The Perception-Adjusted Luce Model

Federico Echenique Kota Saito Gerelt Tserenjigmid
California Institute of Technology

December 22, 2014

Abstract

We develop an axiomatic theory of random choice that builds on Luce's (1959) model to incorporate a role for perception. We identify agents' "perception priorities" from their violations of Luce's axiom of independence from irrelevant alternatives. Using such perception priorities, we adjust choice probabilities to account for the effects of perception. Our axiomatization requires that the agents' adjusted random choice conforms to Luce's model. The theory can explain the attraction, compromise, similarity effects, and violations of stochastic transitivity, which are very well-documented behavioral phenomena in individual choice. We also discuss implications of the model related the effects of having to choose on what choice is made.

1 Introduction

We study the role of perception in individual random choice. The main novelty is to identify a *perception priority order* from an agent's violations of independence from irrelevant alternatives (IIA), the rationality axiom behind Luce's (1959) model of choice. In other words, we attribute any violation of Luce's model to the role of perception. Our model, a *perception-adjusted Luce model (PALM)*, can explain the attraction, compromise, and similarity effects: these are the best known deviations from Luce's model in experiments.

In our model, an agent perceives alternatives differently. In particular, she goes through the different alternatives in sequence, following a perception priority order. The perception priority order could represent differences in familiarity, or salience, of the different objects of choice. In Psychology, there is clear evidence of the effects of perception, namely salience and perception, on decision making: see for example Shaw and Shaw (1977), Milosavljevic et al. (2012), Towal et al. (2013), Sheng et al. (2005), and Ratneshwar et al. (1987). In our model, each time an alternative is under consideration, it may be chosen with a probability that is obtained from Luce's model. As a result of the priority order, if one adds alternatives to a choice set, the relative probabilities of choosing the various alternatives, may change.

The main idea is that *high priority alternatives are hurt by adding options.* Suppose that you add an alternative z to the choice set consisting of x and y. In principle, the addition of z should intuitively decrease the probability of choosing x, and the probability of choosing y (this monotonicity property is called "regularity" in the literature). But suppose that x is more salient than y, so it has higher priority and is therefore considered before y. Then the addition of z has an indirect *positive* effect on y, in addition to the initial negative effect: Given that y has lower priority than x, y will only be chosen when x is not chosen, and the addition of z has decreased the probability that x is going to be chosen. This is good for y. Hence, the addition of z should hurt x relatively more than y.

Now, in our paper we use the idea we have just mentioned as a way of inferring the perception priority order. The priority order is not observable, but we can use violations of Luce's IIA to infer the priority order. The IIA axiom says that the probability of choosing x relative to that of choosing y is not affected by the presence of alternatives other than x or y. We use an agent's violations of IIA to identify his perception priority order by attributing the effect of additional alternatives on the relative choice probabilities to a priority ranking.

We adjust the agent's random choice using his perception priority order. The priority order defines a *hazard rate*: the probability of choosing an object, conditional on not choosing any of the objects with higher perception priority. So hazard rates incorporate the effects of perception.

The resulting model of choice is termed the *perception-adjusted Luce model* (PALM). PALM connects perception and choice probabilities based on the idea that perception priority captures the direction of violations of Luce's IIA model. In PALM, an agent who is faced with a choice problem considers the different alternatives in order of their priority. Each time one alternative is considered, it is chosen with probability dictated by an underlying Luce model. So the probability that a given alternative is chosen depends both on the its utility (as in Luce) and on its priority in perception.

We show that PALM is characterized by three axioms on choice behavior. The first axiom requires that the agent's perception priority order be well-behaved: it must be complete and transitive. The other two axioms are imposed on hazard rates. The effect of perception on IIA has been accounted for in the hazard rates, so we require that hazard rates satisfy IIA. We also assume that hazard rates satisfy a second standard property of Luce's model, the regularity axiom. The regularity axiom says that the probability of choosing x from a set A cannot be larger than the probability of choosing x from a subset of A.

Despite having a tight axiomatic characterization, PALM is very flexible. PALM can explain many behavioral phenomena, including some of the best known violations of Luce's model in experiments. It can explain attraction effects, compromise effects, similarity effects, and violations of stochastic transitivity: Sections 4.2-4 have all the details. Some of these effects stem from violations of the regularity axiom: PALM can violate the regularity axiom, as regularity is only imposed on hazard rates. In Section 5, we use PALM to explain recent experimental findings on how forcing agents to make a choice affects their choices. Moreover, we discuss necessary and sufficient conditions to have positive correlation between utility and perception.

It may be instructive to illustrate how PALM can accommodate the attraction effect. Doyle et al. (1999) is a representative experiment in evidence of the attraction effect: Doyle et al. present customers in a grocery store in the UK of a choice of baked beans. The first choice is between two types of baked beans: x and y; x is Heinz baked beans, while y is a local cheap brand called Spar. In this problem y was chosen 19% of the time. The authors then introduced a third option, z, which was a more expensive version of the local brand Spar. After z was introduced, y was chosen 33% of the time. This pattern cannot be explained by Luce's model; indeed it cannot be explained by any model of random utility. It can, however, be explained by PALM.

Suppose that perception is related to the familiarity of the beans. Since x is the well-known Heinz brand, it is likely to be the highest priority alternative. Also, y is at least as familiar as z because y and z are the same brands, and z is introduced later. With this perception priority, if the utility of x is large enough, PALM produces the attraction effect in Doyle et al.'s experiment.

As we explained above, the addition of z hurts in principle both x and y, but, while x does not benefit from y's potential decrease, y does benefit from the decrease in the probability of choosing x because y has lower priority than x. The magnitude of this positive effect depends on the utility of x; if the utility of x is large enough, then the indirect positive effect overcomes the direct negative effect, and that is how PALM produces an increase in the probability of choosing y.

There are models within the economic axiomatic literature that explain some of these deviations from Luce's model. We are not aware of any existing model that can explain them all. Section 6 discusses the related literature.

2 Primitives and Luce's model

Let X be a nonempty set of *alternatives*, and $\mathscr{A}$ be the set of finite and nonempty subsets of X.[1] We model an agent who makes a probabilistic choice from $A \cup \{x_0\}$, with $A \in \mathscr{A}$. The element $x_0 \notin X$ represents an outside option that is always available to the agent. Choosing the outside option can simply mean that the agent does not make a choice. We shall also allow the agent to never choose the outside option x_0: the probability of choosing x_0 can be zero for all A.

Definition: A function $\rho : X \cup \{x_0\} \times \mathscr{A} \to [0,1]$ is called *a stochastic choice function* if

$$\sum_{a \in A \cup \{x_0\}} \rho(a, A) = 1$$

for all $A \in \mathscr{A}$. A stochastic choice function ρ is *nondegenerate* if $\rho(a, A) \in (0,1)$ for all $A \in \mathscr{A}$ with $|A| \geq 2$ and $a \in A$.

We write $\rho(B, A)$ for $\sum_{b \in B} \rho(b, A)$, and say that $\rho(\emptyset, A) = 0$.

Note that we allow for $\rho(x_0, A) = 0$. So it is possible that the outside option is never chosen with positive probability, even when ρ is nondegenerate.

Definition: A stochastic choice function ρ *satisfies Luce's independence of irrelevant alternatives (IIA) axiom at* $a, b \in X$ if, for any $A \in \mathscr{A}$,

$$\frac{\rho(a, \{a,b\})}{\rho(b, \{a,b\})} = \frac{\rho(a, A)}{\rho(b, A)}.$$

Moreover, ρ *satisfies IIA* if ρ satisfies IIA at a, b for all $a, b \in X$.

Luce (1959) proves that, if a non-degenerate stochastic choice function satisfies IIA, then it can be represented by the following model (also referred to as multinomial logit):

Definition: ρ satisfies Luce's model if there exists a real-valued function u on $X \cup \mathscr{A}$

[1] Our results hold for smaller classes of subsets of X.

such that

$$\rho(a, A) = \frac{u(a)}{\sum_{a' \in A} u(a') + u(A)}. \tag{1}$$

Luce presented his model in the absence of an outside option. When $\rho(x_0, A) = 0$, then $u(A) = 0$. Here we allow for an outside option, and present the Luce model in which not choosing in A is possible. The number $u(A)$ can be interpreted as the utility of abstaining from choosing an element from A. For later reference it is important to note that

$$u(A) = \sum_{a \in A} u(a) \left(\frac{1}{\sum_{a \in A} \rho(a, A)} - 1 \right). \tag{2}$$

When $u(A) = 0$ for all A, we obtain Luce's model without an outside option.

Luce's model without an outside option satisfies a monotonicity property: $\rho(x, A) \geq \rho(x, B)$, if $A \subset B$. This property is called *regularity.* In general, when an outside option is present, we need an assumption on u in order for Luce's model to satisfy regularity. Luce's model with an outside option may violate regularity.

3 Axioms

We introduce the perception priority order derived from ρ, and the resulting "perception adjusted" random choice function: the hazard rate function.

<u>Perception priority.</u> We capture the role of perception through a weak order $\succsim$. The idea is that when $a \succ b$, then a tends to be perceived sooner than b. In particular, we denote by $\succsim^*$ the (revealed) priority relation that we obtain from the data in ρ. To define $\succsim^*$, first we identify the directly revealed priority relation $\succsim^0$ from ρ. Then $\succsim^*$ is defined as the transitive closure of $\succsim^0$.

We shall attribute all violations of IIA to the role of perception. That is, we require that $a \sim^0 b$ when IIA holds at a and b. In other words, when two alternatives a and b do not exhibit a violation of IIA then we impose that they are equivalent from the view point of perception: they have the same perception priority.

In contrast, if a and b are such that IIA fails at a and b, then we shall require that a and b are strictly ordered by $\succ^0$: we shall require that either $a \succ^0 b$ or that $b \succ^0 a$. Which possibility of the two, $a \succ^0 b$ or $b \succ^0 a$, is determined by the nature of the violation of IIA.

Suppose that IIA fails at a and b because there is some c such that

$$\frac{\rho(a,\{a,b\})}{\rho(b,\{a,b\})} > \frac{\rho(a,\{a,b,c\})}{\rho(b,\{a,b,c\})}. \tag{3}$$

In words, the presence of c lowers the probability of choosing a relative to that of b. When does adding an option hurt one alternative more than another? We claim that *high priority alternatives are hurt by adding options.* The reason is that, by adding certain alternatives c we are "muddying the waters." We are making the choice between a and b less clear than before, and thus diluting the advantage held by a over b.

As we explained in the introduction, we have in mind a model where perception priority dictates the order in which alternatives are considered. Adding c to $\{a,b\}$ would in principle decrease the probability of choosing both a and b; but when a has higher priority than b, then the very fact that a's probability decreases means that b becomes more likely. The alternative b is only chosen when a is not chosen, so the decrease in the probability of choosing a increases the probability of choosing b. Of course, the initial effect of adding c on choosing b is still negative, so the *net* effect on the probability of choosing b is not determined. However, we know unambiguously that $\rho(a,\{a,b\})/\rho(b,\{a,b\}) > \rho(a,\{a,b,c\})/\rho(b,\{a,b,c\})$. And thus the direction of violation of Luce's IIA is dictated by perception priority.

Definition: Let a and b be arbitrary elements in X.

(i)

$$a \sim^0 b \quad \text{if} \quad \frac{\rho(a,\{a,b\})}{\rho(b,\{a,b\})} = \frac{\rho(a,\{a,b,c\})}{\rho(b,\{a,b,c\})},$$

for all $c \in X$;

(ii)

$$a \succ^0 b \quad \text{if} \quad \frac{\rho(a,\{a,b\})}{\rho(b,\{a,b\})} > \frac{\rho(a,\{a,b,c\})}{\rho(b,\{a,b,c\})},$$

for all $c \in X$ such that $c \nsim^0 a$ and $c \nsim^0 b$, and if there is at least one such c. We write $a \succsim^0 b$ if $a \sim^0 b$ or $a \succ^0 b$.

(iii) Define $\succsim^*$ as the *transitive closure* of $\succsim^0$: that is, $a \succsim^* b$ if there exist $c_1, \ldots, c_k \in X$ such that

$$a \succsim^0 c_1 \succsim^0 \cdots c_k \succsim^0 b.$$

The binary relation $\succsim^*$ is called the *revealed perception priority* derived from ρ.

We shall impose the following condition on ρ:

Axiom (Weak Order) The relation $\succsim^*$ derived from ρ is a weak order.

<u>HAZARD RATE.</u> The second important component of our analysis is the hazard rate function. The hazard rate is the probability of choosing an object, conditional on not choosing any of the objects with higher perception priority.

Definition (Hazard Rate): For all $a \in X$ and $A \in \mathscr{A}$, define

$$q(a,A) = \frac{\rho(a,A)}{1-\rho(A_a,A)},$$

where $A_a = \{b \in A | b \succ^* a\}$, $A \in \mathscr{A}$ and $a \in A$. q is called ρ's *hazard rate function.*

We ascribe all violations of IIA to the role of perception, and consider the hazard rate function to be a sort of adjusted choice probability. Hazard rates are adjusted for the role of perception, and hence for the phenomena that in our theory lie behind violations of IIA. We shall therefore impose IIA and regularity as conditions on hazard rates.

Axiom (Hazard Rate IIA) The hazard rate function q satisfies Luce's IIA.

Axiom (Hazard Rate Regularity) $q(a,\{a,b\}) \geq q(a,\{a,b,c\})$, for all $a,b,c \in X$; and $q(a,\{a,b\}) > q(a,\{a,b,c\})$ when $b \nsim^0 c$.

In the regularity axiom above, the first condition is standard. In the second condition, $q(a, \{a, b\}) > q(a, \{a, b, c\})$ is only required when $b \nsim^0 c$. This is because if $b \sim^0 c$, then adding c does not affect the decision maker's perception priority so that it does not affect the corrected choice probability.

4 Theorem

A PALM decision maker is described by two parameters: a weak order $\succsim$ and a utility function u. She perceives each element of a set A sequentially according to the perception priority $\succsim$. Each perceived alternative is chosen with probability described by μ, a function that depends on utility u according to Luce's formula (1). Formally, the representation is as follows.

Definition: A perception-adjusted Luce model (PALM) is a pair $(u, \succsim)$ of a weak order $\succsim$ on X, and a function $u : X \cup \mathscr{A} \to \mathbf{R}$ such that

$$\rho(a, A) = \mu(a, A) \prod_{\alpha \in A/\succsim :\, \alpha \succ a} (1 - \sum_{c \in A :\, c \in \alpha} \mu(c, A)), \tag{4}$$

where

$$\mu(a, A) = \frac{u(a)}{\sum_{b \in A} u(b) + u(A)}.$$

The notation $A/\succsim$ is standard: $A/\succsim$ is the set of equivalence classes in which $\succsim$ partitions A. That is, (i) if $A/\succsim = \{\alpha_i\}_{i \in I}$, then $\cup_{i \in I} \alpha_i = A$; and (ii) $x \sim y$ if and only if $x, y \in \alpha_i$ for some $i \in I$. The notation $\alpha \succ a$ means that $x \succ a$ for all $x \in \alpha$.

For any PALM $(u, \succsim)$, we denote by $\rho_{(u,\succsim)}$ the stochastic choice defined through (4). (When there is no risk of confusion, we write ρ instead of $\rho_{(u,\succsim)}$.)

The PALM has a procedural interpretation. Consider the following procedure. First, take the highest priority alternatives of the menu A, and choose each of them with Luce probability $(\mu(\cdot, A))$. If none of them are chosen, then move on to the second highest priority alternatives and choose each of them with Luce probability.

And so on and so forth.

For example, consider the menu $A = \{x, y, z\}$ with $x \succ y \succ z$. In the PALM, the agent first looks at x and chooses it with Luce probability $\mu(x, A)$. Second, if x is not chosen, she moves on to y and chooses it with Luce probability $\mu(y, A)$. So the probability of choosing y is $\mu(y, A)(1 - \mu(x, A))$. Finally, the probability of choosing z is equal to $\mu(z, A)(1 - \mu(x, A))(1 - \mu(y, A))$. If, instead of having $x \succ y \succ z$, we have that $x \sim y \succ z$ then the probability of choosing z is equal to $\mu(z, A)(1-\mu(x, A)-\mu(y, A))$. The idea is that x and y are perceived, and considered, simultaneously. So the probability of choosing an option that has higher priority than z is $\mu(x, A) + \mu(y, A)$.

Before stating the theorem, we discuss two properties of a PALM model. Firstly, we are interested in stochastic choice for which μ satisfies regularity. As with Luce's model with an outside option, this requires an assumption on utility u.

Definition: A PALM $(u, \succsim)$ is *regular* if for all $a, b, c \in X$, $u(c) \geq u(\{a, b\}) - u(\{a, b, c\})$ and, moreover, $u(c) > u(\{a, b\}) - u(\{a, b, c\})$ if $b \not\succ c$.

Regularity means that, given two alternatives a and b, the impact of c *on the utility of not choosing*, cannot be greater than the utility of c.

The second property is a technical "richness" axiom. Richness requires that X has infinitely many alternatives, but we do not need this assumption to prove the sufficiency of the axioms for the representation. We need it to prove the necessity of the axioms, in particular, the result that $\succsim = \succsim^*$

Axiom (Richness) For any pair (a, b), there is $c \in X$ with $c \succ a$ or $b \succ c$.

Theorem 1 *If a nondegenerate stochastic choice function ρ satisfies Weak Order, Hazard Rate IIA, and Hazard Rate regularity, then there is a regular PALM $(u, \succsim)$ such that $\succsim^* = \succsim$ and $\rho = \rho_{(u, \succsim)}$.*

Conversely, if $(u, \succsim)$ is a regular PALM, and $\succsim$ satisfies Richness, then $\rho_{(u, \succsim)}$ satisfies Weak Order, Hazard Rate IIA, and Hazard Rate regularity, and $\succsim = \succsim^*$.

The proof of the theorem is in Section 7. The sufficiency of the axioms for the representation is straightforward. The converse of Theorem 1 states, not only that PALM satisfies the axioms, but that $\succsim$ must coincide with $\succsim^*$. The perception priority is therefore identified, given data on stochastic choice. Therefore, u is unique up to multiplication by a positive scalar. The bulk of the proof is devoted to establishing that $\succsim = \succsim^*$.

4.1 Discussion of PALM

Luce's model is a special case of PALM, in which $a \sim b$ for all $a, b \in X$. It is useful to compare how Luce and PALM treat the outside option, the probability of not making a choice from a set A.

The utility $u(A)$ for $A \in \mathscr{A}$ has a similar expression to Equation (2), obtained for Luce's model. Indeed,

$$u(A) = \sum_{a \in A} u(a) \left(\frac{1}{\sum_{a \in A} q(a, A)} - 1 \right), \tag{5}$$

with the hazard rates q in place of ρ.

It is interesting to contrast the value of $u(A)$ according to Equation (5) with the utility one would obtain from Equation (2). Given a PALM model $(u, \succsim)$, we can calculate $\hat{u}(A)$ from $\rho_{(u,\succsim)}$ by application of Equation (2). If we do that, we obtain

1. $\hat{u}(A) \geq u(A)$,

2. and $\hat{u}(A) = u(A)$ when $a \sim b$ for all $a, b \in A$.

The inequality $\hat{u}(A) \geq u(A)$ reflects that there are two sources behind choosing the outside option in PALM. One source is the utility $u(A)$ of not making a choice; this is the same as in Luce's model with an outside option. The second source is due to the sequential nature of choice in PALM. When we consider sequentially choosing an option following the priority order $\succsim$, then it is possible that we exhaust the choices in A without making a choice. When that happens, it would seem to inflate (or bias) the value of the outside option; as a result we get that $\hat{u}(A) \geq u(A)$.

Despite the tight behavioral characterization in Theorem 1, PALM is very flexible and can account for some well known behavioral phenomena. We proceed to discuss:

1. similarity and compromise effects (Section 4.2),
2. attraction effect (violations of regularity) (Section 4.3),
3. violations of stochastic transitivity (Section 4.4).

One might suspect that the menu-dependent utility for the outside option is the reason for the flexibility of PALM. However, we can modify PALM so that it has no menu-dependent component and still account for all phenomena.

The presence of the outside option allows us to compare two different environments: an agent has high and low pressure to make choice. The implication of the model is consistent with the experimental result of Dhar and Simonson (2003) on the effects of forced choice on choice (Section 5.1). Moreover, we show that if an agent chooses not to make choice with high probability, then utility and perception are positively correlated (Section 5.2).

4.2 Consistency with Violation of IIA–Similarity Effect and Compromise Effect

The similarity and compromise effects are well-known deviations from Luce's model. See Rieskamp et al. (2006) for a survey. In this section, we demonstrate how PALM can capture each of these phenomena.

The similarity and compromise effects are defined in the same kind of experimental setup. An agent makes choice from the sets $\{x, y\}$ and $\{x, y, z\}$. The "effects" relate to the consequences of adding the alternative z.

4.2.1 Similarity Effect

Suppose that our three alternatives are such that x and z are somehow very similar to each other, and clearly distinct from y. This setup is discussed by Tversky (1972a), building on an example of Debreu (1960). In Debreu's example, x and z are two different recordings of the same Beethoven symphony while y is a suite by Debussy. This effect is called *similarity effect* and can be formalized as follows:

$$\frac{\rho(x, \{x, y, z\})}{\rho(y, \{x, y, z\})} < \frac{\rho(x, \{x, y\})}{\rho(y, \{x, y\})}. \tag{6}$$

The standard explanation of the similarity effect says that x and z somehow jointly compete with y as a "bundle." The agent does not really view x and z as two separate options. Our notion of a perception priority can nicely capture this explanation by assuming that $x \sim z \succ y$. So that x and z are perceived "simultaneously" and before y.

We shall need the following property of μ.

Definition: A hazard rate function μ satisfies *increasing impact at* $(x, y; z)$ if

$$\mu(x, \{x, y\}) - \mu(x, \{x, y, z\}) > \mu(z, \{x, y, z\}).$$

The increasing impact property has a natural implication: the effect on the hazard rate μ of adding z should not be smaller than the magnitude of z. Let us assume that $x \sim z \succ y$: the two Beethoven recordings are more salient from the viewpoint of perception than Debussy.

Proposition 1: *If $x \sim z \succ y$ and μ satisfies increasing impact at $(x, y; z)$, then $\rho_{(u,\succsim)}$ exhibits the similarity effect.*

Proof of Proposition 1: Since $x \succ y$, then $\rho(x, \{x, y\}) = \mu(x, \{x, y\})$ and $\rho(y, \{x, y\}) = \mu(y, \{x, y\})(1 - \mu(x, \{x, y\}))$. Since $x \sim z \succ y$, then $\rho(x, \{x, y, z\}) = \mu(x, \{x, y, z\})$ and $\rho(y, \{x, y, z\}) = \mu(y, \{x, y, z\})(1 - \mu(x, \{x, y, z\}) - \mu(z, \{x, y, z\}))$. Note also that $\frac{\mu(x,\{x,y\})}{\mu(y,\{x,y\})} = \frac{u(x)}{u(y)} = \frac{\mu(x,\{x,y,z\})}{\mu(y,\{x,y,z\})}$. By increasing impact, $\mu(x, \{x, y\}) - \mu(x, \{x, y, z\}) > \mu(z, \{x, y, z\})$; so $1 - \mu(x, \{x, y\}) < 1 - \mu(x, \{x, y, z\}) - \mu(z, \{x, y, z\})$. Hence, $\frac{\rho(x,\{x,y\})}{\rho(y,\{x,y\})} = \frac{u(x)}{u(y)} \frac{1}{1-\mu(x,\{x,y\})} > \frac{u(x)}{u(y)} \frac{1}{1-\mu(x,\{x,y,z\})-\mu(z,\{x,y,z\})} = \frac{\rho(x,\{x,y,z\})}{\rho(y,\{x,y,z\})}$. ■

4.2.2 Compromise Effect

Consider again three alternatives, x, y and z. Suppose that x and z are "extreme" alternatives, while y represents a moderate middle ground, a compromise. In the experiment studied by Simonson and Tversky (1992), x is X-370, a very basic model of Minolta camera; y is MAXXUM 3000i, a more advanced model of the same brand; and z is MAXXUM 7000i, the top of the line offered by Minolta in this class of cameras.

Model	Price ($)	Choices Exp. 1	Choices Exp. 2
x (X-370)	169.99	50%	22%
y (MAXXUM 3000i)	239.99	50 %	57%
z (MAXXUM 7000i)	469.99	N/A	21%

Figure 1: Compromise effect in Simonson and Tversky (1992)

The agent's choice set is $\{x, y\}$ in Experiment 1 and $\{x, y, z\}$ in Experiment 2. The experimental data show that the probability of choosing y increases when moving from Experiment 1 to 2 (see Figure 1). Simonson and Tversky (1992) call this phenomenon the *compromise effect.* As in Rieskamp et al. (2006), the compromise effect can be written as follows:

$$\frac{\rho(x, \{x, y, z\})}{\rho(y, \{x, y, z\})} < 1 \leq \frac{\rho(x, \{x, y\})}{\rho(y, \{x, y\})}. \tag{7}$$

Proposition 2: *When $x \succ y \succsim z$, $\rho_{(u,\succsim)}$ exhibits the compromise effect (i.e., (7)) if and only if $u(y) > u(x)$ and*

$$u(z) + u(\{x, y, z\}) > \frac{u^2(x) - u^2(y) + u(x)u(y)}{u(y) - u(x)} \geq u(\{x, y\}). \tag{8}$$

Proposition 2 results from a straightforward calculation so the proof is omitted. Note that the condition (8) is consistent with regularity: $u(z) + u(\{x, y, z\}) > u(\{x, y\})$.

Simonson and Tversky (1992)'s explanation for the compromise effect is that subjects are averse to extremes, which helps the "compromise" option y when facing the problem $\{x, y, z\}$. PALM can capture the compromise effect when we assume that y is "in between" x and z with respect to priority. One rationale for $x \succ y \succsim z$ is familiarity (see the discussion and references on familiarity in the introduction). The basic camera model may be more familiar, while the top of the line is the least familiar.

A recent marketing study by Mochon and Frederick (2012) attributes the compromise effect to how the options are presented to subjects in experiments. They critique the explanation in Simonson and Tversky, arguing that the effect is not driven by the interpretation of some options as being a compromise. Instead, they find that which options are more salient could lie behind the effect. We note that their point is consistent with PALM as a description for the compromise effect. Mochon and Frederick's explanation could correspond to the assumption of $x \sim z \succ y$, where the extremes are perceived more prominently than the compromise option. In Appendix A.2, we show that PALM can capture the compromise effect with $x \sim z \succ y$.

4.3 Consistency with Violation of Regularity–Attraction Effect

PALM can accommodate violations of regularity. We focus on the attraction effect, a well-known violation of regularity. PALM can explain the attraction effect when one uses *familiarity* to infer perception priority, so the familiar objects are perceived before unfamiliar ones. The role of familiarity in the compromise and attraction effects is documented in Sheng et al. (2005) and Ratneshwar et al. (1987), respectively.

A famous example of the attraction effect is documented by Simonson and Tversky (1992) using the following experiment. Consider our three alternatives again, x, y and z. Suppose now that y and z are different variants of the same good: y is a Panasonic microwave oven (meaning a higher quality and expensive good), while

z is a more expensive version of y: z is dominated by y. The alternative x is an Emerson microwave oven (meaning a lower quality and cheap good). A more recent example, which we discussed in the introduction, is due to Doyle et al. (1999). As we mentioned in the introduction, the findings in Doyle et al.'s experiments fit the story in PALM particularly well.

Option	Choices Exp. 1	Choices Exp. 2
x (Emerson)	57 %	27 %
y (Panasonic I)	43 %	60 %
z (Panasonic II)	N/A	13 %

Figure 2: Attraction effect in Simonson and Tversky (1992)

Simonson and Tversky (1992) (p. 287) asked subjects to choose between x and y in Experiment 1 and to choose among x, y, and z in Experiment 2 (see Figure 2). They found that in Experiment 2, the share of subjects who chose y becomes higher than that in Experiment 1. This effect is called the *attraction effect*. As in Rieskamp et al. (2006), the effect can be described as follows:

$$\rho(y, \{x, y, z\}) > \rho(y, \{x, y\}). \tag{9}$$

Proposition 3: *If $x \succ y \succsim z$ and $u(x)$ is large enough, then $\rho_{(u,\succsim)}$ exhibits the attraction effect (i.e., (9)).*

Proof of Proposition 3: We have

$$\begin{aligned} \rho(y, \{x, y, z\}) > \rho(y, \{x, y\}) &\Leftrightarrow q(y, \{x, y, z\})(1 - q(x, \{x, y, z\})) > q(y, \{x, y\})(1 - q(x, \{x, y\})) \\ &\Leftrightarrow u(x) > \sqrt{(u(y) + u(z) + u(\{x, y, z\}))(u(y) + u(\{x, y\}))} \end{aligned}$$

■

The assumption $x \succ y \succsim z$ means that the Emerson microwave x is more salient than the Panasonic microwaves, perhaps because of its price. The first Panasonic microwave y is at least as salient as z since there are the same brands.[2] In Doyle et al.'s experiments (as discussed in the introduction), perception is related to the familiarity of the brand of beans.

It is natural to consider a symmetric experiment. Consider adding a different alternative, say t, instead of z. The purpose would be to enhance the choice of x. So t could be a more expensive version of x. Heath and Chatterjee (1995) found that one is *less* likely to observe the attraction effect when the third alternative is dominated by the low-quality alternative (x) compared to the high-quality alternative (y). More precisely, one is more likely to have $\rho(y, \{x, y, z\}) > \rho(y, \{x, y\})$ compared to $\rho(x, \{x, y, t\}) > \rho(x, \{x, y\})$. Our model is consistent with this finding: we cannot have $\rho(x, \{x, y, t\}) > \rho(x, \{x, y\})$ when $x \succ y$ because of the regularity of PALM.[3]

4.4 Stochastic Transitivity

Violations of weak stochastic transitivity are well documented in lab experiments. For example, see Tversky (1969), Loomes et al. (1991), and Day and Loomes (2010). In this section, we show that PALM allows for violations of weak stochastic transitivity. We also provide a sufficient condition on PALM that implies weak stochastic transitivity.

Definition (Weak Stochastic Transitivity): For any $a, b, c \in X$, $\rho(a, \{a, b\}) \geq \rho(b, \{a, b\})$ and $\rho(b, \{b, c\}) \geq \rho(c, \{b, c\})$ imply $\rho(a, \{a, c\}) \geq \rho(c, \{a, c\})$.

To see that PALM allows for violations of weak stochastic transitivity, suppose $b \succ c \succ a$, $u(a) = u(c) = x$, $u(b) = \frac{\sqrt{5}-1}{2} x$ for some $x > 0$ and $u(\{a, b\}) = u(\{a, c\}) = u(\{b, c\}) = 0$. Then $\rho(a, \{a, b\}) = \rho(b, \{a, b\}) = \left(\frac{\sqrt{5}-1}{2}\right)^2$ and $\rho(b, \{b, c\}) = \rho(c, \{b, c\}) = \left(\frac{\sqrt{5}-1}{2}\right)^2$, but $\rho(a, \{a, c\}) = \frac{1}{4} < \frac{1}{2} = \rho(c, \{a, c\})$.

[2] It is also reasonable when perception is related to the familiarity of the microwaves. The Emerson microwave x is likely to be the most familiar alternative since it is the cheapest and simplest model. Also, y is at least as familiar as z because y and z are the same brands, and z is introduced later.

[3] However, it is still possible that the relative probability of choosing x increases; that is, $\frac{\rho(x,\{x,y,t\})}{\rho(y,\{x,y,t\})} > \frac{\rho(x,\{x,y\})}{\rho(x,\{x,y\})}$, when t is added where $x \succsim t \succ y$.

However, under reasonable conditions, PALM satisfies weak stochastic transitivity. Formally,

Proposition 4: *Let $a, b, c \in X$. Suppose neither $a \succ b \succ c$, $c \succ a \succ b$, nor $b \succ c \succ a$. If $u(\{a,b\}) = u(\{b,c\}) = u(\{a,c\})$, then ρ satisfies weak stochastic transitivity.*

Proof of Proposition 4: First, we prove the following claim.

Claim: *Take any $x, y \in X$. Let $u(xy) = t$ and $f(z) = \frac{\sqrt{(z+t)^2+4z(z+t)}-(z+t)}{2}$ and $g(z) = \frac{\sqrt{(z-t)^2+4z(z+t)}+(z-t)}{2}$ for all $z \in \mathbb{R}$.*

(i) When $x \sim y$, $\rho(x, \{x,y\}) \geq \rho(y, \{x,y\})$ if and only if $u(x) \geq u(y)$.

(ii) When $x \succ y$, $\rho(x, \{x,y\}) \geq \rho(y, \{x,y\})$ if and only if $u(x) \geq f(u(y))$ if and only if $g(u(x)) \geq u(y)$.

(iii) f and g are increasing functions and $f(z) < z < g(z)$ for any $z > 0$.

Since $x \sim y$, $\rho(x, \{x,y\}) = \frac{u(x)}{u(x)+u(y)+u(xy)} \geq \rho(y, \{x,y\}) = \frac{u(y)}{u(x)+u(y)+u(xy)}$ if and only if $u(x) \geq u(y)$. Hence, (i) holds. When $x \succ y$, $\rho(x, \{x,y\}) = \frac{u(x)}{u(x)+u(y)+t} \geq \rho(y, \{x,y\}) = \frac{u(y)(u(y)+t)}{(u(x)+u(y)+t)^2}$ if and only if $u^2(x) + u(x)(u(y)+t) - u(y)(u(y)+t) \geq 0$ if and only if $u(x) \geq f(u(y))$ if and only if $g(u(x)) \geq u(y)$. Hence, (ii) holds. (iii) follows from a direct calculation.

Proposition follows from the claim directly. Let $u(\{a,b\}) = u(\{b,c\}) = u(\{a,c\}) = t$ and $f(z) = \frac{\sqrt{(z+t)^2+4z(z+t)}-(z+t)}{2}$ and $g(z) = \frac{\sqrt{(z-t)^2+4z(z+t)}+(z-t)}{2}$ for all $z \in \mathbb{R}$.

Consider the case $c \succ b \succ a$. Then by the claim, $\rho(a, \{a,b\}) \geq \rho(b, \{a,b\})$ implies $u(a) \geq g(u(b))$ and $\rho(b, \{b,c\}) \geq \rho(c, \{b,c\})$ implies $u(b) \geq g(u(c))$. Therefore, we obtain $u(a) \geq g(u(b)) \geq g(g(u(c))) > g(u(c))$ which implies $\rho(a, \{a,c\}) > \rho(c, \{a,c\})$ by the claim. The other cases can be proved in the same way (See Appendix A.6). ■

5 Forced Choice and Positive Correlation

5.1 Consistency with Dhar and Simonson (2003)–the Effect of Forced Choice

Dhar and Simonson (2003) run choice experiments in which agents may not have to make a choice. In their design, "no-choice" and "forced choice" are two experimental treatments. Under the no-choice option, subjects can opt not to make a choice. Under the forced-choice treatment, subjects must make a choice. The authors show that the introduction of the no choice option strengthens the attraction effect, weakens the compromise effect, and decrease the relative share of an option that is "average" on all dimensions. In our model, no choice corresponds to choosing the default option x_0. We proceed to illustrate how PALM can capture the evidence presented by Dhar and Simonson (2003).

In the following, instead of simply comparing the no choice and forced choice cases, we will consider the intermediate case in which a decision maker chooses "no choice option" (i.e., the default option x_0) more or less.

Formally, we consider two PALM $\rho = \rho_{(u,\succsim)}$ and $\rho^* = \rho_{(u^*,\succsim^*)}$ which may only differ in $u(A)$, for $A \in \mathscr{A}$. Thus, $u(x) = u^*(x)$ for any $x \in X$ and $\succsim = \succsim^*$. We assume that $\rho(x_0, A) > \rho^*(x_0, A)$ for all $A \in \mathscr{A}$. Roughly speaking, in the PALM ρ^*, a decision maker chooses the default option less often.

Condition ♠: $\rho(x_0, \{x, y\}) > \rho^*(x_0, \{x, y\})$ if and only if $u(\{x, y\}) > u^*(\{x, y\})$ and $\rho(x_0, \{x, y, z\}) > \rho^*(x_0, \{x, y, z\})$ if and only if $u(\{x, y, z\}) > u^*(\{x, y, z\})$.

Condition ♠ has a natural interpretation. Forcing the decision maker to choose means making his utility of the outside option smaller. This would be true in experiments where subjects are not allowed to leave the experiment without making a choice, in which case their payoff must be at least as much as they would make if they were to not participate in the experiment.

Proposition 5 : *Any PALM model satisfies Condition ♠.*

We show Proposition 5 in Appendix A.3.

We now turn to the findings of Dhar and Simonson. Fix three alternatives $x, y, z \in X$. Suppose that $x \succ y \succ z$. So y can be interpreted as an "average" option. Given our assumption on ρ and ρ^*, and under Condition ♠, we assume that $u(\{x, y\}) > u^*(\{x, y\})$ and $u(\{x, y, z\}) > u^*(\{x, y, z\})$.

In the first place, PALM can capture Dhar and Simonson's (2003) finding that the no-choice option decreases the relative share of an average alternative (recall that ρ represents the case where subjects exercise the outside "no-choice" option more):

Proposition 6 :

$$\frac{\rho(y, \{x, y\})}{\rho(x, \{x, y\})} > \frac{\rho^*(y, \{x, y\})}{\rho^*(x, \{x, y\})} \text{ and } \frac{\rho(y, \{x, y, z\})}{\rho(x, \{x, y, z\})} > \frac{\rho^*(y, \{x, y, z\})}{\rho^*(x, \{x, y, z\})}.$$

Proof of Proposition 6: By a direct calculation, $\frac{\rho(x,\{x,y\})}{\rho(y,\{x,y\})} = \frac{u(x)}{u(y)}\left(1 + \frac{u(x)}{u(y)+u(\{x,y\})}\right)$ and $\frac{\rho^*(x,\{x,y\})}{\rho^*(y,\{x,y\})} = \frac{u(x)}{u(y)}\left(1 + \frac{u(x)}{u(y)+u^*(\{x,y\})}\right)$. Since $f(t) = \frac{u(x)}{u(y)}\left(1 + \frac{u(x)}{u(y)+t}\right)$ is decreasing in t, we obtain $u(\{x, y\}) > u^*(\{x, y\})$ if and only if $\frac{\rho(x,\{x,y\})}{\rho(y,\{x,y\})} < \frac{\rho^*(x,\{x,y\})}{\rho^*(y,\{x,y\})}$. Similarly, $\frac{\rho(x,\{x,y,z\})}{\rho(y,\{x,y,z\})} = \frac{u(x)}{u(y)}\left(1+\frac{u(x)}{u(y)+u(z)+u(\{x,y,z\})}\right)$ and $\frac{\rho^*(x,\{x,y,z\})}{\rho^*(y,\{x,y,z\})} = \frac{u(x)}{u(y)}\left(1+\frac{u(x)}{u(y)+u(z)+u^*(\{x,y,z\})}\right)$. Since $g(t) = \frac{u(x)}{u(y)}\left(1+\frac{u(x)}{u(y)+u(z)+t}\right)$ is decreasing in t, we obtain $u(\{x, y, z\}) > u^*(\{x, y, z\})$ if and only if $\frac{\rho(x,\{x,y,z\})}{\rho(y,\{x,y,z\})} < \frac{\rho^*(x,\{x,y,z\})}{\rho^*(y,\{x,y,z\})}$. ■

In second place, PALM can capture Dhar and Simonson's (2003) finding that the no-choice option weakens the compromise effect as follows:

Proposition 7: *If* $u(\{x, y\}) - u(\{x, y, z\}) \geq u^*(\{x, y\}) - u^*(\{x, y, z\})$, *then*

$$\frac{\rho^*(x, \{x, y\})}{\rho^*(y, \{x, y\})} \Big/ \frac{\rho^*(x, \{x, y, z\})}{\rho^*(y, \{x, y, z\})} > \frac{\rho(x, \{x, y\})}{\rho(y, \{x, y\})} \Big/ \frac{\rho(x, \{x, y, z\})}{\rho(y, \{x, y, z\})}.$$

The proof is in Appendix A.4.

5.2 Correlation between Utility u and Perception Priority $\succsim$

Perception and utility are two independent parameters in PALM. Therefore, PALM allows us to model scenarios where perception is positively correlated with utility, negatively correlated, or simply unrelated.

In experimental settings, Reutskaja et al. (2011) find no intrinsic correlation between utility and perception (a similar finding is reported in Krajbich and Rangel (2011)). High-utility items are not *per se* more likely to be perceived more prominently than others. It is therefore important that PALM not force a particular relation between perception and utility.

However, we argue that when an agent chooses the outside option with high probability, it is likely that utility and perception are positively correlated.[4] Now we give some conditions under which $u(a) > u(b)$ if and only if $a \succ b$.

Proposition 8: *Suppose $a \not\sim b$ and $u(a) \neq u(b)$. If $\rho_{(u,\succsim)}(x_0, \{a,b\}) \geq \min\{\rho_{(u,\succsim)}(a, \{a,b\}), \rho_{(u,\succsim)}(b, \{a,b\})\}$ and $u(\{a,b\}) \leq 0$, then $u(a) > u(b)$ if and only if $a \succ b$.*

Proof of Proposition 8: First, we show that if $u(a) > u(b)$ then $a \succ b$. By way of contradiction, suppose $b \succ a$. By calculation, $\rho(a, \{a,b\}) = \frac{u(a)(u(a)+u(\{a,b\}))}{(u(a)+u(b)+u(\{a,b\}))^2}$, $\rho(b, \{a,b\}) = q(b, \{a,b\}) = \frac{u(b)}{u(a)+u(b)+u(\{a,b\})}$, and $\rho(x_0, \{a,b\}) = 1 - \rho(b, \{a,b\}) - \rho(a, \{a,b\}) = \frac{(u(a)+u(\{a,b\}))(u(b)+u(\{a,b\}))}{(u(a)+u(b)+u(\{a,b\}))^2}$.

First consider the case when $\rho(a, \{a,b\}) = \min(\rho(a, \{a,b\}), \rho(b, \{a,b\}))$. Then $\rho(x_0, \{a,b\}) \geq \rho(a, \{a,b\})$ if and only if $\frac{(u(a)+e(\{a,b\}))(u(b)+u(\{a,b\}))}{(u(a)+u(b)+u(\{a,b\}))^2} \geq \frac{u(a)(u(a)+u(\{a,b\}))}{(u(a)+u(b)+u(\{a,b\}))^2}$ if and only if $u(b) + u(\{a,b\}) \geq u(a)$. Therefore, since $u(\{a,b\}) \leq 0$, $\rho(x_0, \{a,b\}) \geq \rho(a, \{a,b\})$ implies $u(b) \geq u(a)$. Contradiction.

Second consider the case when $\rho(b, \{a,b\}) = \min(\rho(a, \{a,b\}), \rho(b, \{a,b\}))$. Then $\rho(x_0, \{a,b\}) \geq \rho(b, \{a,b\})$ if and only if $\frac{(u(a)+u(\{a,b\}))(u(b)+u(\{a,b\}))}{(u(a)+u(b)+u(\{a,b\}))^2} \geq \frac{u(b)}{u(a)+u(b)+u(\{a,b\})}$ if and only if $(u(a)+u(\{a,b\}))(u(b)+u(\{a,b\})) \geq u(b)(u(a)+u(b)+u(\{a,b\}))$. Therefore, since $u(\{a,b\}) \leq 0$, $\rho(x_0, \{a,b\}) \geq \rho(b, \{a,b\})$ implies $(u(a)+u(\{a,b\}))(u(b)+$

[4] Our model implies negative correlation when $u_0(A) = 0$. However, we will not discuss it in detail because it is not robust to small change in choice probabilities.

$u(\{a,b\})) \geq (u(b) + u(\{a,b\}))(u(a) + u(b) + u(\{a,b\}))$, i.e., $u(a) + u(\{a,b\}) \geq u(a) + u(b) + u(\{a,b\})$. Contradiction. Therefore, we proved that $a \succ b$.

Finally, we show that if $a \succ b$ then $u(a) > u(b)$. Suppose $u(b) > u(a)$. Then by the previous part, $u(b) > u(a)$ implies $b \succ a$. Contradiction. ■

The condition that $\rho(x_0, \{a,b\}) \geq \min\{\rho(a, \{a,b\}), \rho(b, \{a,b\})\}$ means that the probability of choosing the outside option must be large enough. This property is necessary to achieve positive correlation, as evidenced in the following result.

Proposition 9: *If $a \succ b$, $u(a) > u(b)$, and $u(b) - u(a) \leq u(\{a,b\})$, then*

$$\rho_{(u,\succsim)}(a, \{a,b\}) > \rho_{(u,\succsim)}(x_0, \{a,b\}) \geq \rho_{(u,\succsim)}(b, \{a,b\}) = \min\{\rho_{(u,\succsim)}(a, \{a,b\}), \rho_{(u,\succsim)}(b, \{a,b\})\}.$$

The proof is in Appendix A.5.

6 Related Literature

Sections 4.2-4 explain how PALM relates to the relevant empirical findings, including the similarity, compromise, and attraction effects. We now proceed to discuss the relation between PALM and some of the most important theoretical models of stochastic choice.

There is a non-axiomatic literature that proposes several models which can explain similarity, compromise and attraction. Rieskamp et al. (2006) is an excellent survey. Examples are Tversky (1972b), Roe et al. (2001) and Usher and McClelland (2004). The latter two papers propose *decision field theory*, which allows for violations of Luce's regularity axiom. The recent paper by Natenzon (2010) presents a learning model, in which an agent learns about the utility of the different alternatives and makes a choice with imperfect knowledge of these utilities. Learning is random, hence choice is stochastic. Natenzon's model can explain all three effects.

We shall not discuss these papers here, and focus instead on the more narrowly related axiomatic literature in economics.

1) The benchmark economic model of rational behavior for stochastic choice is the random utility model. Luce's model is a special case of both PALM and random utility. So PALM and random utility are not mutually exclusive; PALM is, however, not always a random utility model.

The random utility model is described by a probability measure over preferences over X; $\rho(x, A)$ is the probability of drawing a utility that ranks x above any other alternative in A. The random utility model is famously difficult to characterize behaviorally: see the papers by Falmagne (1978), McFadden and Richter (1990), and Barberá and Pattanaik (1986).

There are instances of PALM which violate the regularity axiom. A random utility model must always satisfy regularity. Thus PALM is not a special case of random utility. Moreover, when there is no outside option, Luce's is a random utility model and a special case of PALM. So PALM and random utility intersect, but they are distinct.

2) The recent paper by Gul et al. (2014) presents a model of random choice in which object attributes play a key role. Object attributes are obtained endogenously from the observed stochastic choices. Their model has the Luce form, but it applies sequentially, first for choosing an attribute and then for choosing an object. In terms of its empirical motivation, the model seeks to address the similarity effect.

Gul, Natenzon and Pesendorfer's model is a random utility model (in fact they show that any random utility model can be approximated by their model). There are therefore instances of PALM that cannot coincide with the model in Gul et al. (2014). (Importantly, PALM can explain violations of the regularity axiom.) On the other hand, Luce's model is a special case of their model and of PALM. So the two models obviously intersect.

3) Manzini and Mariotti (2014) study a stochastic choice model where *attention* is the source of randomness in choice. In their model, preferences are deterministic,

but choice is random because attention is random. Manzini and Mariotti's model takes as parameters a probability measure g on X, and a linear order $\succ_M$. Their representation is then

$$\rho(a, A) = g(a) \prod_{a' \succ_M a} (1 - g(a')).$$

In PALM, perception is described by the (non-stochastic) perception priority relation $\succsim$. Choice is stochastic because it is dictated by utility intensities, similarly to Luce's model. In Manzini and Mariotti, in contrast, attention is stochastic, but preference is deterministic.

Manzini and Mariotti's representation looks superficially similar to ours, but the models are in fact different to the point of not being compatible, and seek to capture totally different phenomena. Manzini and Mariotti's model implies that IIA is violated for any pair x and y, so their model is incompatible with Luce's model. PALM, in contrast, has Luce as a special case. Appendix A.1 shows that the two models are disjoint. Any instance of their model must violate the PALM axioms, and no instance of PALM can be represented using their model. So their model and ours seek to capture completely different phenomena.

4) A closely related paper is Tserenjigmid (2013). In this paper, an order on alternative also matters for random choice, and the model can explain the attraction, similarity and compromise effects. The source of violations of IIA is not perception, but instead a sort of menu-dependent utility.

5) The paper by Fudenberg et al. (2013) considers a decision maker who chooses a probability distribution over alternatives so as to maximize expected utility, with a cost function that ensures that probabilities are non-degenerate. One version of their model can accommodate the attraction effect, and one can accommodate the compromise effect.

6) Some related studies use the model of non-stochastic choice to explain some of the experimental results we describe in Sections 4.2-4. This makes them quite

different, as the primitives are different. The paper by De Clippel and Eliaz (2012) is important to mention; it gives an axiomatic foundation for models of non-stochastic choice that can capture the compromise effect. PALM gives a different explanation for the compromise effect, in the context of stochastic choice.

Another related paper is Lleras et al. (2010). (See also Masatlioglu et al. (2012) for a different model of attention and choice.) They attribute violations of IIA to the role of attention. They elicit revealed preference (not perception priority, but preference) in a similar way to ours. When the choice from $\{x, y, z\}$ is x and from $\{x, z\}$ is z, then they conclude that x is revealed preferred to z (this is in some sense, the opposite of the inference we make).

7) Some papers study deliberate stochastic choice due to non-expected utility or uncertainty aversion. Machina (1985) proposes a model of stochastic choice of lotteries. In Machina's paper, an agent deliberately randomizes his choices due to his non-expected utility preferences. Machina does not provide an axiomatization. Saito (2014) axiomatizes a model of stochastic choice of act. In Saito's model, an agent deliberately randomizes his choices because of non-unique priors over the set of states. Saito's primitives is preferences over sets of acts (i.e., payoff-profiles over the set of states).

7 Proof of Theorem 1

7.1 Necessity

We start by proving the converse statement. Let $(u, \succsim)$ be a regular PALM in which $\succsim$ satisfies Richness. Let $\succsim^*$ be derived revealed perception priority from $\rho_{(u,\succsim)}$. We shall first prove that $\succsim^* = \succsim$. The next lemma is useful throughout this section.

Lemma 1 *If $c \succ a \succ b$, or $a \succ b \succ c$, then* $\dfrac{\rho(a, \{a, b, c\})}{\rho(b, \{a, b, c\})} < \dfrac{\rho(a, \{a, b\})}{\rho(b, \{a, b\})}$.

Proof: Let $a \succ b$.

Case 1: $c \succ a \succ b$. Since $b \not\sim c$,

$$\frac{\rho(a,\{a,b,c\})}{\rho(b,\{a,b,c\})}\Big/\frac{\rho(a,\{a,b\})}{\rho(b,\{a,b\})} = \frac{\left(\frac{\mu(a,\{a,b,c\})(1-\mu(c,\{a,b,c\}))}{\mu(b,\{a,b,c\})(1-\mu(c,\{a,b,c\}))(1-\mu(a,\{a,b,c\}))}\right)}{\left(\frac{\mu(a,\{a,b\})}{\mu(b,\{a,b\})(1-\mu(a,\{a,b\}))}\right)}$$
$$= \frac{(1-\mu(a,\{a,b\}))}{(1-\mu(a,\{a,b,c\}))}\left[\frac{u(a)}{u(b)}/\frac{u(a)}{u(b)}\right] < 1,$$

where the last strict inequality is by Regularity.

Case 2: $a \succ b \succ c$. Since $b \not\sim c$,

$$\frac{\rho(a,\{a,b,c\})}{\rho(b,\{a,b,c\})}\Big/\frac{\rho(a,\{a,b\})}{\rho(b,\{a,b\})} = \frac{\mu(a,\{a,b,c\})}{\mu(b,\{a,b,c\})(1-\mu(a,\{a,b,c\}))}\Big/\frac{\mu(a,\{a,b\})}{\mu(b,\{a,b\})(1-\mu(a,\{a,b\}))}$$
$$= \frac{1-\mu(a,\{a,b\})}{1-\mu(a,\{a,b,c\})} < 1,$$

where the last strict inequality is by Regularity. ■

First, we prove $a \sim b$ if and only if $a \sim^* b$. Then, we prove $a \succ b$ if and only if $a \succ^* b$.

Lemma 2 *$a \sim b$ if and only if $a \sim^* b$.*

Proof of Lemma 2:

Step 1: If $a \sim b$ then $a \sim^* b$.

Proof of Step 1: Fix $c \in X$ to show $\frac{\rho(a,\{a,b,c\})}{\rho(b,\{a,b,c\})}/\frac{\rho(a,\{a,b\})}{\rho(b,\{a,b\})} = 1$.

Case 1: $a \sim b \succsim c$.

$$\frac{\rho(a,\{a,b,c\})}{\rho(b,\{a,b,c\})}\Big/\frac{\rho(a,\{a,b\})}{\rho(b,\{a,b\})} = \frac{\mu(a,\{a,b,c\})}{\mu(b,\{a,b,c\})}\Big/\frac{\mu(a,\{a,b\})}{\mu(b,\{a,b\})} = \frac{u(a)}{u(b)}\Big/\frac{u(a)}{u(b)} = 1.$$

Case 2: $c \succ a \sim b$.

$$\frac{\rho(a,\{a,b,c\})}{\rho(b,\{a,b,c\})}\Big/\frac{\rho(a,\{a,b\})}{\rho(b,\{a,b\})} = \frac{\mu(a,\{a,b,c\})(1-\mu(c,\{a,b,c\}))}{\mu(b,\{a,b,c\})(1-\mu(c,\{a,b,c\}))}\Big/\frac{\mu(a,\{a,b\})}{\mu(b,\{a,b\})} = \frac{u(a)}{u(b)}\Big/\frac{u(a)}{u(b)} = 1.$$

□

Step 2: If $a \succ b$, then $a \nsim^0 b$.

Proof of Step 2: By Richness, there is c with $c \succ a \succ b$ or $a \succ b \succ c$. In either case, by Lemma 1, $\frac{\rho(a,\{a,b,c\})}{\rho(b,\{a,b,c\})} < \frac{\rho(a,\{a,b\})}{\rho(b,\{a,b\})}$. Hence, $a \nsim^0 b$. □

Step 3: If $a \succsim^0 b$, then $a \succsim b$.

Proof of Step 3: We show that if $a \not\succsim b$ then $a \not\succsim^0 b$. Let $a \not\succsim b$. Then by completeness, $b \succ a$. By Richness, there is c with $c \succ b \succ a$ or $b \succ a \succ c$. Suppose without loss of generality that $c \succ b \succ a$. By Lemma 1, we have $\frac{\rho(b,\{a,b,c\})}{\rho(a,\{a,b,c\})} < \frac{\rho(b,\{a,b\})}{\rho(a,\{a,b\})}$. Moreover, since $c \succ b$ and $c \succ a$, Step 2 shows that $c \nsim^0 a$ and $c \nsim^0 b$. Hence, $b \succ^0 a$, so that $a \not\succsim^0 b$. □

Step 4: If $a \sim^* b$ then $a \sim b$.

Proof of Step 4: Let $a \sim^* b$. By the definition of $\sim^*$, $a \succsim^* b$ and $b \succsim^* a$. Then $a \succsim^* b$ implies that there exist $c_1, \ldots, c_k$ such that $a = c_1 \succsim^0 c_2 \succsim^0 \ldots \succsim^0 c_k = b$. By Step 3 and the transitivity of $\succsim$, we have that $a \succsim b$. Similarly, $b \succsim^* a$ implies that $b \succsim a$. Thus $a \sim b$. □

■

In the following, we prove that $a \succ b$ if and only if $a \succ^* b$.

Lemma 3 *If $a \succ^* b$ then $a \succ b$.*

Proof: Let $a \succ^* b$. It suffices to consider the following two cases.

Case 1: $a \succ^0 b$. Suppose, towards a contradiction, $a \not\succ b$. By the completeness of $\succsim$, $b \succsim a$. Note that $a \succ^0 b$ implies $a \nsim^0 b$, so $a \nsim b$ by Lemma 2. Then $b \succ a$. By Richness there is c such that $c \succ b \succ a$ or $b \succ a \succ c$. In either case, $\frac{\rho(a,\{a,b,c\})}{\rho(b,\{a,b,c\})} / \frac{\rho(a,\{a,b\})}{\rho(b,\{a,b\})} > 1$ by Lemma 1, in contradiction with $a \succ^0 b$.

Case 2: There exist $c_1, \ldots, c_k \in X$ such that $a \succ^0 c_1 \succ^0 \cdots \succ^0 c_k \succ^0 b$. Then, by the proof in Case 1, $a \succ c_1 \succ \cdots \succ c_k \succ b$. Hence, by transitivity, $a \succ b$. ■

The next lemma shows the converse.

Lemma 4 *If $a \succ b$ then $a \succ^* b$.*

Proof: To simplify the exposition, we use the following notation in this proof: $a \vdash b$ if $a \succ b$ and there is no $c \in X$ with $a \succ c \succ b$.

Let $a \succ b$.

Case 1: $a \vdash b$. It suffices to show that $a \succ^0 b$. By Richness, there exists c such that $c \succ a \succ b$ or $a \succ b \succ c$; so there is at least one c such that $a \not\sim c$ and $b \not\sim c$. By Lemma 2, $a \not\sim^* c$ and $b \not\sim^* c$.

Choose any $d \in X$ such that $a \not\sim^* d$ and $b \not\sim^* d$. By Lemma 2, $a \not\sim d$ and $b \not\sim d$. Since $a \vdash b$, it is not true that $a \succ d \succ b$. That is, either $d \succ a$ or $b \succ d$. Since $a \succ b$, then $d \succ a \succ b$ or $a \succ b \succ d$. In either case, by Lemma 1, $\frac{\rho(a,\{a,b,c\})}{\rho(b,\{a,b,c\})} / \frac{\rho(a,\{a,b\})}{\rho(b,\{a,b\})} < 1$. Thus $a \succ^0 b$. Hence, $a \succ^* b$.

Case 2: $a \not\vdash b$. There exist $c_1, \ldots, c_k \in X$ such that $a \vdash c_1 \vdash \cdots \vdash c_k \vdash b$. By the argument in Case 1, $a \succ^0 c_1 \succ^0 \cdots \succ^0 c_k \succ^0 b$. Therefore, $a \succ^* b$. ■

Finally, to complete the proof of the necessity, we prove that ρ satisfies Hazard Rate Regularity if and only if $\rho_{(u,\succsim)}$ satisfies the regularity. Since $\frac{1}{\mu(a,\{a,b,c\})} - \frac{1}{\mu(a,\{a,b\})} = \frac{u(a)+u(b)+u(c)+u(\{a,b\})}{u(a)} - \frac{u(a)+u(b)+u(\{a,b\})}{u(a)}$, $\mu(a,\{a,b\}) \geq \mu(a,\{a,b,c\})$ if and only $u(c) \geq u(\{a,b\}) - u(\{a,b,c\})$. Moreover, $b \nsim c$ if and only if $b \not\sim^0 c$.

7.2 Sufficiency

In this section, we prove sufficiency. Choose a nondegenerate stochastic choice function ρ that satisfies the axioms in the theorem. Let $\succsim^*$ be the derived revealed perception priority.

For all $A \in \mathscr{A}$ and $a \in A$, define

$$\nu(a, A) = \frac{q(a, A)}{\sum_{a \in A} q(a, A)}.$$

Since ρ is nondegenerate, $1 > \rho(a, A) > 0$ for all $a \in A$. Remember that $A_a = \{b \in A | b \succ^0 a\}$. Since $a \notin A_a$, $1 - \rho(A_a, A) > 0$. Hence, q is well defined. Moreover, $q(a, A) > 0$ because ρ is non-degenerate. Thus, ν is also well defined.

Step 1: There exists $u : X \to \mathbb{R}_{++}$ such that $q(a, A) = \frac{u(a)}{\sum_{a' \in A} u(a')} \sum_{a \in A} q(a, A)$.
Proof of Step 1: First, we show that ν satisfies Luce's IIA. For all $a, b, c \in X$

$$\frac{\nu(a, \{a, b\})}{\nu(b, \{a, b\})} = \frac{q(a, \{a, b\})}{q(b, \{a, b\})} = \frac{q(a, \{a, b, c\})}{q(b, \{a, b, c\})} = \frac{\nu(a, \{a, b, c\})}{\nu(b, \{a, b, c\})}.$$

Moreover, $\sum_{a \in A} \nu(a, A) = 1$. Therefore, by Luce's theorem (Luce (1959)), there exists $u : X \to \mathbb{R}_{++}$ such that $\nu(a, A) = \frac{u(a)}{\sum_{a' \in A} u(a')}$. Hence, by definition, $q(a, A) = \nu(a, A) \sum_{a \in A} q(a, A) = \frac{u(a)}{\sum_{a' \in A} u(a')} \sum_{a \in A} q(a, A)$. □

For all $A \in \mathscr{A}$, define

$$u(A) = \sum_{a \in A} u(a) \Big(\frac{1}{\sum_{a \in A} q(a, A)} - 1 \Big).$$

Step 2: $q(a, A) = \frac{u(a)}{\sum_{a' \in A} u(a') + u(A)}$.
Proof of Step 2: By a direction calculation,

$$\begin{aligned}
\frac{\sum_{a' \in A} u(a') + u(A)}{u(a)} &= \frac{\sum_{a' \in A} u(a')}{u(a)} + \frac{\sum_{a' \in A} u(a')}{u(a)} \Big(\frac{1}{\sum_{a' \in A} q(a', A)} - 1 \Big) \\
&= \frac{\sum_{a' \in A} u(a')}{u(a)} \frac{1}{\sum_{a' \in A} q(a', A)} \\
&= \frac{1}{q(a, A)},
\end{aligned}$$

where the last equality holds by Step 1. □

Step 3: $\rho = \rho_{(u, \succsim^*)}$.

Proof of Step 3: Choose any $A \in \mathscr{A}$. Since $\succsim^*$ is a weak order, therefore the indifference relation $\sim^*$ is transitive. Then, the set of equivalence classes $A/_{\succsim^*}$ is well defined and finite. That is, there exists a partition $\{\alpha^1, \alpha^2, \dots \alpha^k\}$ of A such that $a_j \succ^* a_i$ for all $a_i \in \alpha^i$ and $a_j \in \alpha^j$ with $j > i$ and $a_i \sim^* a_{i'}$ for all $a_i, a_{i'} \in \alpha^i$.

Define $p_i \equiv \rho(\alpha^i, A) = \sum_{a' \in \alpha^i} \rho(a', A)$. Then for $a \in \alpha^i$, $q(a, A) = \frac{\rho(a,A)}{1-\sum_{j>i} p_j}$. Therefore,

$$\sum_{a \in \alpha^i} q(a, A) = \sum_{a \in \alpha^i} \frac{\rho(a, A)}{1 - \sum_{j>i} p_j} = \frac{\sum_{a \in \alpha^i} \rho(a, A)}{1 - \sum_{j>i} p_j} = \frac{p_i}{1 - \sum_{j=i+1}^{k} p_j}.$$

Hence,

$$1 - \sum_{a \in \alpha^i} q(a, A) = 1 - \frac{p_i}{1 - \sum_{j=i+1}^{k} p_j} = \frac{1 - \sum_{j=i+1}^{k} p_j - p_i}{1 - \sum_{j=i+1}^{k} p_j} = \frac{1 - \sum_{j=i}^{k} p_j}{1 - \sum_{j=i+1}^{k} p_j}.$$

Therefore, for any $s \in \{1, \dots, k\}$,

$$\prod_{i=s+1}^{k} (1 - \sum_{a \in \alpha^i} q(a, A)) = \prod_{i=s+1}^{k} \frac{1 - \sum_{j=i}^{k} p_j}{1 - \sum_{j=i+1}^{k} p_j} = \frac{1 - \sum_{j=s+1}^{k} p_j}{1} = 1 - \rho(A_a, A).$$

For all $a \in A$ and $A \in \mathscr{A}$, define $\mu(a, A) = q(a, A)$.

Choose $a \in A$. Without loss of generality assume that $a \in \alpha^s$. Then, $\rho(a, A) = q(a, A)(1 - \rho(A_a, A)) = \mu(a, A)(1 - \rho(A_a, A)) = \mu(a, A) \prod_{i=s+1}^{k} (1 - \sum_{a' \in \alpha^i} \mu(a', A)) \equiv \rho_{(u, \succsim^*)}(a, A)$. □

■

A Appendix: Supplements

A.1 Relation to Manzini and Mariotti

The model of Manzini and Mariotti (2014) is specified by a probability measure g on X, and a linear order $\succ_M$. Their representation is then

$$\rho(a, A) = g(a) \prod_{a' \succ_M a} (1 - g(a')).$$

Superficially, this representation looks similar to ours, but it is actually very different: It is incompatible with our model, in the sense that the set of stochastic choices that satisfy our model is disjoint from the set of stochastic choices in Manzini and Mariotti's model. We now proceed to prove this fact.

Let ρ have a Manzini and Mariotti (2014) representation as above and let X have at least three elements. Suppose, towards a contradiction that it also has a representation using our model.

We are going to prove that the two models differ in a strong sense, because we are going to show that there is no subset of X of three elements on which the two models can coincide.

Let $a, b, c \in X$. The preference relation $\succ_M$ is a linear order. Suppose, without loss of generality, that $a \succ_M b \succ_M c$. Given the Manzini-Mariotti representation, then

$$\rho(a, \{a, b, c\}) = \rho(a, \{a, b\}) = \rho(a, \{a, c\}) = g(a),$$

and

$$\rho(b, \{a, b, c\}) = \rho(b, \{a, b\}) = g(b)(1 - g(a)).$$

We have assumed that ρ has a PALM representation given by some $(u, \succsim)$. Now consider how a, b, c are ordered by $\succsim$.

There are seven cases to consider; each one of these cases end in a contradiction.

1. $a \succsim b$, $a \succsim c$, and $b \not\sim c$: By Regularity, since $b \not\sim c$, $\rho(a, \{a, b, c\}) =$

$q(a, \{a, b, c\}) < \rho(a, \{a, b\}) = q(a, \{a, b\})$.

2. $b \succsim a$, $b \succsim c$ and $a \not\sim c$: By Regularity, since $a \not\sim c$, $\rho(b, \{a, b, c\}) = q(b, \{a, b, c\}) < \rho(b, \{a, b\}) = q(b, \{a, b\})$.

3. $c \succ a \succsim b$: By Regularity, $\rho(a, \{a, b, c\}) = q(a, \{a, b, c\})(1 - q(c, \{a, b, c\})) < q(a, \{a, b, c\}) \leq q(a, \{a, b\}) = \rho(a, \{a, b\})$.

4. $a \succ b \sim c$: By Regularity, since $\rho(a, \{a, b, c\}) = q(a, \{a, b, c\}) = \rho(a, \{a, b\}) = q(a, \{a, b\})$ and $q(b, \{a, b, c\}) < q(b, \{a, b\})$ because $a \not\sim c$, $\rho(b, \{a, b, c\}) = q(b, \{a, b, c\})(1 - q(a, \{a, b, c\})) < \rho(b, \{a, b\}) = q(b, \{a, b\})(1 - q(a, \{a, b\}))$.

5. $b \succ a \sim c$: By Regularity, $\rho(a, \{a, b, c\}) = q(a, \{a, b, c\})(1 - q(b, \{a, b, c\})) < q(a, \{a, b, c\}) \leq q(a, ac) = \rho(a, ac)$.

6. $c \succ b \succ a$: By Regularity, $\rho(b, \{a, b, c\}) = q(b, \{a, b, c\})(1 - q(c, \{a, b, c\})) < q(b, \{a, b, c\}) \leq \rho(b, \{a, b\}) = q(b, \{a, b\})$.

7. $a \sim b \sim c$: In this case, Luce's IIA is cannot be violated in PALM. However, in Manzini and Marriott's Model, there is always at least one violation of Luce's IIA.

A.2 Alternative Condition for Compromise Effect

In Proposition 2, we show that PALM can capture the compromise effect under that condition that $x \succ y \succsim z$.

Remember our example of the compromise effect in which x and z are the extremes and y is the compromised option. One may think that the extremes are perceived more prominently than the compromised option, and hence, $x \sim z \succ y$.

Proposition A2: *When $x \sim z \succ y$, ρ exhibits the compromise effect at $\{x, z\}$ with respect to y if and only if*

$$u(y) > u(x) \text{ and } u(\{x, y, z\}) - \frac{u(x)u(z)}{u(y) - u(x)} > \frac{u^2(x) - u^2(y) + u(x)u(y)}{u(y) - u(x)} > u(\{x, y\}).$$

The proof is by straightforward calculations.

A.3 Supplement to Section 5.1

Observation 1: For any PALM ρ,

$$\rho(x_0, \{x, y\}) > \rho^*(x_0, \{x, y\}) \text{ if and only if } u(\{x, y\}) > u^*(\{x, y\})$$

Proof: By a direct calculation, $\rho(x_0, \{x, y\}) = \frac{\big(u(x)+u(\{x,y\})\big)\big(u(y)+u(\{x,y\})\big)}{\big(u(x)+u(y)+u(\{x,y\})\big)^2}$ and $\rho^*(x_0, \{x, y\}) = \frac{\big(u(x)+u^*(\{x,y\})\big)\big(u(y)+u^*(\{x,y\})\big)}{\big(u(x)+u(y)+u^*(\{x,y\})\big)^2}$. Let $g(t) = \frac{\big(u(x)+t\big)\big(u(y)+t\big)}{\big(u(x)+u(y)+t\big)^2}$. Since $g'(t) = \frac{t\big(u(x)+u(y)\big)+u^2(x)+u^2(y)}{\big(u(x)+u(y)+t\big)^3}$, g is increasing in t when $t > -\frac{u^2(x)+u^2(y)}{u(x)+u(y)}$.

Now it is enough to prove that $u(\{x, y\})$ is larger than $-\frac{u^2(x)+u^2(y)}{u(x)+u(y)}$. First, $\rho(y, \{x, y\}) = \frac{u(y)\big(u(y)+u(\{x,y\})\big)}{\big(u(x)+u(y)+u(\{x,y\})\big)^2} > 0$ implies that $u(\{x, y\}) > -u(y)$. Second, $\rho(x_0, \{x, y\}) = \frac{\big(u(x)+u(\{x,y\})\big)\big(u(y)+u(\{x,y\})\big)}{\big(u(x)+u(y)+u(\{x,y\})\big)^2} \geq 0$ and $u(\{x, y\}) > -u(y)$ imply $u(\{x, y\}) \geq -u(x)$. Then we obtain $u(\{x, y\}) > -\frac{u(x)+u(y)}{2} \geq -\frac{u^2(x)+u^2(y)}{u(x)+u(y)}$. ■

Observation 2: For any PALM ρ,

$$\rho(x_0, \{x, y, z\}) > \rho^*(x_0, \{x, y, z\}) \text{ if and only if } u(x, \{x, y, z\}) > u^*(x, \{x, y, z\}).$$

Proof: By a direct calculation,

$$\rho(x_0, \{x, y, z\}) = \frac{\big(u(x) + u(y) + u(\{x, y, z\})\big)\big(u(x) + u(z) + u(\{x, y, z\})\big)\big(u(y) + u(z) + u(\{x, y, z\})\big)}{\big(u(x) + u(y) + u(z) + u(\{x, y, z\})\big)^3}$$

and $p^*(x_0, \{x, y, z\}) = \frac{\big(u(x)+u(y)+u^*(\{x,y,z\})\big)\big(u(x)+u(z)+u^*(\{x,y,z\})\big)\big(u(y)+u(z)+u^*(\{x,y,z\})\big)}{\big(u(x)+u(y)+u(z)+u^*(xyz)\big)^3}$.

Let $s(t) = \frac{\big(u(x)+u(y)+t\big)\big(u(x)+u(z)+t\big)\big(u(y)+u(z)+t\big)}{\big(u(x)+u(y)+u(z)+t\big)^3}$. Also, let $A \equiv u(x)+u(y)+u(z)$, $B \equiv u^2(x)+u^2(y)+u^2(z)+u(x)u(y)+u(x)u(z)+u(y)u(z)$, and $C \equiv u^3(x)+u^3(y)+$

$u^3(z) + u^2(x)u(y) + u^2(x)u(z) + u^2(y)u(x) + u^2(y)u(z) + u^2(z)u(x) + u^2(z)u(y) + 3u(x)u(y)u(z)$.

Then we obtain $s'(t) = \frac{t^2 \cdot A + 2t \cdot B + C}{(t+A)^4}$ which implies that s is increasing when $t > -\frac{B - \sqrt{B^2 - AC}}{A}$.

Now it is enough to prove that $u(\{x, y, z\})$ is larger than $-\frac{B-\sqrt{B^2-AC}}{A}$. First, $\rho(y, \{x, y, z\}) = \frac{u(y)\big(u(y)+u(z)+u(\{x,y,z\})\big)}{\big(u(x)+u(y)+u(z)+u(\{x,y,z\})\big)^2} > 0$ implies $u(\{x, y, z\}) > -\big(u(y) + u(z)\big)$. Second,

$\rho(z, \{x, y, z\}) = \frac{u(z)\big(u(y)+u(z)+u(\{x,y,z\})\big)\big(u(x)+u(z)+u(\{x,y,z\})\big)}{\big(u(x)+u(y)+u(z)+u(\{x,y,z\})\big)^3} > 0$ and $u(\{x, y, z\}) > -\big(u(y) + u(z)\big)$ imply $u(\{x, y, z\}) > -\big(u(x) + u(z)\big)$. Lastly, $p(x_0, \{x, y, z\}) = \frac{\big(u(x)+u(y)+u(\{x,y,z\})\big)\big(u(x)+u(z)+u(\{x,y,z\})\big)\big(u(y)+u(z)+u(\{x,y,z\})\big)}{\big(u(x)+u(y)+u(z)+u(\{x,y,z\})\big)^3} \geq 0$ implies that $u(\{x, y, z\}) \geq -\big(u(x) + u(y)\big)$. When $u(x) = u(y) = u(z) = a$, it is obvious that $u(\{x, y, z\}) > -\frac{B-\sqrt{B^2-AC}}{A} = -2a$. Now it is enough to prove that

$$-\min\big(u(x) + u(z);\, u(x) + u(z);\, u(x) + u(z)\big) \geq -\frac{B - \sqrt{B^2 - AC}}{A}.$$

Since the inequality is completely symmetric, WLOG, let assume $u(z) \geq u(y) \geq u(x)$. Now we shall prove that $\frac{B-\sqrt{B^2-AC}}{A} \geq u(x) + u(y)$.

$B - \sqrt{B^2 - AC} \geq A\big(u(x) + u(y)\big)$ if and only if $u^2(z) - u(x)u(y) \geq \sqrt{B^2 - AC}$ if and only if

$u^4(z) - 2u^2(z)u(x)u(y) + u^2(x)u^2(y) \geq u^2(x)u^2(y) + u^2(y)u^2(z) + u^2(x)u^2(z) - u(x)u(y)u(z)\big(u(x) + u(y) + u(z)\big)$;

equivalently, $u^4(z) + u^2(x)u(y)u(z) + u^2(y)u(x)u(z) \geq u^2(x)u^2(z) + u^2(y)u^2(z) + u^2(z)u(x)u(y)$

if and only if $u(z)\big(u(z) - u(y)\big)\big(u(z) - u(x)\big)\big(u(x) + u(y) + u(z)\big) \geq 0.$

■

A.4 Proof of Proposition 7

Proof of Proposition 7: By direct calculations, we obtain

$$\frac{\rho(x,\{x,y\})}{\rho(y,\{x,y\})}\Big/\frac{\rho(x,\{x,y,z\})}{\rho(y,\{x,y,z\})}=1+\frac{u(x)\big(u(z)-u(\{x,y\})+u(\{x,y,z\})\big)}{\big(u(y)+u(\{x,y\})\big)\big(u(x)+u(y)+u(z)+u(\{x,y,z\})\big)} \text{ and}$$

$$\frac{\rho^*(x,\{x,y\})}{\rho^*(y,\{x,y\})}\Big/\frac{\rho^*(x,\{x,y,z\})}{\rho^*(y,\{x,y,z\})}=1+\frac{u(x)\big(u(z)-u^*(\{x,y\})+u^*(\{x,y,z\})\big)}{\big(u(y)+u^*(\{x,y\})\big)\big(u(x)+u(y)+u(z)+u^*(\{x,y,z\})\big)}.$$

By Regularity of ρ and ρ^*, $u(z)-u(\{x,y\})+u(\{x,y,z\})>0$ and $u(z)-u^*(\{x,y\})+u^*(\{x,y,z\})>0$ respectively. Therefore, $\frac{\rho^*(x,\{x,y\})}{\rho^*(y,\{x,y\})}/\frac{\rho^*(x,\{x,y,z\})}{\rho^*(y,\{x,y,z\})} > \frac{\rho(x,\{x,y\})}{\rho(y,\{x,y\})}/\frac{\rho(x,\{x,y,z\})}{p(\rho,\{x,y,z\})}$ if and only if

$$\frac{u(z)-u^*(\{x,y\})+u^*(\{x,y,z\})}{u(z)-u(\{x,y\})+u(\{x,y,z\})}>\frac{\big(u(y)+u^*(\{x,y\})\big)\big(u(x)+u(y)+u(z)+u^*(\{x,y,z\})\big)}{\big(u(y)+u(\{x,y\})\big)\big(u(x)+u(y)+u(z)+u(\{x,y,z\})\big)}.$$

Since $u(\{x,y\})-u(\{x,y,z\})\geq u^*(\{x,y\})-u^*(\{x,y,z\})$, we obtain $\frac{u(z)-u^*(\{x,y\})+u^*(\{x,y,z\})}{u(z)-u(\{x,y\})+u(\{x,y,z\})}\geq 1$. Since $u(\{x,y\})>u^*(\{x,y\})$ and $u(\{x,y,z\})>u^*(\{x,y,z\})$, we obtain

$$1>\frac{\big(u(y)+u^*(\{x,y\})\big)\big(u(x)+u(y)+u(z)+u^*(\{x,y,z\})\big)}{\big(u(y)+u(\{x,y\})\big)\big(u(x)+u(y)+u(z)+u(\{x,y,z\})\big)}.$$

■

A.5 Proof of Proposition 9

Proof of Proposition 9: By calculation, we obtain $\rho(a,\{a,b\})=q(a,\{a,b\})=\frac{u(a)}{u(a)+u(b)+u(\{a,b\})}$, $\rho(b,\{a,b\})=q(b,\{a,b\})(1-q(a,\{a,b\}))=\frac{u(b)(u(b)+u(\{a,b\}))}{(u(a)+u(b)+u(\{a,b\}))^2}$, and $\rho(x_0,\{a,b\})=1-\rho(a,\{a,b\})-\rho(b,\{a,b\})=\frac{(u(a)+u(\{a,b\}))(u(b)+u(\{a,b\}))}{(u(a)+u(b)+u(\{a,b\}))^2}$.

First, $\rho(a,\{a,b\})>\rho(x_0,\{a,b\})$ if and only if $\frac{u(a)}{u(a)+u(b)+u(\{a,b\})}>\frac{(u(a)+u(\{a,b\}))(u(b)+u(\{a,b\}))}{(u(a)+u(b)+u(\{a,b\}))^2}$ if and only if $u(a)(u(a)+u(b)+u(\{a,b\}))>(u(a)+u(\{a,b\}))(u(b)+u(\{a,b\}))$. Since $u(a)+u(b)+u(\{a,b\})>u(b)+u(\{a,b\})$, we obtain $(u(a)+u(\{a,b\}))(u(b)+u(\{a,b\}))\geq(u(a)+u(\{a,b\}))(u(a)+u(b)+u(\{a,b\}))$. Therefore, $\rho(a,\{a,b\})>$

$\rho(x_0, \{a,b\})$. Second, $\rho(x_0, \{a,b\}) \geq \rho(b, \{a,b\})$ if and only if $\frac{(u(a)+u(\{a,b\}))(u(b)+u(\{a,b\}))}{(u(a)+u(b)+u(\{a,b\}))^2} \geq \frac{u(b)(u(b)+u(\{a,b\}))}{(u(a)+u(b)+u(\{a,b\}))^2}$ if and only if $u(a)+u \geq u(b)$. Therefore, $\rho(x_0, \{a,b\}) \geq \rho(b, \{a,b\}) = \min(\rho(a, \{a,b\}), \rho(b, \{a,b\}))$. ■

A.6 Omitted Proof of Proposition 4

1. $a \sim b \sim c$: $\rho(a, \{a,b\}) \geq \rho(b, \{a,b\})$ implies $u(a) \geq u(b)$ and $\rho(b, \{b,c\}) \geq \rho(c, \{b,c\})$ implies $u(b) \geq u(c)$. Therefore, we obtain $u(a) \geq u(c)$ which implies $\rho(a, \{a,c\}) \geq \rho(c, \{a,c\})$.

2. $a \sim b \succ c$: $\rho(a, \{a,b\}) \geq \rho(b, \{a,b\})$ implies $u(a) \geq u(b)$ and $\rho(b, \{b,c\}) \geq \rho(c, \{b,c\})$ implies $u(b) \geq f(u(c))$. Therefore, we obtain $u(a) \geq f(u(c))$ which implies $\rho(a, \{a,c\}) \geq \rho(c, \{a,c\})$.

3. $c \succ a \sim b$: $\rho(a, \{a,b\}) \geq \rho(b, \{a,b\})$ implies $u(a) \geq u(b)$ and $\rho(b, bc) \geq \rho(c, bc)$ implies $u(b) \geq g(u(c))$. Therefore, we obtain $u(a) \geq g(u(c))$ which implies $\rho(a, \{a,c\}) \geq \rho(c, \{a,c\})$.

4. $a \succ b \sim c$: $\rho(a, \{a,b\}) \geq \rho(b, \{a,b\})$ implies $u(a) \geq f(u(b))$ and $\rho(b, \{b,c\}) \geq \rho(c, \{b,c\})$ implies $u(b) \geq u(c)$. Therefore, we obtain $u(a) \geq f(u(b)) \geq f(u(c))$ which implies $\rho(a, \{a,c\}) \geq \rho(c, \{a,c\})$.

5. $b \sim c \succ a$: $\rho(a, \{a,b\}) \geq \rho(b, \{a,b\})$ implies $u(a) \geq g(u(b))$ and $\rho(b, \{b,c\}) \geq \rho(c, \{b,c\})$ implies $u(b) \geq u(c)$. Therefore, we obtain $u(a) \geq g(u(b)) \geq g(u(c))$ which implies $\rho(a, \{a,c\}) \geq \rho(c, \{a,c\})$.

6. $a \sim c \succ b$: $\rho(a, \{a,b\}) \geq \rho(b, \{a,b\})$ implies $u(a) \geq f(u(b))$ and $\rho(b, \{b,c\}) \geq \rho(c, \{b,c\})$ implies $f(u(b)) \geq u(c)$. Therefore, we obtain $u(a) \geq u(c)$ which implies $\rho(a, \{a,c\}) \geq \rho(c, \{a,c\})$.

7. $b \succ a \sim c$: $\rho(a, \{a,b\}) \geq \rho(b, \{a,b\})$ implies $u(a) \geq g(u(b))$ and $\rho(b, \{b,c\}) \geq \rho(c, \{b,c\})$ implies $g(u(b)) \geq u(c)$. Therefore, we obtain $u(a) \geq u(c)$ which implies $\rho(a, \{a,c\}) \geq \rho(c, \{a,c\})$.

8. $a \succ c \succ b$: $\rho(a, \{a, b\}) \geq \rho(b, \{a, b\})$ implies $u(a) \geq f(u(b))$ and $\rho(b, \{b, c\}) \geq \rho(c, \{b, c\})$ implies $f(u(b)) \geq u(c)$. Therefore, we obtain $u(a) \geq u(c) > f(u(c))$ which implies $\rho(a, \{a, c\}) > \rho(c, \{a, c\})$.

9. $b \succ a \succ c$: $\rho(a, \{a, b\}) \geq \rho(b, \{a, b\})$ implies $u(a) \geq g(u(b))$ and $\rho(b, \{b, c\}) \geq \rho(c, \{b, c\})$ implies $g(u(b)) \geq u(c)$. Therefore, we obtain $u(a) \geq u(c) > f(u(c))$ which implies $\rho(a, \{a, c\}) > \rho(c, \{a, c\})$.

References

Barberá, S. and P. K. Pattanaik (1986): "Falmagne and the Rationalizability of Stochastic Choices in Terms of Random Orderings," *Econometrica*, 54, pp. 707–715.

Day, B. and G. Loomes (2010): "Conflicting violations of transitivity and where they may lead us," *Theory and decision*, 68, 233–242.

De Clippel, G. and K. Eliaz (2012): "Reason-based choice: A bargaining rationale for the attraction and compromise effects," *Theoretical Economics*, 7, 125–162.

Debreu, G. (1960): "Review of RD Luce, Individual choice behavior: A theoretical analysis," *American Economic Review*, 50, 186–188.

Dhar, R. and I. Simonson (2003): "The effect of forced choice on choice," *Journal of Marketing Research*, 146–160.

Doyle, J. R., D. J. O'Connor, G. M. Reynolds, and P. A. Bottomley (1999): "The robustness of the asymmetrically dominated effect: Buying frames, phantom alternatives, and in-store purchases," *Psychology & Marketing*, 16, 225–243.

Falmagne, J. (1978): "A representation theorem for finite random scale systems," *Journal of Mathematical Psychology*, 18, 52–72.

Fudenberg, D., R. Iijima, and T. Strzalecki (2013): "Stochastic Choice and Revealed Perturbed Utility," Mimeo, Harvard University.

Gul, F., P. Natenzon, and W. Pesendorfer (2014): "Random Choice as Behavioral Optimization," Forthcoming, Econometrica.

HEATH, T. B. AND S. CHATTERJEE (1995): "Asymmetric decoy effects on lower-quality versus higher-quality brands: Meta-analytic and experimental evidence," *Journal of Consumer Research*, 268–284.

KRAJBICH, I. AND A. RANGEL (2011): "Multialternative drift-diffusion model predicts the relationship between visual fixations and choice in value-based decisions," *Proceedings of the National Academy of Sciences*, 108, 13852–13857.

LLERAS, J. S., Y. MASATLIOGLU, D. NAKAJIMA, AND E. Y. OZBAY (2010): "When More is Less: Choice with Limited Consideration," Mimeo, U. of Michigan.

LOOMES, G., C. STARMER, AND R. SUGDEN (1991): "Observing violations of transitivity by experimental methods," *Econometrica: Journal of the Econometric Society*, 425–439.

LUCE, R. D. (1959): *Individual Choice Behavior a Theoretical Analysis*, John Wiley and sons.

MACHINA, M. J. (1985): "Stochastic choice functions generated from deterministic preferences over lotteries," *The economic journal*, 575–594.

MANZINI, P. AND M. MARIOTTI (2014): "Stochastic Choice and Consideration Sets," *Econometrica*, 82, 1153–1176.

MASATLIOGLU, Y., D. NAKAJIMA, AND E. Y. OZBAY (2012): "Revealed attention," *The American Economic Review*, 102, 2183–2205.

MCFADDEN, D. AND M. RICHTER (1990): "Stochastic rationality and revealed stochastic preference," in *Preferences, Uncertainty, and Optimality, Essays in Honor of Leo Hurwicz, Westview Press: Boulder, CO*, 161–186.

MILOSAVLJEVIC, M., V. NAVALPAKKAM, C. KOCH, AND A. RANGEL (2012): "Relative visual saliency differences induce sizable bias in consumer choice," *Journal of Consumer Psychology*, 22, 67–74.

MOCHON, D. AND S. FREDERICK (2012): "Compromise or response order effects?" Mimeo, Yale University School of Management.

Natenzon, P. (2010): "Random choice and learning," Working paper, Princeton University.

Ratneshwar, S., A. D. Shocker, and D. W. Stewart (1987): "Toward understanding the attraction effect: The implications of product stimulus meaningfulness and familiarity," *Journal of Consumer Research*, 520–533.

Reutskaja, E., R. Nagel, C. F. Camerer, and A. Rangel (2011): "Search Dynamics in Consumer Choice under Time Pressure: An Eye-Tracking Study," *The American Economic Review*, 101, 900–926.

Rieskamp, J., J. R. Busemeyer, and B. A. Mellers (2006): "Extending the bounds of rationality: evidence and theories of preferential choice," *Journal of Economic Literature*, 631–661.

Roe, R. M., J. R. Busemeyer, J. T. Townsend, et al. (2001): "Multialternative decision field theory: A dynamic connectionist model of decision making," *PSYCHOLOGICAL REVIEW-NEW YORK-*, 108, 370–392.

Saito, K. (2014): "Preference for Flexibility and Preference for Randomization under Ambiguity," *Forthcoming, The American Economic Review.*

Shaw, M. L. and P. Shaw (1977): "Optimal allocation of cognitive resources to spatial locations." *Journal of Experimental Psychology: Human Perception and Performance*, 3, 201.

Sheng, S., A. M. Parker, and K. Nakamoto (2005): "Understanding the mechanism and determinants of compromise effects," *Psychology & Marketing*, 22, 591–609.

Simonson, I. and A. Tversky (1992): "Choice in context: tradeoff contrast and extremeness aversion." *Journal of marketing research.*

Towal, R. B., M. Mormann, and C. Koch (2013): "Simultaneous modeling of visual saliency and value computation improves predictions of economic choice," *Proceedings of the National Academy of Sciences*, 110, E3858–E3867.

Tserenjigmid, G. (2013): "The Order Dependent Luce Model," Mimeo, Caltech.

Tversky, A. (1969): "Intransitivity of preferences." *Psychological Review; Psychological Review*, 76, 31.

——— (1972a): "Choice by elimination," *Journal of mathematical psychology*, 9, 341–367.

——— (1972b): "Elimination by aspects: A theory of choice." *Psychological review*, 79, 281.

Usher, M. and J. L. McClelland (2004): "Loss aversion and inhibition in dynamical models of multialternative choice." *Psychological review*, 111, 757.