not completely disappear if the correlation between alternatives is made explicit. However, the fraction of paradoxical choices reduces from above 60% (Kahneman and Tversky, 1979) to around 20%. 55
Example 3 Let $v(x) = \text{Id}(x)$, $\lambda = 1$, so that
\[
\phi(x, y) = \begin{cases} 
\frac{e^x}{e^y} - 1 & x \geq y \\
1 - \frac{e^y}{e^x} & x < y.
\end{cases}
\]

Recall that in the Allais paradox the marginal distributions are:
\[
p_1 = (2500, 0.33; 0, 0.01; z, 0.66) \\
p_2 = (2400, 0.34; z, 0.66).
\]

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

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

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

and this prediction is consistent with the experimental evidence generated in BGS.\[56\]
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.

References


