Forthcoming, Mathematical Social Sciences.

The Perception-Adjusted Luce Model

Federico Echenique*[1], Kota Saito†[1], and Gerelt Tserenjigmid‡[2]

[1]Division of the Humanities and Social Sciences, Caltech

[2]Department of Economics, Virginia Tech

October 15, 2017

Abstract

We develop an axiomatic theory of random choice that builds on Luce's (1959) model to incorporate a role for perception. We capture the role of perception through *perception priorities;* priorities that determine whether an object or alternative is perceived sooner or later than other alternatives. We identify agents' perception priorities from their violations of Luce's axiom of *independence from irrelevant alternatives* (IIA). The direction of the violation of IIA implies an orientation of agents' priority rankings. We adjust choice probabilities to account for the effects of perception, and impose that adjusted choice probabilities satisfy IIA. So all violations of IIA are accounted for by the perception order. The theory can explain some very well-documented behavioral phenomena in individual choice.

*1200 E California Blvd, MC 228-77 Pasadena, CA 91125, USA. E-mail: fede@hss.caltech.edu
Echenique thanks the National Science Foundation for its support through the grants SES 1558757 and CNS-1518941.

†Corresponding Author. 1200 E California Blvd, MC 228-77 Pasadena, CA 91125, USA. E-mail: saito@caltech.edu Saito thanks the National Science Foundation for its support through the grant SES 1558757.

‡3016 Pamplin Hall, 880 West Campus Drive, Blacksburg, VA 24061, USA. E-mail: gerelt@vt.edu

JEL Classification: D01.

Keywords: random choice, Luce's model, IIA, perception priorities, regularity.

1 Introduction

We study the role of perception in individual stochastic choice. Perception is captured through *priority orders,* which determine whether an alternative, or object of choice, is perceived sooner or later than other alternatives. The perception priority order could represent differences in familiarity, or salience, of the objects of choice.

Our main contribution 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. We attribute any violation of IIA to the role of perception, and use these violations to back out a perception order. Our model, a *perception-adjusted Luce model (PALM)*, reduces to Luce's when perception plays no role.

In PALM, an agent makes choices as if she were following a sequential procedure. In the procedure, the agent considers different alternatives in sequence, following a perception priority order. The probability of choosing an alternative depends on the probability of not choosing an alternative that has been perceived before. The probability of choosing an alternative also depends on relative utility, just as in Luce's model. If none of the alternatives is chosen, then the outside option will be chosen. The sequential nature of PALM allows us to explain violations of stochastic transitivity and regularity and choice overload (see Section 5).

We use stochastic choice data to construct a perception priority order. We start from a primitive stochastic choice, and when the choice satisfies certain axioms, we can construct a PALM model. The perception priority order comes from the observed violations of Luce's IIA. Luce's IIA says that the relative choice probabilities of alternative a over b should not be affected by adding a third alternative c. So suppose that we have a violation of IIA, and that adding c changes the probability of choosing a relative to that of choosing b. What can we conclude about perception? We claim that a *decrease* in the relative probability of choosing a over b is an indication that a has *higher* perception priority than b.

The reasoning is as follows. Suppose a has the highest perception priority and c has the lowest perception priority among a, b, and c. Adding c has the following two effects. First, since c directly competes with a and b, as in Luce's model c proportionately decreases the probabilities of choosing a and b if we ignore the effect of perception (Luce effect). Second, when a has a higher perception priority than b, then the very fact that a is chosen with lower probability means that b has a higher chance of being perceived. So there is a second effect of adding c, and it favors choosing b (a perception effect). Taking both effects into account means that adding c provokes a *larger decrease* in the probability of choosing a than in the probability of choosing b. This means that the resulting violation of Luce's IIA takes the form of a decrease in the relative probability of choosing a over b.[1]

The second idea in our construction is to use the perception priority order to define a *hazard rate*. The hazard rate is 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. We impose two axioms. The first requires that the perception priority be complete and transitive. The second axiom is imposed on hazard rates, and says that hazard rates must satisfy the IIA. Since hazard rates are obtained from choices by accounting for priority, and hazard rates equal the primitive choice probability where priority does not matter, our axiom means that perception explains all the deviations from IIA.

The resulting model of choice is what we call PALM, the perception-adjusted Luce 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 its utility (as in

[1]Our strategy for identifying priority from violations of IIA is why the resulting order can be called a perception priority. It follows from the role of perception in our theory of choice.

Luce) and on its priority in perception.

Despite having a tight axiomatic characterization, PALM is quite flexible and can accommodate many behavioral phenomena: Section 5 has the details. In particular, PALM can violate the regularity axiom and stochastic transitivity. We also use PALM to explain choice overload. An increase in the number of objects can lead to an increased probability of not making a choice (i.e., an increased probability of choosing the outside option), when the objects are similar to each other.

It is instructive to see how PALM can accommodate violations of regularity. Doyle et al. (1999) is a representative experiment with evidence in favor of the attraction effect, a well-known violation of regularity. Doyle et al. present customers with a choice of baked beans. The first choice is between two types of baked beans: a and b; a is Heinz baked beans, while b is a local cheap brand called Spar. In the experiment, b was chosen 19% of the time. The authors then introduced a third option, c, a more expensive version of the local brand Spar. After c was introduced, b was chosen 33% of the time. This pattern (i.e., an increase in the probability of choosing b) of choices 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 brand of beans. Since a is the well-known Heinz brand, it is likely to be the highest priority alternative. Also, b and c have the same perception priority because they are the same brands. Given this perception priority, if the utility of a is large enough, PALM produces the attraction effect in Doyle et al.'s experiment. As we explained above, the addition of c in principle hurts the choice probabilities of both a and b. However, while a does not benefit from b's potential decrease, b does benefit from the decrease in the probability of choosing a because b has lower priority than a (a perception effect). The magnitude of this positive effect depends on the utility of a; if the utility of a 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 b. This increase in the probability of choosing b is a violation of regularity.

2 Primitives and Luce's model

Let X be a countable and nonempty set of *alternatives*, and $\mathscr{A}$ be a set of finite and nonempty subsets of X. Suppose that $\mathscr{A}$ includes all sets with two and three elements. We model an agent who makes a probabilistic choice from $A_0 \equiv A \cup \{x_0\}$, 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 be interpreted as the agent not making a choice. Let $X_0 \equiv X \cup \{x_0\}$.

Definition: A function $\rho : X_0 \times \mathscr{A} \to [0,1]$ is called a *stochastic choice function* if

$$\sum_{a \in A_0} \rho(a, A) = 1$$

for all $A \in \mathscr{A}$. A stochastic choice function ρ is *nondegenerate* if $\rho(a, A) \in (0,1)$ for all $A \in \mathscr{A}$ and $a \in A_0$.

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

Definition: A stochastic choice function ρ *satisfies Luce's independence of irrelevant alternatives (IIA) axiom at* $a, b \in X_0$ if, for any $A \in \mathscr{A}$ with $a, b \in A_0$,

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

Moreover, ρ *satisfies IIA* if ρ satisfies IIA at a, b for all $a, b \in X_0$.

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 the (extended) Luce's model if there exists $u : X_0 \to \mathbf{R}_{++}$ such that for any $A \in \mathscr{A}$ and $a \in A_0$,

$$\rho(a, A) = \frac{u(a)}{\sum_{a' \in A} u(a') + u(x_0)}. \tag{1}$$

Luce presented his model with no outside option. Here we allow for an outside option, and use the version of Luce's model in which not choosing in A is possible.

Luce's model satisfies a monotonicity property: $\rho(x, A) \geq \rho(x, B)$, if $A \subset B$. This property is called *regularity*.

2.1 PALM

Perception priority. 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 before b, and when $a \sim b$, then a and b are perceived simultaneously. We should mention that other interpretations of $\succsim$ are possible. One can, for example, think that all alternatives are perceived simultaneously, but considered in order.

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 $u : X_0 \to \mathbf{R}_{++}$ such that for any $A \in \mathscr{A}$ and $a \in A$,

$$\rho(a, A) = \mu(a, A) \prod_{\alpha \in A/\succsim:\, \alpha \succ a} \big(1 - \mu(\alpha, A)\big), \tag{2}$$

where

$$\mu(a, A) = \frac{u(a)}{\sum_{b \in A} u(b) + u(x_0)}.$$

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) for any $x, y \in A$, $x \sim y$ 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 $x \in \alpha$. Luce's model is a special case of PALM, in which $a \sim b$ for all $a, b \in X$.

For any PALM $(u, \succsim)$, we denote by $\rho_{(u,\succsim)}$ the stochastic choice defined through (2).

(When there is no risk of confusion, we write ρ instead of $\rho_{(u,\succsim)}$.)

The PALM has a procedural interpretation. First consider the highest-priority alternatives in A, and choose each of them with probability given by $\mu(\cdot, A)$; these probabilities obey a Luce formula. This means that if α is the set of highest-priority elements of A, then each $a \in \alpha$ is chosen with probability $\mu(a, A)$. With probability $1 - \mu(\alpha, A)$ none of the elements in α is chosen. If none of the elements of α are chosen, then move on to the second-highest priority alternatives, and choose each of them with the Luce probability specified by μ. 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 x with "Luce probability" $\mu(x, A)$. With probability $1 - \mu(x, A)$, x is not chosen, and the agents moves on to consider y, the second-highest priority element. She chooses y with probability $\mu(y, A)$. This means that the probability of choosing y is $\mu(y, A)\big(1 - \mu(x, A)\big)$. Finally, the probability of choosing z is equal to $\mu(z, A)\big(1 - \mu(x, A)\big)\big(1 - \mu(y, A)\big)$. 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 $\mu(z, A)\big(1 - \mu(x, A) - \mu(y, A)\big)$. The idea captured by $x \sim y$ is that x and y are perceived, and considered, simultaneously. So the probability of choosing an option that has higher priority than z is $\mu(x, A) + \mu(y, A)$.

Note that $x \succ y$ means that the agent considers x before y, but it does not mean that the agent completely ignores y. Therefore, x is chosen with probability $\mu(x, A)$ instead of $\mu(x, \{x\})$ since she is aware of all x, y, z. We are certainly not modelling agents who miss or ignore y and z even when there are only three alternatives.

3 Axioms

We introduce the revealed perception priority order derived from ρ, and the resulting hazard rate function. The hazard rate function will be a "perception adjusted" random choice function. It coincides with the random choice function except where violations of Luce's IIA are present. When there are violations of Luce's IIA, they will be attributed to the role of perception. So in our model the hazard rate will satisfy IIA, even when the primitive stochastic choice violates IIA.

Revealed perception priority. We denote by $\succsim^*$ the revealed priority relation that we obtain from the data in ρ. To define $\succsim^*$, first we identify the direct revealed priority relation $\succsim^0$ from ρ. The revealed priority relation $\succsim^*$ 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, meaning that there is some third alternative whose presence affects the relative probability of choosing a over 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 of the two orderings, $a \succ^0 b$ 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

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

In words, the presence of c lowers the probability of choosing a relative to the probability of choosing b. When does adding an option hurt one alternative relatively more than another? We claim that this happens when a has *higher priority* than b. The reason is that by adding 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 the high priority a over the low priority b.

As we explained in the introduction, we seek to model perception through an order in which alternatives are considered. Adding c to $\{a,b\}$ would in principle decrease the probability of choosing both a and b because c competes with a and b; but when a has higher priority than b, then the sole fact that a's choice probability decreases implies that choosing b becomes more likely. The reason is that 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, by adding c we may also be decreasing the probability of choosing b because c and b are in competition, so the *net* effect on the probability of choosing b is not determined. However, we do know that $\frac{\rho(a,\{a,b\})}{\rho(b,\{a,b\})} > \frac{\rho(a,\{a,b,c\})}{\rho(b,\{a,b,c\})}$. 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.

It is important to note that

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

does not always imply that $a \succ b$. It will imply that $a \succ b$ only when c has either more or less priority than both a and b. When c is inbetween, then its presence may also disproportionally hurt b, as it has higher priority than b.[2]

[2]To illustrate, consider the case $a \succeq c \succ b$. As we explained before, adding c to $\{a,b\}$ has negative effects on the choice probabilities of both a and b because c competes with a and b. It also

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

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

where $A_a = \{b \in A | b \succ^* a\}$. For the outside option, we also define $q(x_0, A) = 1 - \sum_{a \in A} q(a, A)$. Here q is called ρ's *hazard rate function.*

It is important to note that the Hazard Rate function q can be defined independent of the Weak Order axiom. In fact, In Appendix A.4, we show that Weak Order and the next axiom on q, *Hazard Rate IIA*, are independent.

We ascribe all violations of IIA to the role of perception, and the hazard rate is the tool that we use to that purpose.

Axiom (Hazard Rate IIA) The hazard rate function q satisfies Luce's IIA; that is, for any $a, b \in X_0$, and $A \in \mathscr{A}$ with $a, b \in A_0$,

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

---

has a positive effect on b because b will be chosen only after a is not chosen, and a is not chosen with higher probability after we add c. However, when $a \succeq c \succ b$, then c also directly hurts b (but not a) because b will be chosen only after c is not chosen. Therefore, when $a \succeq c \succ b$, we *can* have

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

In particular, when either $a \sim c \succ b$ or $a \succ c \succ b$ and the utility of c is large enough, we will have

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

Observe that the definition of $\succ^0$ involves $c \in X$ such that $c \nsim^0 a$ and $c \nsim^0 b$. The subtlety in the definition of $\succ^0$ is to rule out the case $a \sim c \succ b$.

The idea behind Hazard Rate IIA is that all violations of Luce's IIA are explained by the perception priority order. The definition of q implies that

$$\frac{q(a,A)}{q(b,A)} = \frac{\rho(a,A)}{\rho(b,A)} \frac{\rho(\overline{A}_b, A)}{\rho(\overline{A}_a, A)}. \tag{4}$$

(Where $\overline{A}_a = A_0 \setminus A_a$ and $\overline{A}_b = A_0 \setminus A_b$.) If Luce's IIA is violated, we must have a change in the "relative probability" of choosing a over b: $\frac{\rho(a,A)}{\rho(b,A)} \neq \frac{\rho(a,\{a,b\})}{\rho(b,\{a,b\})}$. Hazard Rate IIA implies that the "relative hazard rate" stays the same, $\frac{q(a,\{a,b\})}{q(b,\{a,b\})} = \frac{q(a,A)}{q(b,A)}$. This means that the far-right term of (4), $\frac{\rho(\overline{A}_b,A)}{\rho(\overline{A}_a,A)}$, must change as well.

Now, if $a \sim^* b$ then $A_a = A_b$, and Hazard Rate IIA implies the Luce IIA formula for a and b. *Therefore Hazard Rate IIA only differs from Luce's IIA for alternatives that are strictly ordered by perception priority.*

So suppose that a has higher priority than b, and that the relative probability of choosing a over b is smaller when the choice set is $A \cup \{c\}$ than when the choice set is A. Hazard Rate IIA means that the perception priority explains the change in relative probabilities: we must have a compensating decrease in the probability of choosing an element that is perceived before b, relative to the probability of choosing an element that is perceived before a. The explanation is that a was "hurt" relative to b because the choice of a or b depends in part on the probability of choosing an element with higher perception priority, and the addition of c decreased the relative probability of choosing an element with higher priority than b.

In other words, the relative probability of choosing a over b decreased, and therefore Luce's IIA was violated, because the probability of choosing an element that is perceived before b increased relative to the probability of choosing an element that is perceived before a. Hazard Rate IIA means that the only permissible violations of Luce's IIA are those that can be explained in this fashion by the perception priority order.

4 Theorem

Before stating the theorem, we also define an additional technical condition called "richness". Richness requires that X has infinitely many alternatives. We do not need this condition to prove the sufficiency of the axioms: that the axioms imply a PALM representation. We need it to prove the necessity of the axioms, in particular, the result that $\succsim = \succsim^*$.

Richness: Let $X = \{a_i\}_{i \in \mathbb{Z}}$. For any $i \in \mathbb{Z}$, $a_i \succeq a_{i+1}$, and for any pair $a_i, a_j \in X$ with $a_i \succ a_j$, there is $a_k \in X$ with $a_k \succ a_i$ or $a_j \succ a_k$.[3]

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

Conversely, for a given PALM $(u, \succsim)$, if $\succsim$ satisfies Richness, then $\rho_{(u,\succsim)}$ satisfies Weak Order and Hazard Rate IIA, 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 thus identified from 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 the Outside Option

It is useful to compare how Luce and PALM treat the outside option, the probability of not making a choice from a set A.

For PALM, the utility of the outside option is:

$$u(x_0) = \sum_{a \in A} u(a) \left(\frac{1}{\sum_{a \in A} q(a, A)} - 1 \right). \tag{5}$$

[3]We can prove Theorem 1 when X is finite by slightly modifying the revealed perception priority order $\succeq^*$. See Appendix A.2.

In (extended) Luce's model, the utility of the outside option has a similar expression. Indeed,

$$\hat{u}(x_0) = \sum_{a\in A} u(a) \left(\frac{1}{\sum_{a\in A} \rho(a, A)} - 1 \right), \tag{6}$$

with ρ in place of the hazard rates q.

It is interesting to contrast the value of $u(x_0)$ according to Equation (5) with what one would obtain from Equation (6). Given a PALM model $(u, \succsim)$, we can calculate $\hat{u}(x_0)$ from $\rho_{(u,\succsim)}$ by application of Equation (6). If we do that, we obtain

1. $\hat{u}(x_0) \geq u(x_0)$,

2. and $\hat{u}(x_0) = u(x_0)$ when $a \sim b$ for all $a, b \in A$.

The inequality $\hat{u}(x_0) \geq u(x_0)$ reflects that there are two sources behind choosing the outside option in PALM. One source is the utility $u(x_0)$ 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 an agent that chooses sequentially, following the priority order $\succsim$, then it is possible that we exhaust the elements 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 $\hat{u}(x_0) \geq u(x_0)$. For example, when the utility of the outside option is zero, the outside option will not be chosen in Luce's model. However, in PALM, the outside option will be chosen with positive probability because of the second source behind choosing the outside option.

5 Behaviors Consistent with PALM

5.1 Choice Overload

The outside option in PALM allows us to capture various behavioral phenomena. One example is "choice overload:" the idea that a subject may be inclined to make *no choice* when presented with many alternatives. The paper by Iyengar and Lepper (2000) is a well known study of choice overload. Iyengar and Lepper run an experiment where subjects had to choose among a large set of nearly identical alternatives. They find that a large fraction of subjects make no choice whatsoever, and that the fraction of subjects who make no choice increases from 26% to 40% as the number of alternatives increases. These results are easily captured by PALM.

Let $A = \{a_1, \ldots, a_n\}$ be a menu with n elements, each of which provide the same Luce utility; so $u(a_1) = u(a_2) = \ldots = u(a_n) > 0$. Suppose that the n elements in A are strictly ordered by the perception priority $\succsim$, and that $u(x_0) = 0$. Then the probability of choosing the outside option is

$$\rho(x_0, A) = (1 - 1/n)^n,$$

which is monotone increasing in n. In other words, the probability of not making a choice in A increases as the cardinality of A increases. Moreover, $\rho(x_0, A)$ goes from about 25% to $\frac{1}{e} \approx 37\%$ as n increases, consistent with the numbers 26% and 40% in the Iyengar and Lepper experiment.

5.2 Violation of Regularity

Regularity is often assumed in the literature of random choice. However, violations of regularity are widely documented in the literature. For example, as we discussed in the introduction using the experimental result of Doyle et al. (1999), the attraction effect is a well-known violation of regularity.[4]

The following proposition shows that when a third alternative z is added to a menu $\{x, y\}$ we can obtain an increase in the probability of choosing y over x when the utility of x is large enough.

[4]The attraction effect is first documented in Huber et al. (1982) and later confirmed by many studies such as Simonson (1989), Simonson and Tversky (1992), and Herne (1997).

Proposition 1: *If $x \succ y \succsim z$ and $u(x)$ is large enough, then*

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

Proof of Proposition 1: We have

$$\begin{aligned} \rho(y, \{x, y, z\}) > \rho(y, \{x, y\}) &\Leftrightarrow q(y, \{x, y, z\})(1 - q(x, \{x, y, z\})) > q(y, \{x, y\})(1 - q(x, \{x, y\})) \\ &\Leftrightarrow u(x) > \sqrt{(u(y) + u(z) + u(x_0))(u(y) + u(x_0))} \end{aligned}$$

■

More generally, we can show that $\rho(y, A \cup \{z\}) > \rho(y, A)$ holds as long as $y \succsim z$ and the sum of the utilities of alternatives in A that have higher perception priorities than y is large enough.[5]

5.3 Comparing High Perception Priority with High Utility and Violations of Stochastic Transitivity

All other aspects being equal, an increase in the utility or an increase in the perception priority of a given alternative will lead to an increase in the probability of choosing that alternative. In this subsection, we discuss which increase will lead to a higher increase in the probability of choosing that alternative.

To illustrate, assume a menu $A = \{a_1, \ldots, a_n\}$ with $a_1 \succ a_2 \succ \ldots \succ a_n$. If the utilities of all the alternatives are the same, then PALM predicts that $\rho(a_i, A) > \rho(a_{i+1}, A)$ for all i. As the utility of $u(a_{i+1})$ becomes larger than $u(a_i)$, the probability $\rho(a_{i+1}, A)$ becomes closer to $\rho(a_i, A)$. In order to study how the effect of the perception priority on random choice can be diminished by utilities, let us calculate the lower bound of the utility of a_{i+1} such that $\rho(a_i, A) = \rho(a_{i+1}, A)$. For simplicity,

[5] Appendix A.3 discusses a modification of PALM which avoids the outside option. In the modification of PALM, whenever the agent chooses no alternative, she repeats the sequential procedure of PALM until she chooses some alternative. Using this modification, we illustrate that the outside option does not really play a role in explaining violations of regularity and stochastic transitivity, but the sequential procedure does.

let $u(x_0) = 0$ and $u(a_j) = u(a_i)$ for all $j \neq i+1$. Then we have

$$\rho(a_i, A) = \mu(a_i, A) \prod_{k<i} (1 - \mu(a_k, A)) = \rho(a_{i+1}, A) = \mu(a_{i+1}, A) \prod_{k<i+1} (1 - \mu(a_k, A));$$

if and only if

$$\frac{u(a_{i+1})}{u(a_i)} = \frac{\sqrt{(n-3)^2 + 4(n-1)} - (n-3)}{2}.$$

For example, when $n = 2$, $\frac{\sqrt{(n-3)^2+4(n-1)}-(n-3)}{2} \approx 1.6$. Therefore, when only a_i and a_{i+1} are available, the utility of a_{i+1} must be about 60% higher than that of a_i to achieve $\rho(a_i, A) = \rho(a_{i+1}, A)$. Similarly, when $n = 3$, $u(a_{i+1})$ must be about 40% higher than $u(a_i)$ to achieve $\rho(a_i, A) = \rho(a_{i+1}, A)$. Moreover, when n is large, since $\frac{\sqrt{(n-3)^2+4(n-1)}-(n-3)}{2} \approx 1 + \frac{1}{n}$, the utility of a_{i+1} must be about $\frac{100}{n}$% higher than that of a_i to achieve $\rho(a_i, A) = \rho(a_{i+1}, A)$. Interestingly, since $(1 + \frac{1}{n})^n \approx e$, $u(x_n)$ must be about $e - 1 \approx 1.71$ times higher than $u(a_1)$ to achieve $p(a_1, A) = p(a_n, A)$.

By the above argument, it is not difficult to see how PALM allows violations of stochastic transitivity.[6] Consider three alternatives x, y, z with $x \succ y \succ z$. Let $1.5\,u(x) = u(y)$ and $1.5\,u(y) = u(z)$ and $u(x_0) = 0$. By the above argument, $\rho(x, \{x, y\}) > \rho(y, \{x, y\})$ since $x \succ y$ and $1.6\,u(x) > u(y)$. Similarly, $\rho(y, \{y, z\}) > \rho(z, \{y, z\})$ since $y \succ z$ and $1.6\,u(y) > u(z)$. However, $\rho(x, \{x, z\}) < \rho(z, \{x, z\})$ since $1.6\,u(x) < u(z) = 2.25\,u(x)$.

6 Related Literature

Section 5 explains how PALM relates to the relevant empirical findings. 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 proposing models that can explain violations of regularity, IIA, and stochastic transitivity. Rieskamp et al. (2006) is an excellent

[6] Violations of stochastic transitivity are well documented in lab experiments. For example, see Tversky (1969) and Loomes et al. (1991).

survey. Examples are Tversky (1972), Roe et al. (2001) and Usher and McClelland (2004). The latter two papers propose *decision field theory*, which allows for violations of regularity. 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 violations of regularity as well as the well-known violations of IIA, the *similarity effect* (Tversky (1972)) and the *compromise effect* (Simonson (1989)).

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. 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).

As we have seen in Section 5, there are instances of PALM that violate the regularity axiom. A random utility model must always satisfy regularity. Thus PALM is not a special case of random utility. Moreover, Luce's model is a random utility model, and a special case of PALM. So the class of PALM and random utility models 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 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). Therefore, there are 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. It turns out that, Manzini and Mariotti's model is also a random utility model. Therefore, they do not allow for violations of regularity.

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 and compromise effects. The source of violations of IIA is not perception, but instead a sort of menu-dependent utility.

5) Fudenberg et al. (2015) considers a decision maker who make a random choice

to maximize expected utility minus some cost. In their model, because of the cost function, the decision maker's choice is random. Versions of their model can accommodate the attraction effect, and the compromise effect. Their model also can describe choice overload.

6) Some related studies use the model of non-stochastic choice to explain some of the experimental results we describe in Section 5. This makes them quite different, as the primitives are different. The paper by Lleras et al. (2010) is important to mention. (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 the choice 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). The main difference is that in their model alternatives are completely ignored if they are not in the decision maker's consideration set. On the other hand, in our model, the decision maker is aware of all alternatives, but she considers (or perceives) them sequentially.

7) Ravid (2015) studies a random choice model of the following sequential procedure. First, an agent picks an option at random from the choice set; the option becomes "focal." Second, she compares the focal option to each other alternative in the set. Third, the agent chooses the focal option if it passes all binary comparisons favorably. Otherwise, the agent draws a new focal option with replacement. Ravid (2015) characterizes the procedures by an relaxation of IIA termed Independence of Shared Alternatives (ISA). His model is also consistent with choice overload and violations of regularity, IIA, and stochastic transitivity.

7 Proof of Theorem 1

7.1 Necessity

We start by proving the converse statement. Let $(u, \succsim)$ be a 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: $c \succ a \succ b$.

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

where the last strict inequality is by Luce's regularity on μ; that is, $\mu(a,\{a,b\}) = \frac{u(a)}{u(a)+u(b)+u(x_0)} > \mu(a,\{a,b,c\}) = \frac{u(a)}{u(a)+u(b)+u(c)+u(x_0)}$.

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

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

where the last strict inequality is by Luce's regularity on μ. ■

First, we prove that $a \sim b$ if and only if $a \sim^* b$. Then, we prove that $a \succ b$ 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 $a \sim^0 b$.

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

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, $b \succ a$. First, by Step 2, we have $a \nsim^0 b$. Second, by Richness, there is c with $c \succ b \succ a$ or $b \succ a \succ c$. Then, 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 $c \not\sim b$ and $c \not\sim a$, Step 2 shows that $c \nsim^0 a$ and $c \nsim^0 b$. Hence, $a \not\succ^0 b$. So we have $a \not\succsim^0 b$. ■

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 Step 1 of 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: By the definition of $\succeq^*$, there exist $c_1, \ldots, c_k \in X$ such that $a \succeq^0 c_1 \succeq^0 \cdots \succeq^0 c_k \succeq^0 b$ (at least one strict relation). Then, by Step 3 of Lemma 2 and Case 1, $a \succeq c_1 \succeq \cdots \succeq c_k \succeq b$ (at least one strict relation). Hence, by transitivity, $a \succ b$.

■

The next lemma shows the converse.

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

Proof: Let $X = \{a_i\}_{i \in \mathbb{Z}}$. By Richness, for any $i \in \mathbb{Z}$, $a_i \succeq a_{i+1}$, and for any pair $a_i, a_j \in X$ with $a_i \succ a_j$, there is $a_k \in X$ such that $a_k \succ a_i$ or $a_j \succ a_k$. We shall prove that for any a_i, a_j, if $a_i \succ a_j$, then $a_i \succ^* a_j$. 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 $a \succ c \succ b$.

Case 1: $a_i \vdash a_j$.

It suffices to show that $a_i \succ^0 a_j$. Take any $a_t \in X$ such that $a_i \nsim^0 a_t$ and $a_j \nsim^0 a_t$ (by Richness such a_t exists). By Lemma 2, $a_i \nsim a_t$ and $a_j \nsim a_t$. Therefore, either $a_t \succ a_i$ or $a_j \succ a_t$. Since $a_i \succ a_j$, then $a_t \succ a_i \succ a_j$ or $a_i \succ a_j \succ a_t$. In either case, by Lemma 1, $\frac{\rho(a_i,\{a_i,a_j,a_t\})}{\rho(a_j,\{a_i,a_j,a_t\})} / \frac{\rho(a_i,\{a_i,a_j\})}{\rho(a_j,\{a_i,a_j\})} < 1$. Thus $a_i \succ^0 a_j$. Hence, $a_i \succ^* a_j$.

Case 2: $a_i \not\vdash a_j$. Since there is only a finite number of alternatives between a_i, a_j, we can find a sequence $a^1, \ldots, a^m \in \{a_{i+1}, \ldots, a_{j-1}\}$ such that $a \vdash a^1 \vdash \ldots \vdash a^m \vdash a_j$. By the argument in Case 1, $a_i \succ^0 a^1 \succ^0 \ldots \succ^0 a^m \succ^0 a_j$. Therefore, $a_i \succ^* a_j$. ■

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.

Step 1: there exists $u : X_0 \to \mathbb{R}_{++}$ such that $q(a, A) = \frac{u(a)}{\sum_{a' \in A} u(a') + u(x_0)}$.

Proof of Step 1: Since q satisfies Luce's IIA and $\sum_{a \in A_0} q(a, A) = 1$, by Luce's theorem (Luce (1959)), there exists $u : X_0 \to \mathbb{R}$ such that $q(a, A) = \frac{u(a)}{\sum_{a' \in A} u(a') + u(x_0)}$. Since ρ is nondegenerate, $u(a) > 0$ for any $a \in X_0$. ■

Step 2: $\rho = \rho_{(u, \succsim^*)}$.

Proof of Step 2: 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 $\{\alpha^1, \alpha^2, \ldots \alpha^k\}$ 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 $s \in \{1, \ldots, k\}$,

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

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. $a \succsim b$ and $a \succsim c$: Since $u(c) > 0$, $\rho(a, \{a, b, c\}) = \mu(a, \{a, b, c\}) < \rho(a, \{a, b\}) = \mu(a, \{a, b\})$.

2. $b \succsim a$ and $b \succsim c$: Since $u(c) > 0$, $\rho(b, \{a, b, c\}) = \mu(b, \{a, b, c\}) < \rho(b, \{a, b\}) = \mu(b, \{a, b\})$.

3. $c \succ a \succsim b$: Since $u(c) > 0$, $\rho(a, \{a, b, c\}) = \mu(a, \{a, b, c\})(1 - \mu(c, \{a, b, c\})) < \mu(a, \{a, b, c\}) < \mu(a, \{a, b\}) = \rho(a, \{a, b\})$.

4. $c \succ b \succ a$: Since $u(c) > 0$, $\rho(b, \{a, b, c\}) = \mu(b, \{a, b, c\})(1 - \mu(c, \{a, b, c\})) < \mu(b, \{a, b, c\}) < \rho(b, \{a, b\}) = \mu(b, \{a, b\})$.

A.2 Finite X

Here we consider the case where X is finite. In order to obtain the necessity part of Theorem 1, we will modify $\succsim^*$ in the following way:

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)

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

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.

(iii) $a \succ^0 b$ if aPb and $a'Pb'$ for any a', b' with $a' \sim^0 a$ and $b' \sim^0 b$. We write $a \succsim^0 b$ if $a \sim^0 b$ or $a \succ^0 b$.

(iv) Define $\succsim^*$ be 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.

Now we can prove Theorem 1 when X is finite. Partition X into equivalence classes for $\succeq$: Let $X = \cup_{i=1}^n \alpha_i$ with $a_i \succ a_{i+1}$ for any $a_i \in \alpha_i$ and $a_{i+1} \in \alpha_{i+1}$.

Richness*: $n \geq 3$, $|\alpha_1 \cup \alpha_n| \geq 3$, and there are $a \in \alpha_1$, $b \in \alpha_n$, and $c \notin \alpha_1 \cup \alpha_n$ such that $u(a) \geq u(b) + u(c) + u(x_0)$; i.e., $\rho(a, \{a, b, c\}) \geq \frac{1}{2}$.

Theorem 2 *If a nondegenerate stochastic choice function ρ satisfies Weak Order and Hazard Rate IIA, then there is a PALM $(u, \succsim)$ such that $\succsim^* = \succsim$ and $\rho = \rho_{(u,\succsim)}$.*

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

Proof: The sufficiency part of Theorem 2 is identical to that of Theorem 1. For the necessity part of Theorem 2, we only need to prove that $\succeq^* = \succeq$.

Lemma 1*: If $c \succ a \succ b$, or $a \succ b \succ c$, then $\frac{\rho(a,\{a,b\})}{\rho(b,\{a,b\})} > \frac{\rho(a,\{a,b,c\})}{\rho(b,\{a,b,c\})}$.

The proof of Lemma 1* is identical to the proof of Lemma 1 of Theorem 1.

Lemma 2*: If $a \sim b$, then $a \sim^0 b$.

The proof of Lemma 2* is identical to the proof of Step 1 of Lemma 2.

Lemma 3*: If $a \not\sim b$, then $a \not\sim^0 b$.

Take any a, b with $a \succ b$. Since $|\alpha_1 \cup \alpha_n| \geq 3$ by Richness*, there is c such that $c \succeq a$ or $b \succeq c$. Then $\frac{\rho(a,\{a,b\})}{\rho(b,\{a,b\})} \neq \frac{\rho(a,\{a,b,c\})}{\rho(b,\{a,b,c\})}$.

First, by Lemmas 2*-3*, we have $a_i \sim^0 b_i$, $a_j \sim^0 b_j$, and $a_i \not\sim^0 a_j$ for any $a_i, b_i \in \alpha_i$ and $a_j, b_j \in \alpha_j$. Then Lemma 1* implies that $a_i P a_{i+1}$ for any $a_i \in \alpha_i$ and $a_{i+1} \in \alpha_{i+1}$. Consequently, by the definition of $\succ^0$, we have that $a_i \succ^0 a_{i+1}$ for any $a_i \in \alpha_i$ and $a_{i+1} \in \alpha_{i+1}$. Since $\succeq^*$ is the transitive closure of $\succeq^0$, we have $a_i \succ^* a_j$

for any $a_i \in \alpha_i$ and $a_j \in \alpha_j$ with $i < j$. Finally, the following lemma concludes the proof of $\succsim^* = \succsim$.

Lemma 4*: $a_j \not\succ^0 a_i$ for any $a_i \in \alpha_i$ and $a_j \in \alpha_j$ with $i < j$.

Take any $a_i \in \alpha_i, a_j \in \alpha_j$ with $i < j$.

Case 1. $i \neq 1$.

Take any $a_1 \in \alpha_1$. Since $a_1 \succ a_i \succ a_j$, by Lemma 1*, we have $\frac{\rho(a_j,\{a_j,a_i\})}{\rho(a_i,\{a_j,a_i\})} < \frac{\rho(a_j,\{a_j,a_i,a_1\})}{\rho(a_i,\{a_j,a_i,a_1\})}$. Therefore, we cannot have $a_j P a_i$. Consequently, $a_j \not\succ^0 a_i$.

Case 2. $j \neq n$.

Take any $a_n \in \alpha_n$. Since $a_i \succ a_j \succ a_n$, by Lemma 1*, we have $\frac{\rho(a_j,\{a_j,a_i\})}{\rho(a_i,\{a_j,a_i\})} < \frac{\rho(a_j,\{a_j,a_i,a_n\})}{\rho(a_i,\{a_j,a_i,a_n\})}$. Therefore, we cannot have $a_j P a_i$. Consequently, $a_j \not\succ^0 a_i$.

Case 3. $i = 1$ and $j = n$.

By Richness*, there is $a \in \alpha_1$, $b \in \alpha_n$, and $c \notin \alpha_1 \cup \alpha_n$ such that $u(a) \geq u(b) + u(c) + u(x_0)$. It turns out that, $u(a) \geq u(b) + u(c) + u(x_0)$ implies $\frac{\rho(a,\{a,b\})}{\rho(b,\{a,b\})} > \frac{\rho(a,\{a,b,c\})}{\rho(b,\{a,b,c\})}$. Therefore, we cannot have bPa. Finally, since $a, a_1 \in \alpha_1$ and $b, a_n \in \alpha_n$, by the definition of $\succ^0$, $a_n \not\succ^0 a_1$.

■

A.3 A modification without the outside option

In this section we show that by modifying PALM, we can dispense with the outside option. In the modified model, whenever the agent chooses no alternative, she repeats the sequential procedure of PALM until she chooses some alternatives. This modified PALM is represented by the following representation:

$$\rho(a, A) = \frac{\mu(a, A) \prod_{\alpha \in A/\succsim : \alpha \succ a} \big(1 - \mu(\alpha, A)\big)}{\sum_{b \in A} \mu(b, A) \prod_{\alpha \in A/\succsim : \alpha \succ b} \big(1 - \mu(\alpha, A)\big)}$$

where

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

Let us now show that this modified PALM can rationalize violations of regularity. In fact, we obtain the following observation which is very similar to Proposition 1.

Observation 1: *If $x \succ y \succ z$ and $u(x)$ is large enough, then*

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

Proof of Observation 1: We have $\rho(y, \{x, y, z\}) > \rho(y, \{x, y\})$ iff

$$\frac{\mu(y, \{x, y, z\})(1 - \mu(x, \{x, y, z\}))}{\mu(x,\{x, y, z\})+\mu(y,\{x, y, z\})(1-\mu(x,\{x, y, z\})) + \mu(z,\{x, y, z\})(1-\mu(x,\{x, y, z\}))(1-\mu(y,\{x, y, z\}))}$$

$$> \frac{\mu(y, \{x, y\})(1 - \mu(x, \{x, y\}))}{\mu(x, \{x, y\})+\mu(y, \{x, y\})(1-\mu(x, \{x, y\}))} \text{ iff}$$

$$\frac{u(y)(1 - \mu(x, \{x, y, z\}))}{u(x)+u(y)(1-\mu(x, \{x, y, z\})) + u(z)(1-\mu(x, \{x, y, z\}))(1-\mu(y, \{x, y, z\}))}$$

$$> \frac{u(y)(1 - \mu(x, \{x, y\}))}{u(x)+u(z)(1-\mu(x, \{x, y\}))}.$$

By direct calculations, we obtain that $\rho(y, \{x, y, z\}) > \rho(y, \{x, y\})$ iff

$$u^2(x) > u(y)\big(u(y) + u(z)\big) \frac{u(x) + u(z)}{u(x) + u(y) + u(z)}.$$

Since $1 > \frac{u(x)+u(z)}{u(x)+u(y)+u(z)}$, if $u(x) > \sqrt{u(y)\big(u(y) + u(z)\big)}$, then we have $\rho(y, \{x, y, z\}) > \rho(y, \{x, y\})$. Therefore, when $u(x)$ is large enough, we can have violations of regularity.

■

The above observation illustrates that the outside option does not really play a role in explaining violations of regularity, but the sequential procedure does. We leave the axiomatic characterization of the modification as an open question.

A.4 Independence of Axioms

We shall prove that Weak Order and Hazard Rate IIA are independent.

Weak Order does not imply Hazard Rate IIA: We shall find ρ that satisfies Weak Order, but violates Hazard Rate IIA. Consider the following ρ on X: there exists a function $u : X_0 \to \mathbb{R}_{++}$ such that $u(a) \neq u(b)$ for some $a, b \in X$, and for all $A \in \mathcal{A}$ and $a \in A_0$,

$$\rho(a, A) = \begin{cases} \frac{u(a)}{\sum_{b \in A_0} u(b)} & \text{when } |A| \leq 3, \\ \frac{1}{|A|+1} & \text{when } |A| \geq 4. \end{cases}$$

Since Luce's IIA is satisfied at any menu A with $|A| \leq 3$, $a \sim^0 b$ for all $a, b \in X$. Therefore, Weak Order is satisfied. Moreover, since $q(a, A) = \rho(a, A)$, Hazard Rate IIA is equivalent to Luce's IIA. However, Luce's IIA is violated at all a, b with $u(a) \neq u(b)$ since

$$\frac{\rho(a, \{a, b\})}{\rho(b, \{a, b\})} = \frac{u(a)}{u(b)} \neq \frac{\rho(a, A)}{\rho(b, A)} = 1 \text{ when } a, b \in A \text{ and } A \geq 4.$$

Hazard Rate IIA does not imply Weak Order: We now shall find ρ that violates Weak Order, but satisfies Hazard Rate IIA.

Take any $X = \{a_i\}_{i \in \mathbb{Z}} \cup \{x\}$ and $\succ$ on X such that for each $i \in Z$, $a_i \succ a_{i+1}$ and neither $a_i \succ x$ nor $x \succ a_i$. Consider the following ρ on X: there exists a function $u : X_0 \to \mathbb{R}_{++}$ such that for all $A \in \mathcal{A}$,

$$\rho(a, A) = \mu(a, A) \prod_{b \succ a} \big(1 - \mu(b, A)\big) \text{ for any } a \in A \setminus \{x\}$$

and

$$\rho(x, A) = \mu(x, A), \text{ where } \mu(a, A) = \frac{u(a)}{\sum_{b \in A} u(b) + u(x_0)}.$$

The following four steps will prove that Weak Order is violated.

Step 1. Note that $a_i \not\succ^0 x$ for any $a_i \in X \setminus \{x\}$ since

$$\frac{\rho(a_i,\{a_i,x\})}{\rho(x,\{a_i,x\})}=\frac{u(a_i)}{u(x)}\neq\frac{\rho(a_i,\{a_i,a_{i-1},x\})}{\rho(x,\{a_i,a_{i-1},x\})}=\frac{u(a_i)(1-\mu(a_{i-1},\{a_i,a_{i-1},x\}))}{u(x)}.$$

Step 2. Note that $a_i \succ^0 a_{i+1}$ for each i since

$$\begin{aligned}\frac{\rho(a_i,\{a_i,a_{i+1}\})}{\rho(a_{i+1},\{a_i,a_{i+1}\})}&=\frac{u(a_i)}{u(a_{i+1})(1-\mu(a_i,\{a_i,a_{i+1}\}))}\\ >\frac{\rho(a_i,\{a_i,a_{i+1},x\})}{\rho(a_{i+1},\{a_i,a_{i+1},x\})}&=\frac{u(a_i)}{u(a_{i+1})(1-\mu(a_i,\{a_i,a_{i+1},x\}))}\end{aligned}$$

and for any $t\in\mathbb{Z}$ with either $t\geq 2$ or $t\leq -1$,

$$\begin{aligned}\frac{\rho(a_i,\{a_i,a_{i+1}\})}{\rho(a_{i+1},\{a_i,a_{i+1}\})}&=\frac{u(a_i)}{u(a_{i+1})(1-\mu(a_i,\{a_i,a_{i+1}\}))}\\ >\frac{\rho(a_i,\{a_i,a_{i+1},a_{i+t}\})}{\rho(a_{i+1},\{a_i,a_{i+1},a_{i+t}\})}&=\frac{u(a_i)}{u(a_{i+1})(1-\mu(a_i,\{a_i,a_{i+1},a_{i+t}\}))}.\end{aligned}$$

Step 3. Note that neither $a_i \succ^0 x$ nor $x \succ^0 a_i$ for any $a_i \in X\setminus\{x\}$ since $a_{i+1} \not\succ^0 x$, $a_{i+1} \not\succ^0 a_i$, and

$$\frac{\rho(a_i,\{a_i,x\})}{\rho(x,\{a_i,x\})}=\frac{u(a_i)}{u(x)}=\frac{\rho(a_i,\{a_i,a_{i+1},x\})}{\rho(x,\{a_i,a_{i+1},x\})}=\frac{u(a_i)}{u(x)}.$$

Step 4. Note that $a_{i+k} \not\succeq^0 a_i$ for each i and $k\geq 1$ since and $x \not\succ^0 a_i$, $x \not\succ^0 a_{i+k}$, and

$$\begin{aligned}\frac{\rho(a_i,\{a_i,a_{i+k}\})}{\rho(a_{i+k},\{a_i,a_{i+k}\})}&=\frac{u(a_i)}{u(a_{i+k})(1-\mu(a_i,\{a_i,a_{i+k}\}))}\\ >\frac{\rho(a_i,\{a_i,a_{i+k},x\})}{\rho(a_{i+k},\{a_i,a_{i+k},x\})}&=\frac{u(a_i)}{u(a_{i+k})(1-\mu(a_i,\{a_i,a_{i+k},x\}))}.\end{aligned}$$

Therefore, $\succeq^*$ is complete and transitive on $X\setminus\{x\}$ with $a_i \succ^* a_j$ for all $i<j$, but neither $a_i \succeq^* x$ nor $x \succeq^* a_i$ for each $i\in Z$. In other words, Weak Order is violated.

We now shall prove that Hazard Rate IIA is satisfied. Since ρ is a PALM with

$(\succ, u)$ on $X \setminus \{x\}$, Hazard Rate IIA is satisfied on $X \setminus \{x\}$. We will now prove that Hazard Rate IIA is satisfied at any menu $A \cup \{x\}$.

First, note that $(A \cup \{x\})_a = \{b \in A \cup \{x\} : b \succ^* a\} = \{b \in A : b \succ^* a\} = A_a = \{b \in A : b \succ a\}$ and $(A \cup \{x\})_x = \emptyset$. Therefore,

$$q(a, A \cup \{x\}) = \frac{\rho(a, A \cup \{x\})}{1 - \rho(A_a, A \cup \{x\})} = \mu(a, A \cup \{x\}) \text{ as in the PALM,}$$

and

$$q(x, A \cup \{x\}) = \rho(x, A \cup \{x\}) = \mu(x, A \cup \{x\}).$$

Therefore, Hazard Rate IIA is satisfied at $A \cup \{x\}$.

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.

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 (2015): "Stochastic choice and revealed perturbed utility," *Econometrica*, 83, 2371–2409.

Gul, F., P. Natenzon, and W. Pesendorfer (2014): "Random Choice as Behavioral Optimization," *Econometrica*, 1873–1912.

Herne, K. (1997): "Decoy alternatives in policy choices: Asymmetric domination and compromise effects," *European Journal of Political Economy*, 13, 575–589.

Huber, J., J. W. Payne, and C. Puto (1982): "Adding asymmetrically dominated alternatives: Violations of regularity and the similarity hypothesis," *Journal of consumer research*, 90–98.

Iyengar, S. S. and M. R. Lepper (2000): "When choice is demotivating: Can one desire too much of a good thing?" *Journal of personality and social psychology*, 79, 995.

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.

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.

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

Ravid, D. (2015): "Focus, Then Compare," Working Paper, Princeton University.

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.

Simonson, I. (1989): "Choice based on reasons: The case of attraction and compromise effects," *Journal of consumer research*, 158–174.

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

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

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

——— (1972): "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.