Stochastic Choice with Categorization

Jun Zhang*

May 15, 2016

Abstract

Observing that people often use categorization to simplify choice problems, we develop and axiomatize a stochastic choice model in which the decision maker first categorizes alternatives into disjoint categories, then considers categories sequentially until making a choice. The model subsumes both the Luce model and the random consideration set rule (Manzini and Mariotti, 2014) as special cases. We also develop and axiomatize a variant of the model by excluding the existence of a default option. The elements of both models are uniquely identified.

JEL Classification: D01, D81

Keywords: Stochastic Choice, Categorization

1 Introduction

People often use categorization to simplify complex choice problems. For example, in supermarkets products are exogenously categorized so that customers can quickly locate the products they want to purchase.[1] When searching for hotels through online-booking websites, people often first categorize hotels according to some criterion such as distance,

*Division of Humanities and Social Sciences, California Institute of Technology. Email: jzhang@hss.caltech.edu. Address: 1200 East California Blvd, MC 228-77, Pasadena, CA 91125. I thank Federico Echenique for guidance, and Marina Agranov and Gerelt Tserenjigmid for comments. All errors are mine.

[1] Mogilner et al. (2008) find that consumers are more satisfied when they face menus that are exogenously categorized.

price or rating. In some situations people may subjectively categorize alternatives, the process of which cannot be directly observed by others. When there are a large number of alternatives in a menu, it is almost impossible for people to pay attention to each alternative carefully. By using categorization people can efficiently allocate their attention: people can first abstractly consider categories without knowing their exact contents; then people can further consider the content of a category if they want to make a choice from it.

On the other hand, people often consider alternatives sequentially instead of simultaneously. It is either because alternatives are physically ordered when they are presented to people, or because they come to people's mind in some sequential manner. In the literature, Simon (1955) introduces the satisficing model in which the decision maker (DM) searches through alternatives until finding an alternative that meets a threshold criterion. Rubinstein and Salant (2006) axiomatize choice functions in which DM's choice depends on the order of alternatives.

In this paper we propose and axiomatize a stochastic choice model that incorporates both features of categorization and sequential consideration. We call the model *random choice rule with categorization* (RCRC). In the model all alternatives are categorized into disjoint categories. Facing a menu, DM considers the categories in the menu sequentially until making a choice from some category. Both categorization and the search order are menu-independent and in our axiomatization both are identified from the dataset. To illustrate RCRC, suppose that DM wants to book a hotel online through a website and categorizes hotels by rating before searching. Then DM first looks at the highest-rating category of hotels and tries to make a choice from it. DM randomly chooses each hotel in the same category with a probability of a Luce-type formula. If DM successfully chooses a hotel, he stops the search. However, DM also randomly fails to make a choice from the highest-rating category such that he moves his attention to the second highest-rating category and so on. If DM fails to make a choice from all categories, we say he chooses the "no choice" option. RCRC does not explain why DM may fail to make a choice from a considered category. In real situations it may be because the alternatives in the same category are too similar or unfamiliar to DM so that DM has a difficulty to make a choice.

RCRC is related to the Luce model in the sense that DM's random choice from the same category satisfies Luce's Independence of Irrelevant Alternatives (IIA). But it is different from the Luce model in that DM considers categories sequentially such that any category considered later cannot affect DM's random choice from any category considered earlier, and the odd ratio of any two alternatives from the same category is not affected by other categories. We use the first feature to identify the categorization and search order of RCRC, and formalize the second feature as an axiom to characterize RCRC. Another difference between RCRC and the Luce model is that RCRC does not necessarily have full support. In RCRC it can happen that an alternative is chosen with zero probability in a menu but is chosen with a positive probability in another menu.

RCRC is also related to the random consideration set rule (RCSR) of Manzini and Mariotti (2014). In particular, RCRC degenerates to RCSR if each category contains only one alternative. As a corollary of our result we provide a new axiomatization of RCSR. It is also interesting to compare RCRC with the attribute rule of Gul et al. (2014). In the attribute rule DM first randomly pays attention to an attribute then randomly chooses an alternative that has the attribute. DM's choice at both stages follows the Luce model. If we interpret each attribute as a category, then the difference between the two models is that in the attribute rule categories need not be disjoint and it is as if that DM considers all categories simultaneously, while in RCRC DM considers categories sequentially.

In the paper we also compare RCRC with other stochastic choice models in the literature. RCRC is not a random utility model and it can explain multiple behavioral anomalies that violate regularity and Luce's IIA.

Since in many real situations and lab experiments a default "no choice" alternative is not available, we also propose and axiomatize a variant of RCRC by excluding the default. In the variant model whenever DM considers the last category in a menu, he will make a choice from the category for sure. Here the last category in each menu plays the role of the default.[2]

The rest of the paper is organized as follows. In Section 2 we define the random choice datasets we study. In Section 3 we define RCRC and discuss its special cases. In Section 4 we identify and axiomatize RCRC. In Section 5 we define and axiomatize the variant of RCRC without default. In Section 6 we apply RCRC to explain multiple

[2] Manzini and Mariotti (2014) suggest a variant of RCSR in which whenever DM misses all alternatives in menu, DM reconsiders the whole menu until making a choice. However they do not axiomatize the model.

behavioral anomalies. In Section 7 we compare RCRC to other stochastic choice models in the literature. In Section 8 we discuss related literature. All proofs are in Appendix.

2 Random Choice Dataset

There are a finite set X of alternatives containing at least three elements, and a *domain* $\mathcal{D}$ of subsets of X. Each element of $\mathcal{D}$ is a menu from which DM needs to make a choice. $\mathcal{D}$ may not contain every subset of X, but it satisfies the following "richness" assumption: $\{a, b, c\} \in \mathcal{D}$ for all distinct $a, b, c \in X$ and for any $A, B \subseteq X$, if $A \subseteq B$ and $B \in \mathcal{D}$ then $A \in \mathcal{D}$. If DM does not choose any alternative from a menu, we say he chooses the "no choice" option a^* (default alternative). Denote $X^* \equiv X \cup \{a^*\}$ and $A^* \equiv A \cup \{a^*\}$ for all $A \in \mathcal{D}$. A random choice dataset is defined as a random choice rule as follows.

Definition 1. *A random choice rule is a map* $p : X^* \times \mathcal{D} \to [0, 1]$ *such that:* $\sum_{a \in A^*} p(a, A) = 1$ *for all* $A \in \mathcal{D}$*;* $p(a, A) = 0$ *for all* $a \notin A^*$*;* $p(a, A) \in [0, 1]$ *for all* $a \in A^*$ *and all* $A \in \mathcal{D} \backslash \emptyset$*;* $p(a^*, A) = 1$ *if and only if* $A = \emptyset$*.*

Here $p(a, A)$ is the probability that DM chooses $a \in A^*$ from A. If A is empty, DM chooses the default a^* for sure. If A is nonempty, since $p(a, A) \in [0, 1]$, we allow DM to choose some alternative in A with zero probability; but we require that DM choose some alternative in A with a positive probability since $p(a^*, A) = 1$ only if $A = \emptyset$. For all $A \in \mathcal{D} \backslash \emptyset$ and all $B \subseteq A^*$, we denote by $p(B, A) \equiv \sum_{a \in B} p(a, A)$ the probability that some alternative in B is chosen from A.

3 The Model

In our model all alternatives in X are categorized into disjoint categories. This categorization can be objective or subjective. To denote it we introduce a transitive *categorization relation* $\sim$ such that for any $a, b \in X$, a and b belong to the same category if and only if $a \sim b$. So all categories form a partition of X, denoted by $\tilde{X} \equiv \{X_1, X_2, \ldots, X_{k_X}\}$. The categories in each menu A are similarly denoted by $\tilde{A} \equiv \{A_1, A_2, \ldots, A_{k_A}\}$. In our model DM considers the categories in each menu sequentially following a menu-independent order. We denote the order by $\succ$ and call it *search order*. Without loss of generality we assume that $X_i \succ X_{i+1}$ for all $1 \leq i < k_X$ and $A_i \succ A_{i+1}$ for all $A \in \mathcal{D} \backslash \{\emptyset\}$ and all

$1 \le i < k_A$. For all $A \in \mathcal{D}\backslash\{\emptyset\}$ and all $a \in A$, we denote by A_a the category in A that contains a. For convenience we define a binary relation $\succsim$ to subsume both $\succ$ and $\sim$: for all $a, b \in X$ and all $A \in \mathcal{D}$, we say $a \succ b$ if $X_a \succ X_b$, $a \succsim b$ if $a \sim b$ or $a \succ b$, and $A \succsim b$ if $a \succsim b$ for all $a \in A$. By this definition $\succsim$ is a weak order on X.

Facing a menu A, DM considers the categories in A sequentially according to the search order $\succ$. So DM first considers A_1 and tries to make a choice from it. DM chooses each $a \in A_1$ with a probability of $\gamma(a, A_1) \equiv \frac{u(a)}{\sum_{b \in A_1} u(b) + u(A_1)}$, which is a Luce-type formula. Here $u(a)$ reflects the desirability (attractiveness) of a relative to other alternatives in A_1. We denote by $\pi(A_1)$ the probability that some alternative in A_1 is chosen. However, DM may choose nothing from A_1, the probability of which is $1 - \pi(A_1) = \frac{u(A_1)}{\sum_{b \in A_1} u(b) + u(A_1)}$. Here $u(A_1)$ is not the measure of the desirability of A_1, instead it is used to capture the chance that DM chooses nothing from A_1. After missing A_1 DM considers the next category A_2 and tries to make a choice from it. If he also misses A_2, he further considers A_3 and so on. If DM misses all categories in A, he chooses a^*. So DM chooses some $a \in A_i$ only if he misses all category A_j such that $j < i$. Hence, DM chooses any $a \in A_i$ with a probability of $\gamma(a, A_i) \prod_{j<i}[1 - \pi(A_j)]$ and chooses a^* with a probability of $\prod_{A_j \in \tilde{A}}[1 - \pi(A_j)]$.

We define the model formally as follows.

Definition 2. *A* ***random choice rule with categorization*** *(RCRC) is a random choice rule* $p_{\succsim,u}$ *for which* $\succsim$ *is a weak order on* X*,* u *is a function defined on* $X \cup \{A_i \in \tilde{A}\}_{A \in \mathcal{D}\backslash\{\emptyset\}}$ *such that* $X \xrightarrow{u} \mathbf{R}_{++}$ *and* $\{A_i \in \tilde{A}\}_{A \in \mathcal{D}\backslash\{\emptyset\}} \xrightarrow{u} \mathbf{R}_{+}$*, and*

$$p_{\succsim,u}(a, A) = \gamma(a, A_a) \prod_{A_i \in \tilde{A}: A_i \succ A_a} [1 - \pi(A_i)] \text{ for all } a \in A \text{ and all } A \in \mathcal{D}\backslash\{\emptyset\} \quad (1)$$

where

$$\gamma(a, A_a) \equiv \frac{u(a)}{\sum_{b \in A_a} u(b) + u(A_a)}, \quad \pi(A_i) \equiv \sum_{b \in A_i} \gamma(b, A_i). \quad (2)$$

In the above definition we require that $u(a) > 0$ for all $a \in X$. This implies that once DM considers a category in a menu, DM chooses each alternative in the category with a positive probability. However we only require that $u(A_i) \ge 0$ for all $A_i \in \tilde{A}$ and all $A \in \mathcal{D}\backslash\{\emptyset\}$. So if $u(A_i) = 0$ for some A_i, then once DM considers A_i, he will make a choice from A_i for sure and not consider any category A_j such that $j > i$. This implies that some alternative can be chosen with zero probability in some menu. It can also happen that some alternative is chosen with zero probability in some menu but with a positive probability in another menu.[3]

Special Cases of RCRC:

Manzini and Mariotti (2014) propose RCSR by assuming that DM has standard preferences but pays attention to alternatives randomly and independently. It is interesting to note that RCSR is a special case of RCRC. In particular, if each category contains only one alternative in RCRC, then the representation of RCRC degenerates to the representation of RCSR:

$$p_{\succ,\gamma}(a, A) = \gamma(a) \prod_{b \in A: b \succ a} [1 - \gamma(b)] \text{ for all } A \in \mathcal{D} \backslash \{\emptyset\} \text{ and all } a \in A^*. \tag{3}$$

where $\gamma(a) \equiv \frac{u(a)}{u(a)+u(\{a\})}$ is the probability that DM pays attention to each $a \in X$. Manzini and Mariotti (MM hereafter) focus on random choice datasets with full support, that is, $p_{\succ,\gamma}(a, A) \in (0, 1)$. We can achieve this by requiring that $u(\{a\}) > 0$ for all $a \in X$.

It is obvious that RCRC subsumes the Luce model as a special case when all alternatives belong to the same category.

4 Characterization of RCRC

In this section we characterize RCRC using three axioms. We first show how to identify DM's categorization relation and search order from a random choice dataset p. Since the search order is menu-independent, any lower-ranked category in the search order does not affect DM's random choice from any higher-ranked category. We use this feature to identify $\succsim$. In particular, if removing an alternative b has no impact on the choice probability of an alternative a from every menu containing them, then we conclude that a must be in a higher-ranked category than b, that is, $a \succ b$. Otherwise, $b \succsim a$.

Definition 3 (Revealed relation)**.** *For all $a, b \in X$,*

(1) $a \succ^* b$ if $p(a, A \backslash \{b\}) = p(a, A)$ for all $A \in \mathcal{D}$;

(2) $b \succsim^* a$ if $p(a, A \backslash \{b\}) \neq p(a, A)$ for some $A \in \mathcal{D}$.

[3]Suppose there are two menus A, B such that $a \in B_2 \cap A_2$. If $u(A_1) = 0$ but $u(B_1) > 0$, then a will be chosen with zero probability in A but with a positive probability in B.

When $a \succsim^* b$ and $b \succsim^* a$, we say $a \sim^* b$. Then $\succ^*$ is the revealed search order and $\sim^*$ is the revealed categorization relation. When $a \succ^* b$, we also say $a \succsim^* b$.

It is easy to see that $\succsim^*$ is complete since either case (1) or case (2) of Revealed relation must happen. However $\succsim^*$ may not be transitive. So our first axiom *Transitivity* requires that $\succsim^*$ be transitive.

Transitivity: If there are $a^1, a^2, \ldots, a^n \in X$ such that $a^1 R^1 a^2 \cdots a^n R^n a^1$ where $R^i \in \{\succ^*, \succsim^*\}$ for all $1 \leq i \leq n$, then $R^i = \succsim^*$ for all $1 \leq i \leq n$.

In RCRC we assume that $u(a) > 0$ for all $a \in X$, so once DM considers a category, he must choose every alternative in the category with a positive probability. This implies that DM's random choice from any menu consisting of a single category has full support. We formalize it as our second axiom.

Single-category Full Support: For all $A \in \mathcal{D} \backslash \{\emptyset\}$ such that $A = A_1$, $p(a, A) > 0$ for all $a \in A$.

Our last axiom is an adaption of Luce's IIA and it includes two parts. RCRC assumes that DM follows a Luce-type formula to make a choice from a category, so the odd ratio of any two alternatives in the same category must satisfy Luce's IIA. We formalize this feature as the first part of the axiom. Moreover, in RCRC whenever DM considers the last category in a menu, his choice is not affected by the alternatives in higher-ranked categories in the menu since he has missed them. So in the second part we require that the odd ratio of the default and any alternative in the last category of a menu does not change if we remove any alternatives in higher-ranked categories of the menu.

I-IIA:

(1) For all $a, b \in X$, if $a \sim^* b$, then

$$\frac{p(a, A)}{p(b, A)} = \frac{p(a, B)}{p(b, B)} \tag{4}$$

for all $A, B \in \mathcal{D}$ such that $\{a, b\} \subseteq A \cap B$ and $p(b, A)p(b, B) > 0$.

(2) For all $a \in X$,

$$\frac{p(a^*, A)}{p(a, A)} = \frac{p(a^*, A \backslash B)}{p(a, A \backslash B)} \tag{5}$$

for all $B \subseteq A \in \mathcal{D}$ such that $a \in A \succsim^* a$, $B \succ^* a$, and $p(a, A)p(a, A \backslash B) > 0$.

Theorem 1 below will prove that a random choice dataset that satisfies Transitivity, Single-category Full Support and I-IIA must be a RCRC. However, the converse is not true because a RCRC may not satisfy Transitivity. We show it through the following example.

Example 1. *$X = \{a, b, c\}$, $\mathcal{D} = 2^X$, $a \sim b \succ c$, $u(\{a\}) = u(b) + u(\{a, b\})$ and $u(\{b\}) = u(a) + u(\{a, b\})$. It is easy to see that $p_{\succsim,u}(a, \{a\}) = p_{\succsim,u}(a, \{a, c\}) = \frac{u(a)}{u(a)+u(\{a\})} = \frac{u(a)}{u(a)+u(b)+u(\{a,b\})} = p_{\succsim,u}(a, \{a, b\}) = p_{\succsim,u}(a, \{a, b, c\})$. Similarly we have $p_{\succsim,u}(b, \{b\}) = p_{\succsim,u}(b, \{b, c\}) = p_{\succsim,u}(b, \{a, b\}) = p_{\succsim,u}(b, \{a, b, c\})$. By Revealed relation, $b \succ^* a$ and $a \succ^* b$. So RCRC does not satisfy Transitivity.*

In the above example, we deliberately choose the function u such that removing either a or b has no impact on the choice probability of the other in every menu. So although $a \sim b$, we do not identify their relation correctly. However, if the function u is perturbed a little bit, we can identify their relation correctly. In general as long as there is one menu in which removing a has an impact on the choice probability of b and there is one menu in which removing b has an impact on the choice probability of a, then we can identify that $a \sim^* b$. Since u satisfies this property generically, we call RCRC that have this property *generic* and denote them by gRCRC.

Definition 4. *A RCRC $p_{\succsim,u}$ is **generic** if for all $a, b \in X$ such that $a \sim b$, there exist $A, B \in \mathcal{D}$ such that $\{a, b\} \subseteq A \cap B$, $A \sim B \sim a$, and $u(A\backslash\{b\}) \neq u(b) + u(A)$ and $u(B\backslash\{a\}) \neq u(a) + u(B)$.*[4]

Now we prove that Transitivity, Single-category Full Support and I-IIA fully characterize gRCRC. Moreover, since gRCRC is generic, $\succsim$ is uniquely identified. Since RCRC satisfies Luce's IIA in each category, it is not surprising that u is uniquely identified up to category-wise positive multiplication.

Theorem 1. *A random choice rule satisfies Transitivity, Single-category Full Support and I-IIA if and only if it is a generic $p_{\succsim,u}$.*

Moreover, $\succsim$ is unique and u is unique up to category-wise positive multiplication. That is, for any generic $p_{\succsim',u'}$ such that $p_{\succsim',u'} = p_{\succsim,u}$, we know that $\succsim' = \succsim$ and there exists some $\beta_{X_i} > 0$ such that $u'(a) = \beta_{X_i} u(a)$ and $u'(A) = \beta_{X_i} u(A)$ for all $X_i \in \tilde{X}$, all $a \in X_i$ and all $A \in \mathcal{D}\backslash\{\emptyset\}$ such that $A \subseteq X_i$.

[4] Here A and B can be the same menu.

The proof of the theorem can be found in Appendix A. The necessity part is obvious, while the main idea behind the sufficiency part is also simple. Since the random choice rule p satisfies Transitivity, the revealed relation $\succsim^*$ is a weak order. So the categories in any nonempty menu A is well-defined and we denote them by $\{A_1, A_2, \ldots, A_{k_A}\}$ such that $A_i \succ^* A_{i+1}$. By Single-category Full Support and I-IIA(1), there exists a positive function u defined on X such that for all $a \sim^* b$, $\frac{p(a,A)}{p(b,A)} = \frac{u(a)}{u(b)}$ if $p(a,A)p(b,A) > 0$. For each A_i we define $u(A_i) \equiv \frac{p(a^*,A_i)}{1-p(a^*,A_i)} \sum_{c \in A_i} u(c)$ so that $p(a, A_i) = \frac{u(a)}{\sum_{c\in A_i} u(c)+u(A_i)}$ for all $a \in A_i$. Now for all $a \in A_1$, Revealed relation implies that $p(a, A) = p(a, A_1)$. So $p(a, A) = \frac{u(a)}{\sum_{c\in A_1} u(c)+u(A_1)}$.

If $u(A_1) > 0$, then $1 - \sum_{c\in A_1} p(c, A) > 0$. Then for all $a \in A_2$, Revealed relation implies that $p(a, A) = p(a, A_1 \cup A_2)$. So

$$
\begin{aligned}
\frac{p(a, A)}{1 - \sum_{c\in A_1} p(c, A)} = \frac{p(a, A_1 \cup A_2)}{1 - \sum_{c\in A_1} p(c, A_1 \cup A_2)} &= \frac{p(a, A_1 \cup A_2)}{p(a^*, A_1 \cup A_2) + \sum_{b\in A_2} p(b, A_1 \cup A_2)} \\
&= \frac{p(a, A_2)}{p(a^*, A_2) + \sum_{b\in A_2} p(b, A_2)} \\
&= p(a, A_2) \\
&= \frac{u(a)}{\sum_{b\in A_2} u(b) + u(A_2)},
\end{aligned}
$$

where the third equality follows I-IIA. Hence,

$$
p(a, A) = \frac{u(A_1)}{\sum_{c\in A_1} u(c) + u(A_1)} \frac{u(a)}{\sum_{b\in A_2} u(b) + u(A_2)}. \tag{6}
$$

If $u(A_1) = 0$, then $p(A_1, A)$=1 and $p(a, A) = 0$ for all $a \in A_2$. So equation (6) still holds. Repeating the above procedure for all A_i we can prove the sufficiency part.

4.1 Characterization of RCSR

Theorem 1 implies a new characterization of MM's RCSR. In particular, since RCRC degenerates to RCSR when each category contains only one alternative, we need an axiom to rule out $a \sim^* b$ for any a, b. So we strengthen Transitivity to *Acyclicity* which requires that $\succsim^*$ be acyclic.

Acyclicity: There do not exist $a^1, a^2, \ldots, a^n \in X$ such that $a^1 R^1 a^2 \cdots a^n R^n a^1$ where $R^i \in \{\succ^*, \succsim^*\}$ for all $1 \le i \le n$.

Single-category Full Support is automatically satisfied since $p(a^*, A) = 1$ if and only if $A = \emptyset$. Since there does not exist $a \sim^* b$, we do not need I-IIA(1) anymore. So we can characterize RCSR as follows.

Corollary 1. *(Characterization of RCSR) A random choice rule satisfies Acyclicity and I-IIA(2) if and only if it is a RCSR* $p_{\succsim,\gamma}$*. Moreover, both* $\succ$ *and* γ *are unique.*

Recall that $\gamma(a) \equiv \frac{u(a)}{u(a)+u(\{a\})}$. Since u is unique up to category-wise positive multiplication, γ is unique.

MM assume that the random choice dataset has full support and use the following two axioms to characterize RCSR.

I-Asymmetry: $\frac{p(a,A\backslash\{b\})}{p(a,A)} \neq 1 \Rightarrow \frac{p(b,A\backslash\{a\})}{p(b,A)} = 1$.

I-Independence: $\frac{p(a,A\backslash\{b\})}{p(a,A)} = \frac{p(a,B\backslash\{b\})}{p(a,B)}$ and $\frac{p(a^*,A\backslash\{b\})}{p(a^*,A)} = \frac{p(a^*,B\backslash\{b\})}{p(a^*,B)}$.

There are several differences between MM's characterization and Corollary 1.

First, we use different identification methods. MM identify $a \succ^* b$ if $p(b, A\backslash\{a\}) > p(b, A)$ for some menu A. That is, removing a from some menu has a positive impact on the choice probability of b from the menu. This definition itself does not guarantee that $\succ^*$ is complete. By contrast, we identify $a \succ^* b$ if removing b has no impact on the choice probability of a in all menus; otherwise we identify $b \succsim^* a$. By our definition $\succsim^*$ is always complete.

Second, we use different axioms. MM need to use I-Asymmetry and I-Independence together to prove that $\succ^*$ is complete, transitive and asymmetric, and the random choice rule p is a RCSR. By contrast, our axioms play straightforward roles in Corollary 1: Acyclicity implies that $\succsim^*$ is transitive and asymmetric, while I-IIA(2) implies that the random choice rule p is a RCSR. If we assume that the random choice dataset has full support as MM do, we can rewrite I-IIA(2) as

$$\frac{p(a, A\backslash\{b\})}{p(a, A)} = \frac{p(a^*, A\backslash\{b\})}{p(a^*, A)} \text{ for all } a \in A \succsim^* a \text{ and } b \succ^* a.$$

Then the difference between I-IIA(2) and MM's I-Independence is that I-Independence requires that the ratios of $\frac{p(a,A\backslash\{b\})}{p(a,A)}$ and $\frac{p(a^*,A\backslash\{b\})}{p(a^*,A)}$ be independent of the menu A, while I-IIA(2) requires that $\frac{p(a,A\backslash\{b\})}{p(a,A)} = \frac{p(a^*,A\backslash\{b\})}{p(a^*,A)}$ hold conditionally for some a, b in any menu A.

MM show that RCSR satisfies another property called I-Neutrality[5], which states that if $c \succ^* a$ and $c \succ^* b$, then $\frac{p(a,A\backslash\{c\})}{p(a,A)} = \frac{p(b,A\backslash\{c\})}{p(b,A)}$. In our model we can treat a^* as belonging

[5] I-Neutrality: $\frac{p(a,A\backslash\{c\})}{p(a,A)}, \frac{p(b,A\backslash\{c\})}{p(b,A)} > 1 \Rightarrow \frac{p(a,A\backslash\{c\})}{p(a,A)} = \frac{p(b,A\backslash\{c\})}{p(b,A)}$.

to the worst category in all menus, then it is easy to see that I-IIA(2) is weaker than I-Neutrality.

In an earlier version of their paper (Manzini and Mariotti, 2011), MM provide another characterization of RCSR using three axioms: Acyclicity, Regularity and Stochastic binariness.[6] Acyclicity plays the same role in their result as in ours. But since we use difference identification methods, our characterization is still independent of theirs.

5 A Variant of RCRC without Default Option

In previous sections we assume that there is a default alternative such that whenever DM does not make a choice from a menu, he chooses the default. However, in many lab experiments there does not exist such a option (e.g., Agranov and Ortoleva, 2015). In this section we propose a variant of RCRC by excluding the existence of a^*.

5.1 RCRC Without Default

In RCRC we assume that DM chooses a^* if he misses all categories in a menu. Since a^* is not available now we assume that whenever DM considers the last category in a menu and is going to miss it, he will reconsider the last category and force himself to make a choice. This leads to the following representation. To differentiate it from RCRC we call it RCRC-D and denote it by $\bar{p}_{\succsim,u}$.

Definition 5. *A* ***random choice rule with categorization but without default*** *(RCRC-D) is a random choice rule $\bar{p}_{\succsim,u}$ for which $\succsim$ is a weak order on X, u is a function defined on $X \cup \{A_i \in \tilde{A} : A_i \nsubseteq X_{k_X}\}_{A \in \mathcal{D}\backslash\{\emptyset\}}$ such that $X \xrightarrow{u} \mathbf{R}_{++}$ and $\{A_i \in \tilde{A} : A_i \nsubseteq X_{k_X}\}_{A \in \mathcal{D}\backslash\{\emptyset\}} \xrightarrow{u} \mathbf{R}_+$, and*

$$\bar{p}_{\succsim,u}(a, A) = \gamma(a, A) \prod_{A_i \in \tilde{A} : A_i \succ A_a} [1 - \pi(A_i)] \text{ for all } a \in A \text{ and all } A \in \mathcal{D}\backslash\{\emptyset\} \quad (7)$$

[6]Regularity: for all $A, B \in \mathcal{D}$ such that $\{a, b\} \subseteq B \subseteq A$, $p(a, A) \leq p(a, B)$.

Stochastic binariness: for all $A \in \mathcal{D}$ and $\{a, b\} \subseteq A$, if $p(a, A\backslash\{b\}) > p(a, A)$ then $\frac{p(a,A)}{p(a,A\backslash\{b\})} = 1 - p(b, \{a, b\})$.

where

$$\pi(A_i) \equiv \frac{\sum_{b\in A_i} u(b)}{\sum_{b\in A_i} u(b) + u(A_i)}, \quad \gamma(a,A) = \begin{cases} \dfrac{u(a)}{\sum_{b\in A_a} u(b) + u(A_a)}, & \text{if } A_a \neq A_{k_A}, \\ \\ \dfrac{u(a)}{\sum_{b\in A_a} u(b)}, & \text{otherwise,} \end{cases}$$

and X_{k_X}, A_{k_A} *are the last categories in* X *and* A *respectively.*

In RCRC-D there is "discontinuity" in the definition of $\gamma(a, A)$ in each menu. That is, $\gamma(a, A) = \frac{u(a)}{\sum_{b\in A_a} u(b)+u(A_a)}$ if $A_a \neq A_{k_A}$ but $\gamma(a, A) = \frac{u(a)}{\sum_{b\in A_a} u(b)}$ if $A_a = A_{k_A}$. This guarantees that DM always makes a choice from each menu.

5.2 Characterization of RCRC-D

In this subsection we provide a characterization of RCRC-D. Since a^* is not available, for the random choice rule p considered in this subsection we have $p(a^*, A) = 0$ for all $A \in \mathcal{D}\backslash\{\emptyset\}$.

Because of the "discontinuity" in the definition of $\gamma(a, A)$, the identification method we use to characterize RCRC does not work here. It is because in RCRC-D removing the last category A_{k_A} from a menu A may change the choice probabilities of alternatives in A_{k_A-1}. For example, suppose $a \succ b$ and $\bar{p}_{\succsim,u}(a, \{a, b\}) < 1$. Then we have $\bar{p}_{\succsim,u}(a, \{a, b\}\backslash\{b\}) = 1 > \bar{p}_{\succsim,u}(a, \{a, b\})$. If using our previous method we would identify that $b \succsim^* a$.

However, if there exists $c \sim b$, then $a \succ b$ implies that $\bar{p}_{\succsim,u}(a, A\backslash\{b\}) = \bar{p}_{\succsim,u}(a, A)$ for all $\{a, b, c\} \subseteq A$ since a does not belong to the last category in both A and $A\backslash\{b\}$. In the following we will use this fact to identify $\succ$. But to do it we first need to identify $\sim$. Using the similar idea behind the definition of gRCRC, we will use the evidence that removing each of a, b changes the choice probability of the other in each menu containing them to identify $a \sim b$. Moreover, if we identify that $a \sim b \sim c$, then we also know that $a \sim c$.

Definition 6. *(Revealed relation-D) For all* $a, b \in X$,

(1) $a \sim^1 b$ *if* $p(a, A) \neq p(a, A\backslash\{b\})$ *and* $p(b, A) \neq p(b, A\backslash\{a\})$ *for all* $\{a, b\} \subseteq A \in \mathcal{D}$.

Let $\sim^*$ *be the transitive closure of* $\sim^1$. *That is,* $a \sim^* c$ *if there exist* $c_1, c_2, \ldots, c_m$ *such that* $a \sim^1 c_1 \sim^1 c_2 \cdots c_m \sim^1 c$.

(2) $a \succ^2 b$ *if (i) there exist* $\{a, b\} \subseteq A \in \mathcal{D}$ *and some* $c \in A \backslash \{a, b\}$ *such that* $b \sim^* c$, *and (ii)* $p(a, A) = p(a, A \backslash \{b\})$ *for all* A *satisfying (i).*

Let $\succsim^*$ be the transitive closure of $\sim^1$ and $\succ^2$:

$a \succsim^* b$ if $\exists c_1, c_2, \ldots, c_m$ such that $aR^1c_1R^2c_2 \cdots c_mR^{m+1}b$ where all $R^i \in \{\sim^1, \succ^2\}$.

If $R^i = \succ^2$ for some i, we say $a \succ^* b$. Without any assumption $\succsim^*$ may be incomplete and intransitive. Our first axiom requires $\succsim^*$ be a weak order.

Weak Order: $\succsim^*$ is a weak order.

Our second axiom is similar with but stronger than Single-category Full Support. It says that if DM chooses a subset of alternatives in a menu with a positive probability, then DM must choose each alternative in the highest-ranked category in the subset with a positive probability. This axiom is equivalent to Single-category Full Support in the presence of other axioms in Section 4.

Highest-cateogry Full Support: For all $A \in \mathcal{D} \backslash \{\emptyset\}$ and all nonempty $B \subseteq A$, if $p(B, A) > 0$, then $p(a, A) > 0$ for all $a \in B$ such that $a \succsim^* B$.

Our third axiom is also an adaption of I-IIA. Its first part is same with I-IIA(1). Since a^* is not available now, its second part is different from I-IIA(2): we require that removing any alternatives in higher-ranked categories have no impact on the odd ratio of any two alternatives in lower-ranked categories. RCRC-D satisfies this because the categories DM has missed from a menu do not affect DM's choice from the remaining categories.

I-IIA-D:

(1) For all $a, b \in X$, if $a \sim^* b$, then

$$\frac{p(a, A)}{p(b, A)} = \frac{p(a, B)}{p(b, B)} \tag{8}$$

for all $A, B \in \mathcal{D}$ such that $\{a, b\} \subseteq A \cap B$ and $p(b, A)p(b, B) > 0$.

(2) For all $a, b \in X$,

$$\frac{p(a, A)}{p(b, A)} = \frac{p(a, A \backslash B)}{p(b, A \backslash B)} \tag{9}$$

for all $B \subseteq A \in \mathcal{D}$ such that $\{a, b\} \subseteq A$, $B \succ^* \{a, b\}$, and $p(b, A)p(b, A \backslash B) > 0$.

We need one more axiom called L-Independence (lower-independence). It says that if a menu A is a strict subset of both menu B and menu C and each alternative in A belongs to a higher-ranked category than the remaining alternatives in B and C, then each alternative in A should be chosen with the same probability in both B and C. RCRC-D satisfies this axiom since in both B and C DM considers the alternatives in A before considering other alternatives. So his random choice from A must coincide in both menus. We do not need this axiom to characterize RCRC since it is implied by the identification method in Section 4.[7]

L-Independence: For all $A, B, C \in \mathcal{D}\backslash\{\emptyset\}$ such that $A \subsetneq B \cap C$, if $A \succ^* (B \cup C)\backslash A$, then $p(a, B) = p(a, C)$ for all $a \in A$.

Before stating the characterization theorem we first define two properties of RCRC-D. For any $\bar{p}_{\succsim,u}$ and any $a, b \in X$ such that $a \sim b$, we say a and b have *mutual impact*, denoted by $a \wedge b$, if removing each of them has an impact on the choice probability of the other in every menu containing them. That is, $\bar{p}_{\succsim,u}(a, A) \neq \bar{p}_{\succsim,u}(a, A\backslash\{b\})$ and $\bar{p}_{\succsim,u}(b, A) \neq \bar{p}_{\succsim,u}(b, A\backslash\{a\})$ for all $\{a, b\} \subseteq A \in \mathcal{D}$. Then we say $\bar{p}_{\succsim,u}$ is *strongly generic*[8] if any two same-category alternatives either have mutual impact, or they can be connected through a sequence of alternatives in which every alternative has mutual impact with its neighbors.

Definition 7. *A RCRC-D $\bar{p}_{\succsim,u}$ is **strongly generic** if for all $a, b \in X$ such that $a \sim b$, either $a \wedge b$ or there exist $c^1, \cdots, c^n \in X$ such that $a \wedge c^1 \wedge c^2 \cdots c^n \wedge b$.*

In Revealed relation-D we identify $a \succ b$ only if there exists another c such that $b \sim c$. That is, b cannot belong to a singleton category. So we say a $\bar{p}_{\succsim,u}$ is *rich* if every category in X other than X_1 contains at least two alternatives.

Definition 8. *A RCRC-D $\bar{p}_{\succsim,u}$ is **rich** if $|X_i| \geq 2$ for all $X_i \in \tilde{X}\backslash\{X_1\}$.*

Now we are ready to present the characterization theorem.

Theorem 2. *A random choice rule satisfies Weak Order, Highest-category Full Support, I-IIA-D and L-Independence if and only if it is a strongly generic and rich RCRC-D $\bar{p}_{\succsim,u}$. Moreover, $\succsim$ is unique and u is unique up to category-wise positive multiplication.*

[7] In Section 4 by Revealed relation, $p(a, B) = p(a, A) = p(a, C)$ for all $a \in A$.

[8] $\bar{p}_{\succsim,u}$ is always generic since for any $a \sim b$, $\bar{p}_{\succsim,u}(a, \{a\}) = \bar{p}_{\succsim,u}(b, \{b\}) = 1$ but $\bar{p}_{\succsim,u}(a, \{a, b\}) < 1$ and $\bar{p}_{\succsim,u}(b, \{a, b\}) < 1$.

6 Applications

6.1 Violations of Regularity: Choice Overload and Attraction Effect

MM have shown that RCSR can explain choice frequency reversals and violations of stochastic transitivity. Since RCSR is a special case of RCRC, RCRC can also explain these phenomena. RCSR is a random utility model and all random utility models must satisfy regularity, which says the choice probability of any alternative in a menu cannot increase if new alternatives are added to the menu. RCRC can violate regularity, so it is not a random utility model.

Specifically, in $p_{\succsim,u}$ if we add a new alternative a to a menu A, it will not affect the choice probability of any alternative $b \in A$ if $b \succ a$. Suppose $u(A_j) > 0$ for all $A_j \in \tilde{A}$ such that $A_j \succ a$. Then if there does not exist $b \in A$ such that $a \sim b$, then adding a to A will create a new category $\{a\}$ in $A \cup \{a\}$ such that all $b \prec a$ will be chosen with smaller probabilities than before. If there exists $b \in A$ such that $a \sim b$, then let $A_i \in \tilde{A}$ be the category such that $A_i \sim a$. Although adding a to A will increase the size of the category A_i, if $u(A_i \cup \{a\})$ is sufficiently bigger than $u(A_i)$, the choice probability of all $b \prec a$ can weakly increase and some can strictly increase. Indeed, it is easy to check that if $[u(A_i \cup \{a\}) - u(A_i)] \sum_{c \in A_i} u(c) > u(a)u(A_i)$, then $p_{\succsim,u}(b, A \cup \{a\}) \geq p_{\succsim,u}(b, A)$ for all $b \prec a$ and $p_{\succsim,u}(b, A \cup \{a\}) > p_{\succsim,u}(b, A)$ for all $b \in A_{i+1}$. On the other hand, if $u(A_i \cup \{a\})$ is sufficiently smaller than $u(A_i)$, the choice probability of all $b \in A_i$ can increase. Indeed, if $u(a) + u(A_i \cup \{a\}) < u(A_i)$, then $p_{\succsim,u}(b, A \cup \{a\}) > p_{\succsim,u}(b, A)$ for all $b \in A_i$.

A well-known violation of regularity is *choice overload*: when DM faces more alternatives, DM fails to make a choice more often. It is confirmed in multiple field experiments.[9] A common feature in these experiments is that DM faces menus consisting of similar alternatives. So we can summarize choice overload as: $p(a^*, A \cup \{a\}) > p(a^*, A)$ if $a \notin A$ and $a \sim A$. As discussed above, RCRC can exhibit choice overload.

[9] Iyengar and Lepper (2000) set a tasting booth at a local grocery store that offers two menus to consumers: one menu consisting of 6 flavors of jam and the other consisting of 24 flavors. They find that 30% of consumers make a choice facing the first menu but only 3% make a choice facing the second menu. Follow-up experiments have similar findings by using chocolates (Chernev, 2003) and pens (Shah and Wolford, 2007).

Another well-known violation of regularity is the *attraction effect* (Simonson and Tversky, 1992). Suppose there are three alternatives $\{a, b, c\}$ such that b clearly dominates c but a does not dominate c, then the attraction effect says that adding c to the menu $\{a, b\}$ will increase the choice probability of b, that is, $p(b, \{a, b, c\}) > p(b, \{a, b\})$. Since b and c are more comparable, if we treat $\{b, c\}$ as a category and treat $\{a\}$ as another category, then RCRC can exhibit the attraction effect.

6.2 Violations of IIA: Duplicates Problem, Similarity Effect and Compromise Effect

The most famous critique on Luce's IIA is probably the "duplicate problem" pointed out by Debreu (1960). It says that if duplicates of an alternative in a menu are added to the menu, it should not change the choice probabilities of other alternatives in the menu. The Luce model violates this property. RCRC can satisfy this property if we treat the duplicates of an alternative as a category and choose the function u properly such that a category is always chosen with a constant probability. That is, $\frac{u(A_i)}{\sum_{a \in A_i} u(a) + u(A_i)}$ is a constant for all $A_i \in \tilde{A}$.

Similarity effect and compromise effect are two other well-known violations of Luce's IIA. Similarity effect says that if there are three alternatives $\{a, b, c\}$ such that c is similar to a but distinct from b, then adding c to the menu $\{a, b\}$ will reduce the odd ratio of a relative to b. Formally,

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

Compromise effect says that if there are three alternatives $\{a, b, c\}$ such that a, c are extreme in some attributes while b is moderate, then adding c to the menu $\{a, b\}$ will make b being chosen more often than a, in contrast to that a is chosen more often than b from $\{a, b\}$. Formally,

$$\frac{p(a, \{a, b, c\})}{p(b, \{a, b, c\})} < 1 \leq \frac{p(a, \{a, b\})}{p(b, \{a, b\})}$$

Since a, c are similar in the similarity effect and are extreme in the comprise effect, if we treat $\{a, c\}$ as a category and treat $\{b\}$ as another category, then RCRC can explain both effects. The details are in Appendix B.

7 Comparison with Related Stochastic Choice Models

7.1 Models with Stochastic Consideration Set

We have shown that RCRC subsumes RCSR as a special case. Actually we may interpret RCRC also as a model with stochastic consideration set: the search order $\succ$ is DM's preference relation over categories; facing a menu A, DM first pays attention to each category A_i randomly and independently with probability $\pi(A_i)$, then DM chooses an alternative from the best category in his consideration set; if the consideration set is empty, he chooses a^*.

Given a RCRC $p_{\succsim,u}$, without other information we do not know whether the above limited attention interpretation or the sequential search interpretation we use throughout the paper is correct. However, we believe that the limited attention interpretation can be inappropriate in the presence of categories for several reasons. First, when categories are objective, even though DM does not make a choice from some categories, it does not mean that DM ignores their existence. For example, even though a consumer does not purchase a coffee maker from a supermarket, he knows that there is a kitchen product area in the supermarket where he can find coffee makers. Second, when the categorizing process happens in DM's mind, it is impossible for DM to subjectively categorize all alternatives but at the same time ignores the existence of some alternatives. Lastly, no matter categories are objective or subjective, the purpose of categorization often is to simplify choice problems. Then the number of categories in many cases should be small enough such that DM is able to pay attention to all categories.

In the following we compare RCRC with other models with stochastic consideration set. Brady and Rehbeck (2016) (BR hereafter) propose the *random conditional choice set rule* (RCCSR) $p_{\succ,\pi}$. Here $\succ$ is DM's strict preference relation over all alternatives. Facing a menu DM first randomly pays attention to a subset of the menu, then chooses the best alternative in the subset. BR assume that DM's attention paid to all possible menus follows a global probability distribution $\pi : \mathcal{D} \to (0,1)$, and DM pays attention to a subset B of a menu A with a probability of $\frac{\pi(B)}{\sum_{C\subseteq A}\pi(C)}$. RCSR is a special case of RCCSR if we let $\pi(A) = \prod_{b\notin A}\gamma(b)\prod_{a\in A}\gamma(a)$ for all $A \in \mathcal{D}$. The representation of RCCSR is as follows.

BR(2016): $p_{\succ,\pi}(a,A) = \frac{\sum_{B\in A_{a\succ}}\pi(B)}{\sum_{C\subseteq A}\pi(C)}$ where $A_{a\succ} \equiv \{B \subseteq A : a = \arg\max_{\succ} B\}$.

RCRC and RCCSR do not nest each other. To prove it we first show that RCCSR does not satisfy I-IIA through the following example.

Example 2. *$X = \{a, b, c\}$, $\mathcal{D} = 2^X$. Suppose there is a RCCSR $p_{\succ,\pi}$ such that $a \succ b \succ c$, $\pi(\{a\}) = \pi(\{b\}) = \pi(\{c\}) = .1$, $\pi(\{a, b\}) = \pi(\{b, c\}) = \pi(\{a, c\}) = .1$, and $\pi(\{a, b, c\}) = \pi(\emptyset) = .2$. It is easy to check that $p_{\succ,\pi}(a, \{a\}) = 1/3$ and $p_{\succ,\pi}(a, \{a, b\}) = p_{\succ,\pi}(a, \{a, c\}) = 2/5$.*

By Revealed relation we have $b \succsim^* a$ and $c \succsim^* a$. Similarly we have $a \succsim^* b$ and $c \succsim^* b$, and $a \succsim^* c$ and $b \succsim^* c$. So $a \sim^* b \sim^* c$. However, $\frac{p_{\succ,\pi}(a,\{a,c\})}{p_{\succ,\pi}(c,\{a,c\})} = 2$ but $\frac{p_{\succ,\pi}(a,\{a,b,c\})}{p_{\succ,\pi}(c,\{a,b,c\})} = 5$. So $p_{\succ,\pi}$ violates I-IIA(1).

RCCSR is not a random utility model and neither satisfies regularity, but its choice probability of the default a^* satisfies *default monotonicity*: $B \subseteq A \Rightarrow p(a^*, B) > p(a^*, A)$. However, in section 6.1 we show that RCRC can explain choice overload. So RCRC does not satisfy default monotonicity. The existence of the default a^* is crucial to calibrate the attention distribution π in RCCSR. However, we show that we can exclude the default from RCRC.

Aguiar (2015) proposes a *fuzzy attention model* (FAM) $p_{\succ,\varphi}$ which also generalizes RCSR. Specifically, $\succ$ is DM's strict preference relation over all alternatives, and $\varphi : 2^X \to [0, 1]$ is a fuzzy measure which Aguiar calls attention capacity. $\varphi(A)$ measures DM's awareness of a menu A, and $\varphi(A) \leq \varphi(B)$ if $A \subseteq B$, that is, bigger sets are more likely to be noticed. The probability that an alternative a is chosen from a menu A is the marginal contribution of a to the awareness of the set $U_A(a) \cup \{a\}$ where $U_A(a)$ is the upper contour set of a in A.

Aguiar(2015): $p_{\succ,\varphi}(a, A) = \varphi(U_A(a) \cup \{a\}) - \varphi(U_A(a))$ where $U_A(a) \equiv \{b \in A : b \succ a\}$.

In FAM DM can pay complex attention, such as substitutable attention and complementary attention, to alternatives. However, this also increases the difficulty of identifying FAM. FAM satisfies default monotonicity, so it does not nest RCRC. On the other hand, RCRC neither nests FAM.

Kovach (2016) extends RCSR by assuming that there exists an observable status quo in each menu such that DM pays attention to it for sure, and DM randomly pays attention to other alternatives only if they are sufficiently better than the status quo. He uses different datasets than ours, and his model and RCRC do not nest each other.

7.2 Variants of Luce Model

We have shown that RCRC subsumes the Luce model as a special case. Now we compare RCRC with some variants of the Luce model in the literature.

Gul et al. (2014) (GNP hereafter) propose the *attribute rule* in which each alternative is characterized by a set of endogenous attributes and DM follows a two-stage procedure to make a choice: DM first randomly chooses an attribute, then randomly chooses an alternative that has the attribute. At the first stage each attribute has a weight such that DM's choice follows a Luce-type formula; at the second stage each alternative has an intensity of the chosen attribute such that DM's choice also follows a Luce-type formula. The representation of the attribute rule is as follows.

GNP(2014): $p_{v,w,\eta}(a, A) = \sum_{x \in v_a} \frac{w_x}{w(v(A))} \frac{\eta_a^x}{\eta^x(A)}$.[10]

In the attribute rule if we interpret each attribute as a category, then each alternative can belong to multiple categories. Then RCRC is different from the attribute rule in that in RCRC DM considers categories sequentially, while in the attribute rule there is no such sequential order. The attribute rule is a random utility model, so it does not nest RCRC. Through the following example we show that RCRC neither nests the attribute rule.

Example 3. *$X = \{a, b, c\}$, $\mathcal{D} = 2^X$. Suppose there is an attribute rule $p_{v,w,\eta}$ such that $v_a = v_b = \{x, y\}$, $v_c = \{x\}$, $w_x = w_y = 1$, $(\eta_a^x, \eta_a^y) = (1, 2)$, $(\eta_b^x, \eta_b^y) = (2, 1)$ and $\eta_c^x = 1$. Then $p_{v,w,\eta}$ can be represented by the following table.*

$p_{v,w,\eta}$	$\{a,b,c\}$	$\{a,b\}$	$\{a,c\}$	$\{b,c\}$	$\{a\}$	$\{b\}$	$\{c\}$
a	$\frac{11}{24}$	$\frac{1}{2}$	$\frac{3}{4}$		1		
b	$\frac{10}{24}$	$\frac{1}{2}$		$\frac{5}{6}$		1	
c	$\frac{3}{24}$		$\frac{1}{4}$	$\frac{1}{6}$			1

Suppose $p_{v,w,\eta}$ is a RCRC. Since removing any alternative from $\{a, b, c\}$ changes the choice probabilities of the other two alternatives, a, b, c must be in the same category. However, $\frac{p_{v,w,\eta}(a,\{a,b,c\})}{p_{v,w,\eta}(b,\{a,b,c\})} \neq \frac{p_{v,w,\eta}(a,\{a,b\})}{p_{v,w,\eta}(b,\{a,b\})}$, which means that $p_{v,w,\eta}$ violates I-IIA. So $p_{v,w,\eta}$ is not a RCRC.

[10] v_a is the set of attributes of a, $v(A) \equiv \cup_{a \in A} v_a$ is the set of all attributes in A, w_x is the weight of attribute x, η_a^x is the intensity of attribute x at a, and $\eta^x(A) \equiv \sum_{a \in A} \eta_a^x$ is the total intensity of attribute x in A.

Echenique et al. (2014) (EST hereafter) propose the *perception-adjusted Luce model* (PALM) $p_{\hat{\succsim},\hat{u}}$ in which DM follows a weak perception order $\hat{\succsim}$ to consider the alternatives in a menu, and once an alternative is considered, it is chosen with a probability same as in the Luce model.

EST(2014): $p_{\hat{\succsim},\hat{u}}(a, A) = \mu(a, A)\prod_{A_\alpha \in \tilde{A}: A_\alpha \hat{\succ} a}(1 - \sum_{b\in A_\alpha} \mu(b, A))$ where $\hat{\succsim}$ is a weak order on X, $\hat{u} : X \cup \mathcal{D} \to \mathbf{R}_{++}$ is a function, and $\mu(a, A) = \frac{\hat{u}(a)}{\sum_{b\in A}\hat{u}(b)+\hat{u}(A)}$.

The representation of PALM looks similar to that of RCRC, but they have a crucial difference: in PALM once an alternative is perceived, its choice probability depends on all alternatives in the menu ($\mu(a, A) = \frac{\hat{u}(a)}{\sum_{b\in A}\hat{u}(b)+\hat{u}(A)}$); but this implicitly implies that all alternatives in the menu are noticed, which kind of contradicts their idea of sequential perception; however, in RCRC once DM considers a category, DM's choice only depends the alternatives in the category ($\gamma(a, A) = \frac{u(a)}{\sum_{b\in A_a} u(b)+u(A_a)}$).

PALM and RCRC do not nest each other. In particular, in RCRC the default a^* can be chosen with zero probability in some menu, while in PALM a^* is always chosen with a positive probability. In the following example we show that RCRC does not nest PALM.

Example 4. $X = \{a, b, c\}$, $\mathcal{D} = 2^X$. *Suppose there is a PALM* $p_{\hat{\succsim},\hat{u}}$ *such that* $a \hat{\succ} b \hat{\succ} c$, $\hat{u}(o) = \hat{u}(A) = 1$ *for all* $o \in X$ *and all* $A \in \mathcal{D}$. *Then* $p_{\hat{\succsim},\hat{u}}$ *can be represented by the following table.*

$p_{\hat{\succsim},\hat{u}}$	$\{a,b,c\}$	$\{a,b\}$	$\{a,c\}$	$\{b,c\}$	$\{a\}$	$\{b\}$	$\{c\}$
a	$\frac{1}{4}$	$\frac{1}{3}$	$\frac{1}{3}$		$\frac{1}{2}$		
b	$\frac{3}{4}\cdot\frac{1}{4}$	$\frac{2}{3}\cdot\frac{1}{3}$		$\frac{1}{3}$		$\frac{1}{2}$	
c	$(\frac{3}{4})^2\cdot\frac{1}{4}$		$\frac{2}{3}\cdot\frac{1}{3}$	$\frac{2}{3}\cdot\frac{1}{3}$			$\frac{1}{2}$
a^*	$(\frac{3}{4})^3$	$(\frac{2}{3})^2$	$(\frac{2}{3})^2$	$(\frac{2}{3})^2$	$\frac{1}{2}$	$\frac{1}{2}$	$\frac{1}{2}$

Suppose $p_{\hat{\succsim},\hat{u}}$ *is a RCRC. Since removing any alternative from* $\{a, b, c\}$ *will change the choice probabilities of the other two alternatives,* a, b, c *must be in the same category. However,* $\frac{p_{\hat{\succsim},\hat{u}}(a,\{a,b,c\})}{p_{\hat{\succsim},\hat{u}}(b,\{a,b,c\})} \neq \frac{p_{\hat{\succsim},\hat{u}}(a,\{a,b\})}{p_{\hat{\succsim},\hat{u}}(b,\{a,b\})}$, *which means that* $p_{\hat{\succsim},\hat{u}}$ *violates I-IIA. So* $p_{\hat{\succsim},\hat{u}}$ *is not a RCRC.*

7.3 Models with Random Search

Aguiar et al. (2015) propose the *general satisficing model* (GSM) in which DM has a fixed utility function and a fixed satisficing level but uses random search order. Random search order results in stochastic choice. Aguiar et al. assume that if DM does not find an alternative passing the satisficing level in a menu, he chooses the best alternative in the menu. So there is no default alternative a^* in their setup. Without any assumption on the distribution of search order, GSM can explain any dataset. By further assuming that any alternative will be searched first with a positive probability in each menu, Aguiar et al. define the *full support satisficing model* (FSSM) and axiomatize it.

RCRC is different from GSM in that the search order in RCRC is deterministic but DM randomly stops his search. RCRC and FSSM do not nest each other. In particular, in FSSM it cannot happen that an alternative is chosen with zero probability in a menu but is chosen with a positive probability in another menu. However it can happen in RCRC. In the following example we show that RCRC does not nest FSSM.

Example 5. $X = \{a, b, c\}$, $\mathcal{D} = 2^X$. *Suppose there is a FSSM p in which all alternatives pass the satisficing level. The distribution of search order in all menus is as follows.*[11]

search order	a, b, c	a, c ,b	b, a, c	b, c, a	c, a, b	c, b, a
probability	1/4	1/4	1/8	1/8	1/8	1/8

Then p can be represented as follows.

p	$\{a,b,c\}$	$\{a,b\}$	$\{a,c\}$	$\{b,c\}$	$\{a\}$	$\{b\}$	$\{c\}$
a	$\frac{1}{2}$	$\frac{5}{8}$	$\frac{5}{8}$		1		
b	$\frac{1}{4}$	$\frac{3}{8}$		$\frac{1}{2}$		1	
c	$\frac{1}{4}$		$\frac{3}{8}$	$\frac{1}{2}$			1

Suppose p is a RCRC. Since removing any alternative from $\{a, b, c\}$ will change the choice probabilities of the other two alternatives, a, b, c must be in the same category. However, $\frac{p(a,\{a,b,c\})}{p(b,\{a,b,c\})} \neq \frac{p(a,\{a,b\})}{p(b,\{a,b\})}$, which means that p violates I-IIA. So p is not a RCRC.

Ravid (2015) proposes another random search model called *focus, then compare* (FTC). In FTC, DM first randomly draws an alternative from a menu as a focal option, then compares the focal option sequentially with other alternatives in the menu. The winner of every binary comparison is random. The focal option is chosen if it wins in all

[11]When DM follows a search order to consider a menu, he will skip the alternatives that not in the menu.

binary comparisons. Otherwise DM randomly draws another focal option with replacement and makes the sequential comparisons again. There is no default alternative a^*. The representation of FTC is as follows.

Ravid(2015): $p_\pi(a, A) = \dfrac{\prod_{b \in A} \pi(a, b)}{\sum_{c \in A} \left[\prod_{b \in A} \pi(c, b)\right]}$ where $\pi(a, b) > 0$ is the probability that a wins in the comparison between a and b when a is focal.

FTC does not nest RCRC because it can only explain datasets with full support. RCRC neither nests FTC, which can be seen through the following example.

Example 6. $X = \{a, b, c\}$, $\mathcal{D} = 2^X$. *Suppose there is a FTC* p_π *in which* π *is shown by the following table. In the table the rows are focal alternatives.*

π	a	b	c
a	1	.5	.5
b	.6	1	.6
c	.7	.7	1

Then p_π *can be shown as follows.*

p_π	$\{a, b, c\}$	$\{a, b\}$	$\{a, c\}$	$\{b, c\}$	$\{a\}$	$\{b\}$	$\{c\}$
a	.22	.45	.42		1		
b	.33	.55		.46		1	
c	.45		.58	.54			1

Suppose p_π *is a RCRC. Since removing any alternative from* $\{a, b, c\}$ *will change the choice probabilities of the other two alternatives,* a, b, c *must be in the same category. However,* $\frac{p_\pi(a,\{a,b,c\})}{p_\pi(b,\{a,b,c\})} \neq \frac{p_\pi(a,\{a,b\})}{p_\pi(b,\{a,b\})}$, *which means that* p_π *violates I-IIA. So* p_π *is not a RCRC.*

7.4 Models with Deliberately Stochastic Choice

Fudenberg et al. (2015a) and Cerreia-Vioglio et al. (2015) both propose models to explain why DM may deliberately randomize his choice. In particular, Fudenberg et al. assume that DM optimally chooses an lottery over a menu to maximize an expected utility function perturbed by a cost function. In the setup of Cerreia-Vioglio et al. each alternative is a lottery such that choosing a lottery over a menu is equivalent to choosing a compound lottery. If DM has a preference relation over compound lotteries, he will deliberately choose the most preferred compound lottery.

RCRC is different from these two models in that RCRC does not rationalize DM's stochastic choice. RCRC belongs to the class of models with bounded rationality.

8 Related Literature

In the choice-theoretical literature several papers have studied the existence of categories. Barbos (2010) proposes a model of reference-dependent choice in the presence of exogenous categories to show how the existence of categories can distort DM's choice and explains context effects. Maltz (2015) studies the endowment bias in the presence of exogenous categories. Manzini and Mariotti (2012) characterize a two-stage model in the presence of subjective categories. At first stage DM chooses undominated categories according to a preference relation over categories, and at stage two DM chooses the best alternative in all undominated categories according to a preference relation over alternatives. In the model categories may not be disjoint. Furtado et al. (2015) consider a similar model as MM's except that DM does not have a preference relation over categories and chooses the union of the best alternative in each category.

In the stochastic choice literature Echenique and Saito (2015) generalize the Luce model by introducing the dominance relation between alternatives such that an alternative is chosen with zero probability in a menu if it is dominated by another alternative in the menu. Ahn et al. (2015) use a modification of the well-known *path independence* to characterize the Luce model in which the primitive is the average of stochastic choices from each menu.[12] Lu (2015) proposes an informational model of stochastic choice in which stochasticity comes from the unobservable private information of DM. Lu proposes a method to identify DM's private information. Fudenberg et al. (2015b) propose a sequential sampling model to explain the joint distribution of DM's choices and decision times.

[12]To be precise, they assume that each alternative a is an lottery and the primitive p^* is a compound lottery for each menu A such that $p^*(A) = \sum_{a \in A} p(a, A)a$.

A Proofs of Theorem 1 and Theorem 2

Proof of Theorem 1

• (Sufficiency) Suppose a random choice rule p satisfies Transitivity, Single-category Full Support and I-IIA, we prove that it is a gRCRC.

First, Transitivity implies that the revealed relation $\succsim^*$ is a weak order. So $\sim^*$ can induce a partition of X, denoted by $\{X_1, X_2, \ldots, X_{k_X}\}$.

For all X_i and all nonempty $A_i \subseteq X_i$, define

$$q(a, A_i) \equiv \frac{p(a, A_i)}{\sum_{b \in A_i} p(b, A_i)}.$$

By Single-category Full Support $\sum_{b \in A_i} p(b, A_i) > 0$, so q is well-defined. For all $\{a, b\} \subseteq A_i \subseteq X_i$, I-IIA(1) implies that

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

So q satisfies Luce's IIA on X_i. Since $\sum_{b \in A_i} q(b, A_i) = 1$, by Luce's (1959) theorem, there exists $u : X \to \mathbf{R}_{++}$ such that $q(a, A_i) = \frac{u(a)}{\sum_{b \in A_i} u(b)}$. Define $u(A_i) \equiv \sum_{b \in A_i} u(b)(\frac{1}{\sum_{b \in A_i} p(b, A_i)} - 1)$ which implies that $\frac{p(a^*, A_i)}{\sum_{b \in A_i} p(b, A_i)} = \frac{u(A_i)}{\sum_{b \in A_i} u(b)}$. Then,

$$p(A_i, A_i) = \frac{\sum_{b \in A_i} u(b)}{\sum_{b \in A_i} u(b) + u(A_i)} \text{ and } p(a, A_i) = \frac{u(a)}{\sum_{b \in A_i} u(b) + u(A_i)}.$$

Now we prove the sufficiency part by induction. For any nonempty menu $A \in \mathcal{D}$ and its categories $\{A_1, A_2, \ldots, A_{k_A}\}$, consider any $a \in A_1$. It is obvious that $p(a, A) = p(a, A_1)$. Otherwise there exist $b \in A_i$ for some $i > 1$ and some $A_1 \subseteq A' \subseteq A$ such that $p(a, A') \neq p(a, A' \backslash \{b\})$. However, this implies that $b \succsim^* a$ which contradicts $a \succ^* b$. Hence,

$$p(a, A) = p(a, A_1) = \frac{u(a)}{\sum_{b \in A_1} u(b) + u(A_1)}.$$

Suppose the theorem holds for all A_k such that $k < i$, then it implies that

$$1 - \sum_{1 \le k < i} p(A_k, A) = \prod_{1 \le k < i} [1 - \pi(A_k)], \quad \text{where} \quad \pi(A_k) = \frac{\sum_{b \in A_k} u(b)}{\sum_{b \in A_k} u(b) + u(A_k)}.$$

Now we prove that the theorem also holds for A_i.

If $u(A_k) > 0$ for all $k < i$, then $1 - \sum_{1 \le k < i} p(A_k, A) > 0$. By a similar argument as before, $p(A_j, A) = p(A_j, A_1 \cup \cdots \cup A_{i-1} \cup A_i)$ for all A_j such that $j \le i$. This implies that $p(A_i, A_1 \cup \cdots \cup A_{i-1} \cup A_i) > 0$. So there exists $b \in A_i$ such that $p(b, A_1 \cup \cdots \cup A_{i-1} \cup A_i) > 0$.

By Single-category Full Support, $p(a, A_i) > 0$ for all $a \in A_i$. Since I-IIA(1) requires $\frac{p(a,A_i)}{p(b,A_i)} = \frac{p(a,A_1\cup\cdots\cup A_{i-1}\cup A_i)}{p(b,A_1\cup\cdots\cup A_{i-1}\cup A_i)}$, $p(a, A_1 \cup \cdots \cup A_{i-1} \cup A_i) > 0$ for all $a \in A_i$. So the fractions appearing in the following equations are well-defined.

For all $a \in A_i$, we have

$$
\begin{aligned}
\frac{p(a, A)}{1 - \sum_{1\le k<i} p(A_k, A)} &= \frac{p(a, A_1 \cup \cdots \cup A_{i-1} \cup A_i)}{1 - \sum_{1\le k<i} p(A_k, A_1 \cup \cdots \cup A_{i-1} \cup A_i)} \\
&= \frac{p(a, A_1 \cup \cdots \cup A_{i-1} \cup A_i)}{p(a^*, A_1 \cup \cdots \cup A_{i-1} \cup A_i) + \sum_{b\in A_i} p(b, A_1 \cup \cdots \cup A_{i-1} \cup A_i)} \\
&= \frac{1}{\frac{p(a^*, A_1 \cup \cdots \cup A_{i-1} \cup A_i)}{p(a, A_1 \cup \cdots \cup A_{i-1} \cup A_i)} + \sum_{b\in A_i} \frac{p(b, A_1 \cup \cdots \cup A_{i-1} \cup A_i)}{p(a, A_1 \cup \cdots \cup A_{i-1} \cup A_i)}} \\
&= \frac{1}{\frac{p(a^*, A_i)}{p(a, A_i)} + \sum_{b\in A_i} \frac{p(b, A_i)}{p(a, A_i)}} \\
&= \frac{p(a, A_i)}{p(a^*, A_i) + \sum_{b\in A_i} p(b, A_i)} \\
&= p(a, A_i).
\end{aligned}
$$

The fourth equality follows I-IIA(1) and I-IIA(2). Hence,

$$
p(a, A) = p(a, A_i)\left[1 - \sum_{1\le k<i} p(A_k, A)\right] = \frac{u(a)}{\sum_{b\in A_i} u(b) + u(A_i)} \prod_{1\le k<i} [1 - \pi(A_k)]. \quad (10)
$$

If $u(A_k) = 0$ for some $k < i$, then by the induction assumption $\sum_{j\le k} p(A_j, A) = 1$. So $p(A_i, A) = 0$. Then equation (10) is still correct since $\pi(A_k) = 1$. So by induction $p = p_{\succsim^*,u}$.

Now we prove that $p_{\succsim^*,u}$ is generic. Suppose the contrary, then for some $a \sim^* b$, we have either $u(A_a\backslash\{b\}) = u(b) + u(A_a)$ or $u(A_a\backslash\{a\}) = u(a) + u(A_a)$ for all $\{a, b\} \subseteq A \in \mathcal{D}$. This implies that either $p_{\succsim^*,u}(a, A) = p_{\succsim^*,u}(a, A\backslash\{b\})$ or $p_{\succsim^*,u}(b, A) = p_{\succsim^*,u}(b, A\backslash\{a\})$ for all $A \in \mathcal{D}$. By Revealed relation either $a \succ^* b$ or $b \succ^* a$, which contradicts $a \sim^* b$. So $p_{\succsim^*,u}$ is generic.

- (Necessity) To prove that a generic $p_{\succsim,u}$ satisfies the axioms, we first prove that $\succsim = \succsim^*$. That is, the weak order $\succsim$ can be perfectly identified. This implies that $p_{\succsim,u}$ satisfies Transitivity.

If $a \succ b$, then it is obvious that $p_{\succsim,u}(a, A) = p_{\succsim,u}(a, A\backslash\{b\})$ for all $A \in \mathcal{D}$. By Revealed relation $a \succ^* b$. If $a \sim b$, since $p_{\succsim,u}$ is generic, we must have $p_{\succsim,u}(a, A) \neq p_{\succsim,u}(a, A\backslash\{b\})$

for some $\{a,b\} \subseteq A \in \mathcal{D}$ and $p_{\succsim,u}(b,B) \neq p_{\succsim,u}(b,B\backslash\{a\})$ for some $\{a,b\} \subseteq B \in \mathcal{D}$. By Revealed relation $a \succsim^* b$ and $b \succsim^* a$. So $a \sim^* b$. This implies that $\succsim = \succsim^*$.

Since $u(a) > 0$ for all $a \in X$, $p_{\succsim,u}$ must satisfy Single-category Full Support. It is obvious that $p_{\succsim,u}$ satisfies I-IIA.

- (Uniqueness) For any generic $p_{\succsim',u'}$ such that $p_{\succsim',u'} = p_{\succsim,u}$, since $\succsim = \succsim^* = \succsim'$, $\succsim$ is unique. Then the set of categories is also unique. By Luce (1959), there exists some $\beta_{X_i} > 0$ such that $u'(a) = \beta_{X_i} u(a)$ and $u'(A) = \beta_{X_i} u(A)$ for all $X_i \in \tilde{X}$, all $a \in X_i$ and all $A \in \mathcal{D}\backslash\{\emptyset\}$ such that $A \subseteq X_i$. So u is unique up to category-wise positive multiplication.

Proof of Theorem 2

- (Sufficiency) Suppose a random choice rule p satisfies Weak Order, Highest-category Full Support, I-IIA-D and L-Independence. We prove it is a rich and strongly generic RCRC-D

Let $\succsim^*$ be the revealed relation. Weak Order implies that $\succsim^*$ is a weak order. Let $\tilde{X}$ and $\tilde{A}$ be the sets of categories induced by $\sim^*$ in X and any menu A respectively.

For all $X_i \in \tilde{X}$ and all $A_i \subseteq X_i$, we have $\sum_{b\in A_i} p(b,A_i) = 1$ and $p(b,A_i) > 0$ for all $b \in A_i$ which is implied by Highest-category Full Support. Since I-IIA-D(1) implies that p satisfies Luce's IIA on X_i, there exists $u : X \to \mathbf{R}_{++}$ such that $p(a,A_i) = \frac{u(a)}{\sum_{b\in A_i} u(b)}$.

We prove the sufficiency part by induction. For any $A \in \mathcal{A}\backslash\{\emptyset\}$, if $A_2 = \emptyset$, then for all $a \in A_1$, $p(a,A) = p(a,A_1) = \frac{u(a)}{\sum_{b\in A_1} u(b)}$.

If $A_2 \neq \emptyset$, for all $a \in A_1$, I-IIA-D(1) implies that $\frac{p(a,A)}{\sum_{b\in A_1} p(b,A)} = \frac{p(a,A_1)}{\sum_{b\in A_1} p(b,A_1)} = p(a,A_1)$, where $p(b,A)p(b,A_1) > 0$ for all $b \in A_1$ is implied by Highest-category Full Support. So $p(a,A) = p(a,A_1)\sum_{b\in A_1} p(b,A)$. We define

$$u(A_1,A) \equiv \sum_{b\in A_1} u(b)\frac{1-p(A_1,A)}{p(A_1,A)}. \tag{11}$$

We prove that $u(A_1,A)$ is independent of A. That is, for any A,B such that $A_1 = B_1$ and $A_2 \neq \emptyset \neq B_2$, we prove that $u(A_1,A) = u(A_1,B)$. However, it is immediately implied by L-Independence since $p(A_1,A) = p(B_1,B)$. So we can define $u(A_1) \equiv u(A_1,A)$. Then equation (11) implies that

$$p(A_1,A) = \frac{\sum_{b\in A_1} u(b)}{\sum_{b\in A_1} u(b) + u(A_1)}.$$

Hence,

$$p(a,A) = p(a,A_1)p(A_1,A) = \frac{u(a)}{\sum_{b\in A_1} u(b) + u(A_1)}.$$

So we finish the proof for A_1. Suppose the representation of RCRC-D holds for all A_k such that $k < i$, then it implies that

$$1 - \sum_{1 \le k < i} p(A_k, A) = \prod_{1 \le k < i} [1 - \pi(A_k)], \quad \text{where} \quad \pi(A_k) = \frac{\sum_{b \in A_k} u(b)}{\sum_{b \in A_k} u(b) + u(A_k)}.$$

If $A_i = \emptyset$, there is nothing to prove. So we prove that the representation of RCRC-D also holds for nonempty A_i.

If $1 - \sum_{1 \le k < i} p(A_k, A) = 0$, then $p(A_i, A) = 0$. So it obvious that for all $a \in A_i$,

$$p(a, A) = \frac{u(a)}{\sum_{b \in A_i} u(b)} \prod_{1 \le k < i} [1 - \pi(A_k)].$$

If $1 - \sum_{1 \le k < i} p(A_k, A) > 0$, since A_i is the highest category in the remaining alternatives of A, Highest-category Full Support implies that $p(a, A) > 0$ for all $a \in A_i$. Similarly as before I-IIA-D(1) implies that $p(a, A) = p(a, A_i)p(A_i, A)$.

If $A_{i+1} = \emptyset$, then $p(A_i, A) = 1 - \sum_{1 \le k < i} p(A_k, A)$. So for all $a \in A_i$,

$$p(a, A) = p(a, A_i)p(A_i, A) = \frac{u(a)}{\sum_{b \in A_i} u(b)} \prod_{1 \le k < i} [1 - \pi(A_k)].$$

If $A_{i+1} \neq \emptyset$, we define

$$u(A_i, A) \equiv \sum_{b \in A_i} u(b) \frac{1 - \sum_{1 \le k \le i} p(A_k, A)}{p(A_i, A)}. \tag{12}$$

We prove that $u(A_i, A)$ is independent of A. For any $d \in A_{i+1}$, by *L-Independence* $p(A_k, A) = p(A_k, A_1 \cup \cdots \cup A_i \cup \{d\})$ for all $k \le i$. Therefore,

$$\begin{aligned} u(A_i, A) &= \sum_{b \in A_i} u(b) \frac{1 - \sum_{1 \le k \le i} p(A_k, A_1 \cup \cdots \cup A_i \cup \{d\})}{p(A_i, A_1 \cup \cdots \cup A_i \cup \{d\})} \\ &= \sum_{b \in A_i} u(b) \frac{p(d, A_1 \cup \cdots \cup A_i \cup \{d\})}{p(A_i, A_1 \cup \cdots \cup A_i \cup \{d\})} \\ &= \sum_{b \in A_i} u(b) \frac{p(d, A_i \cup \{d\})}{p(A_i, A_i \cup \{d\})}, \end{aligned}$$

where the last equality follows $p(b, A_i \cup \{d\})p(b, A_1 \cup \cdots \cup A_i \cup \{d\}) > 0$ for all $b \in A_i$ and I-IIA-D(2).

For any menu B such that $A_i = B_j$ for some j and $B_{j+1} \neq \emptyset$, L-Independence implies that $p(A_i, A_i \cup \{d\}) = p(A_i, A_i \cup \{e\})$ for any $e \in B_{j+1}$, and $p(d, A_i \cup \{d\}) = 1 - p(A_i, A_i \cup \{d\}) = 1 - p(A_i, A_i \cup \{e\}) = p(e, A_i \cup \{e\})$. So $u(A_i, A) = u(A_i, B)$. Hence we define $u(A_i) \equiv u(A_i, A)$. Then equation (12) implies that

$$p(A_i, A) = \frac{\sum_{b \in A_i} u(b)}{\sum_{b \in A_i} u(b) + u(A_i)} \left[1 - \sum_{1 \le k < i} p(A_k, A)\right].$$

Therefore,

$$p(a, A) = p(a, A_i)p(A_i, A) = \frac{u(a)}{\sum_{b \in A_i} u(b) + u(A_i)} \prod_{1 \leq k < i} [1 - \pi(A_k)] .$$

So we finish the proof that A_i also satisfies the representation of RCRC-D. By induction we can prove that all categories in the menu A satisfy the representation of RCRC-D.

Now we prove that $\bar{p}_{\succsim^*, u}$ is rich. For any $i > 1$, by the definition of Revealed relation-D, to identify that $a \succ^* b$ for any $a \in X_1$ and any $b \in X_i$, there must exist $c \in X_i \backslash \{b\}$ such that $c \sim^* b$, which implies that $|X_i| \geq 2$.

$\bar{p}_{\succsim^*, u}$ is also strongly generic. For any $a \sim^* b$, by Revealed relation-D either $a \sim^1 b$ or there exist $c^1, \cdots, c^n$ such that $a \sim^1 c^1 \sim^1 \cdots \sim^1 c^n \sim^1 b$. So either $a \wedge b$ or $a \wedge c^1 \wedge \cdots \wedge c^n \wedge b$.

- (Necessity) For a rich and strongly generic $\bar{p}_{\succsim, u}$, we prove that the revealed relation $\succsim^*$ is a weak order. To prove it we prove that $\succsim^* = \succsim$.

Since $\bar{p}_{\succsim, u}$ is rich, for all $a \in X_i$ such that $i > 1$ there exists $b \in X_i$ such that $a \sim b$. Since $\bar{p}_{(u, \succsim)}$ is strongly generic, either $a \wedge b$ or there exist $c^1, \cdots, c^n$ such that $a \wedge c^1 \wedge \cdots \wedge c^n \wedge b$. By Revealed relation-D, these imply that either $a \sim^1 b$ or $a \sim^1 c^1 \sim^1 \cdots \sim^1 c^n \sim^1 b$. So $a \sim^* b$. If $|X_1| > 1$, for all $a, b \in X_1$ such that $a \sim b$ we can prove $a \sim^* b$ in the same way.

For all $a, b \in X$ such that $a \succ b$, it must be that $b \in X_i$ for some $i > 1$. So there exists $c \in X_i$ such that $c \sim b$ which implies $c \sim^* b$. Since $\mathcal{D}$ contains all tripleton sets, $\{a, b, c\} \in \mathcal{D}$. Then it is easy to check that $\bar{p}_{\succsim, u}(a, A \backslash \{b\}) = \bar{p}_{\succsim, u}(a, A)$ for all $\{a, b\} \subseteq A \in \mathcal{D}$ such that there exists $d \in A \backslash \{a, b\}$ and $b \sim d$. So by Revealed relation-D, $a \succ^2 b$, which implies $a \succ^* b$. So $\succsim^* = \succsim$.

It is easy to check that $\bar{p}_{\succsim, u}$ satisfies Highest-category Full Support, I-IIA-D and L-Independence.

- (Uniqueness) Since we already prove that $\succsim = \succsim^*$, $\succsim$ is unique. For the same reason in Theorem 1, u is unique up to category-wise positive multiplication.

B Similarity Effect and Compromise Effect

Proposition 1. *(1) If $a \sim c \succ b$, then $p_{\succsim,u}$ exhibits similarity effect if and only if $u(\{a,c\}) > u(\{a\})$, and exhibits compromise effect if and only if $\frac{u(a)}{u(\{a,c\})} < \frac{u(b)}{u(b)+u(\{b\})} \leq \frac{u(a)}{u(\{a\})}$.*

(2) If $b \succ a \sim c$, then $p_{\succsim,u}$ exhibits similarity effect if $u(\{a,c\}) + u(b) > u(\{a\})$, and exhibits compromise effect if and only if $\frac{u(a)}{u(a)+u(c)+u(\{a,c\})} < \frac{u(b)}{u(\{b\})} \leq \frac{u(a)}{u(a)+u(\{a\})}$.

Proof. (1) If $a \sim c \succ b$, then

$$\frac{p_{\succsim,u}(a,\{a,b\})}{p_{\succsim,u}(b,\{a,b\})} = \frac{\frac{u(a)}{u(a)+u(\{a\})}}{\frac{u(b)}{u(b)+u(\{b\})}\frac{u(\{a\})}{u(a)+u(\{a\})}} = \frac{u(a)}{u(\{a\})}\frac{u(b)+u(\{b\})}{u(b)},$$

and

$$\frac{p_{\succsim,u}(a,\{a,b,c\})}{p_{\succsim,u}(b,\{a,b,c\})} = \frac{\frac{u(a)}{u(a)+u(c)+u(\{a,c\})}}{\frac{u(b)}{u(b)+u(\{b\})}\frac{u(\{a,c\})}{u(a)+u(c)+u(\{a,c\})}} = \frac{u(a)}{u(\{a,c\})}\frac{u(b)+u(\{b\})}{u(b)}.$$

So $\frac{p_{\succsim,u}(a,\{a,b,c\})}{p_{\succsim,u}(b,\{a,b,c\})} < \frac{p(a,\{a,b\})}{p(b,\{a,b\})}$ if and only if $u(\{a,c\}) > u(\{a\})$, and $\frac{p_{\succsim,u}(a,\{a,b,c\})}{p_{\succsim,u}(b,\{a,b,c\})} < 1 \leq \frac{p_{\succsim,u}(a,\{a,b\})}{p_{\succsim,u}(b,\{a,b\})}$ if and only if $\frac{u(a)}{u(\{a,c\})} < \frac{u(b)}{u(b)+u(\{b\})} \leq \frac{u(a)}{u(\{a\})}$.

(2) If $b \succ a \sim c$, then

$$\frac{p_{\succsim,u}(a,\{a,b\})}{p_{\succsim,u}(b,\{a,b\})} = \frac{\frac{u(\{b\})}{u(b)+u(\{b\})}\frac{u(a)}{u(a)+u(\{a\})}}{\frac{u(b)}{u(b)+u(\{b\})}} = \frac{u(\{b\})}{u(b)}\frac{u(a)}{u(a)+u(\{a\})},$$

and

$$\frac{p_{\succsim,u}(a,\{a,b,c\})}{p_{\succsim,u}(b,\{a,b,c\})} = \frac{\frac{u(\{b\})}{u(b)+u(\{b\})}\frac{u(a)}{u(a)+u(c)+u(\{a,c\})}}{\frac{u(b)}{u(b)+u(\{b\})}} = \frac{u(\{b\})}{u(b)}\frac{u(a)}{u(a)+u(c)+u(\{a,c\})}.$$

So $\frac{p_{\succsim,u}(a,\{a,b,c\})}{p_{\succsim,u}(b,\{a,b,c\})} < \frac{p(a,\{a,b\})}{p(b,\{a,b\})}$ if and only if $u(\{a,c\}) + u(b) > u(\{a\})$, and $\frac{p_{\succsim,u}(a,\{a,b,c\})}{p_{\succsim,u}(b,\{a,b,c\})} < 1 \leq \frac{p_{\succsim,u}(a,\{a,b\})}{p_{\succsim,u}(b,\{a,b\})}$ if and only if $\frac{u(a)}{u(a)+u(c)+u(\{a,c\})} < \frac{u(b)}{u(\{b\})} \leq \frac{u(a)}{u(a)+u(\{a\})}$.

□

References

Agranov, M. and P. Ortoleva (2015): "Stochastic Choice and Preferences for Randomization," *Journal of Political Economy, forthcoming.*

Aguiar, V. (2015): "Stochastic choice and fuzzy attention," *working paper.*

Aguiar, V., M. Boccardi, and M. Dean (2015): "Satisficing and Stochastic Choice," *working paper.*

Ahn, D., F. Echenique, and K. Saito (2015): "On Path Independent Stochastic Choice," *working paper.*

Barbos, A. (2010): "Context effects: A representation of choices from categories," *Journal of Economic Theory*, 145, 1224–1243.

Brady, R. and J. Rehbeck (2016): "Menu-Dependent Stochastic Feasibility," *Econometrica, forthcoming.*

Cerreia-Vioglio, S., D. Dillenberger, P. Ortoleva, and G. Riella (2015): "Deliberately Stochastic," *Manuscript Columbia University.*

Chernev, A. (2003): "When more is less and less is more: The role of ideal point availability and assortment in consumer choice," *Journal of Consumer Research*, 30, 170–183.

Debreu, G. (1960): "Review of 'Individual Choice Behavior: A Theoretical Analysis' by R.D. Luce," *American Economic Review*, 50, 186–188.

Echenique, F. and K. Saito (2015): "General Luce Model," *working paper.*

Echenique, F., K. Saito, and G. Tserenjigmid (2014): "The Perception-Adjusted Luce Model," *working paper.*

Fudenberg, D., R. Iijima, and T. Strzalecki (2015a): "Stochastic choice and revealed perturbed utility," *Econometrica*, 83, 2371–2409.

Fudenberg, D., P. Strack, and T. Strzalecki (2015b): "Stochastic choice and optimal sequential sampling," *working paper.*

Furtado, B., L. Nascimento, and G. Riella (2015): "Rational Choice with Categories," *working paper.*

Gul, F., P. Natenzon, and W. Pesendorfer (2014): "Random Choice as Behavioral Optimization," *Econometrica*, 82, 1873–1912.

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.

Kovach, M. (2016): "Thinking Inside the Box: Status Quo Bias and Stochastic Consideration," *working paper*.

Lu, J. (2015): "Random choice and private information," *working paper*.

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

Maltz, A. (2015): "Rational Choice with Category Bias," *working paper*.

Manzini, P. and M. Mariotti (2011): "Stochastic choice and consideration sets," *working paper*.

——— (2012): "Categorize then choose: Boundedly rational choice and welfare," *Journal of the European Economic Association*, 10, 1141–1165.

——— (2014): "Stochastic choice and consideration sets," *Econometrica*, 82, 1153–1176.

Mogilner, C., T. Rudnick, and S. S. Iyengar (2008): "The mere categorization effect: How the presence of categories increases choosers' perceptions of assortment variety and outcome satisfaction," *Journal of Consumer Research*, 35, 202–215.

Ravid, D. (2015): "Focus, Then Compare," *working paper*.

Rubinstein, A. and Y. Salant (2006): "A model of choice from lists," *Theoretical Economics*, 1, 3–17.

Shah, A. M. and G. Wolford (2007): "Buying behavior as a function of parametric variation of number of choices," *Psychological Science*, 18, 369–370.

Simon, H. A. (1955): "A behavioral model of rational choice," *Quarterly Journal of Economics*, 99–118.

Simonson, I. and A. Tversky (1992): "Choice in context: Tradeoff contrast and extremeness aversion," *Journal of Marketing Research*, 29, 281.