An Axiomatic Approach to Salience Theory

Giacomo Lanzani

MIT

February 2019
I give an axiomatization to the Salience model of Bordalo, Gennaioli, and Shleifer (henceforth BGS).

The decision-theoretic analysis allows understanding:

- The minimum relaxation of Expected Utility (EU) needed for Salient thinking;
- The functional properties implied by the psychological features of Ordering (Kahneman 2003) and Diminishing Sensitivity (Weber’s Law);
- What is new to previous models;
- It gives us axioms that can be tested in the lab.
The motivation for a Salience Model

Motivation: Allais Paradox

- The DM is asked to choose between the two lotteries

\[ L_1^z = \left( 2500, \frac{33}{100}; 0, \frac{1}{100}; z, \frac{66}{100} \right), \quad L_2^z = \left( 2400, \frac{34}{100}; z, \frac{66}{100} \right). \]

- Accordingly to EU theory, the specific value of \( z \) is irrelevant for the comparison.

- However, we have the following experimental findings

\[ L_1^0 \succ L_2^0 \text{ and } L_2^{2400} \succ L_1^{2400}. \]

- **Prospect Theory** already explains this phenomenon quite well.
Allais Paradox for Acts

Instead, suppose that the correlation between the two acts is made explicit:

\[
\begin{array}{|c|c|c|}
\hline
L_1^z \setminus L_2^z & 2400 & z \\
\hline
0 & 0.01 & 0 \\
2500 & 0.33 & 0 \\
z & 0 & 0.66 \\
\hline
\end{array}
\]

This version makes clear that \(L_1^z\) and \(L_2^z\) pay the common consequence \(z\) in the same state.

Experimental evidence (Conlisk 1989, Birnbaum and Schmidt 2010, BGS 2012):

- Very few DMs are reversing their preferences as \(z\) changes.
- No clear pattern in this reversing.
The motivation for a Salience Model

The difference between the two versions of the experiment cannot be explained either by PT or by more recent models as Cautious Expected Utility.

Roughly speaking, Salience explains the phenomenon by assuming that states with an **higher difference in payoffs draws more attention**. Consider an extreme case where the DM only focus on the state where the difference between the two alternatives is higher:

- for $L_1^0$ and $L_2^0$, it is $(2500, 0)$;
- for $L_2^{2400} \sim L_2^{2400}$ it is $(0, 2400)$;
- in the explicitly correlated case, it is always $(0, 2400)$;
- therefore reversal only in the first case. The same holds for less extreme preferences.

The primary alternative model stressing the role of correlation is **Regret Theory**.
Preferences sets

- $M$ set of possible prizes.
- $p \in \Delta_s (M \times M)$ is a simple (i.e. with finite support) probability measures over the product space $M \times M$.
- We consider a preference set $P \subseteq \Delta_s (M \times M)$, with the following interpretation.
- Let $X, Y$ be two random variables with joint distribution $p \in \Delta_s (M \times M)$.
- I (weakly) prefer to receive the amount specified by $X$ to the amount specified by $Y$ if and only if $p \in P$. 
Since it can explain the Allais Paradox, Salience Theory has to relax some of the EU axioms.

Surprisingly, it is **enough to weaken Transitivity** (maintaining Independence and Continuity) to obtain a Salience Theory representation.
Setup

Maintained Axioms

- Given \( p \in \Delta_s (M \times M) \), define the conjugate distribution \( \bar{p} \) as
  \[
  \bar{p} (x, y) = p (y, x).
  \]

- The strict preference set \( \hat{P} \) is given by those \( p \in P \) such that \( \bar{p} \notin P \). We consider the following axioms for \( P \).

**Completeness** If \( p \notin P \) then \( \bar{p} \in P \).

**Independence** For all \( p \in P, q \in \hat{P}, \lambda \in (0, 1) \) we have \( \lambda p + (1 - \lambda) q \in \hat{P} \).

**Archimedean Continuity** If \( p \in \hat{P} \) and \( q \notin P \), there exists \( \alpha \) and \( \beta \) in \( (0, 1) \) such that
  \[
  \alpha p + (1 - \alpha) q \in \hat{P} \text{ and } \beta p + (1 - \beta) q \notin P.
  \]
Representation Theorem

A function \( \phi : M \times M \rightarrow \mathbb{R} \) is skew-symmetric if \( \phi(x, y) = -\phi(y, x) \).

Theorem

\( P \) satisfies Completeness, Independence and Archimedean Continuity if and only there exists a skew-symmetric \( \phi : M \times M \rightarrow \mathbb{R} \) such that

\[
\Phi(p) := \sum_{x,y} p(x, y) \phi(x, y) \geq 0.
\]

Moreover, \( \phi \) is unique up to a positive linear transformation. In this case, we say that \( P \) admits a skew symmetric additive (SSA) representation.

Recall that under EU

\[
p \in P \iff p_1 \preceq^P p_2 \\
\iff \sum_x p_1(x) u(x) \geq \sum_y p_2(y) u(y) \\
\iff \sum_{x,y} p_1(x) p_2(y) (u(x) - u(y)) \geq 0.
\]
BGS Decision criterion

- We revisit the specific criterion proposed by BGS.
- To identify the Salient pairs of payoff, they use the concept of Salience function.
- $\sigma : \mathbb{R}^2 \rightarrow \mathbb{R}_+$ is a Salience function if it satisfies:

1. **Ordering**: If $x' \leq y'$, $x \leq y$ and $[x', y'] \subseteq [x, y]$, then $\sigma(x, y) \geq \sigma(x', y')$;
2. **Diminishing sensitivity**: if $k, x, y \in \mathbb{R}_+$, then $\sigma(x + k, y + k) \leq \sigma(x, y)$;
3. **Symmetry**: $\sigma(x, y) = \sigma(y, x)$

- Their main example is

$$\sigma(x, y) = \frac{|x - y|}{|x + y + 1|}$$
BGS Decision criterion

- Under BGS Theory of Choice, the DM adopts the following \( \delta - \sigma \) decision criterion:

\[
p \in P \iff \sum_{(x,y)} (x - y) \frac{1}{\delta \sigma(x,y)} p(x, y) \geq 0.
\]

- The main idea is that states with higher Salience are overweighted using a distortion \( \sigma(x, y) \). \( \text{EU} \Rightarrow \delta = 1. \)

**Proposition 1** The Salience Theory model of BGS satisfies Completeness, Independence, and Archimedean Continuity.

- It is easy to see that the former can be embedded in the latter: let

\[
\phi(x, y) = (x - y) \frac{1}{\sigma(x,y)}.
\]
So far we have focused on the *weakenings of EU* necessary for a Salience Theory Representation.

Now, we try to understand the *additional restrictions* implied by Salient Thinking.

In particular, we define *Ordering* and *Diminishing Sensitivity* axiomatically, and we characterize them as properties of $\phi$.

From now on focus on monetary consequences, $M = \mathbb{R}$. 
Axiomatization: Ordering

Definition

We say that $P$ satisfies Ordering if for every $x_H \geq x_L \geq y_H \geq y_L$

we have that

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

The axiom captures the idea that, the best way to make the first component more desirable is to have an event with extremely high difference between outcomes $(x_H, y_L)$.

Proposition 2 If $P$ admits a $\delta$-$\sigma$ representation, it $P$ satisfies Ordering if and only if $\sigma$ satisfies Ordering.

- A strict version of Ordering is not compatible with EU.
Axiomatization: Diminishing Sensitivity

Definition
We say that \( P \) satisfies **Diminishing Sensitivity** if for every \( x \geq y \geq 0 \), and \( k \in \mathbb{R}_+ \)

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

- **Proposition 3** If \( P \) admits a \( \delta\)-\( \sigma \) representation with linear utility, \( P \) satisfies Diminishing Sensitivity if and only if \( \sigma \) satisfies Diminishing Sensitivity.

**Proposition 4** Let \( P \) admit an EU Representation. Then \( P \) satisfies Diminishing Sensitivity if and only if it satisfies **risk-aversion** in the gain domain.

Therefore Diminishing Sensitivity is a generalization of Risk Aversion to Non-Transitive Preferences. What makes Salience Theory incompatible with EU is **Ordering**.
The axiomatic approach takes the revealed choice of the DM as the only observable. Under this approach, there is little distinction between Salience and Regret Theory. However, they are extremely different in terms of neurologic underpinning and welfare implications. Supplementing choice data with neural evidence is essential.