Stochastic Choice with Bounded Processing Capacity

Thierry Marchant

Ghent University

Arunava Sen

Indian Statistical Institute, New Delhi

Abstract

We propose and characterize a class of stochastic decision functions for a decision-maker who has a capacity for processing at most k-alternatives at a time. When faced with a menu containing more than k alternatives, she randomly chooses a sub-menu of size k with uniform probability and selects the best alternative according to a strict ordering $\succ$. For smaller menus, she chooses the best alternative according to $\succ$.

1. Introduction

Choosing from a large menu of alternatives typically imposes a significant burden on the cognitive resources of a decision-maker. In these situations, a two-step procedure is natural: the decision-maker first narrows down the menu using a heuristic and then chooses from the smaller set according to the

☆Authors are listed alphabetically and have contributed equally.

☆☆We would like to thank Abhinash Borah, Federico Echenique, Sean Horan, Yusufcan Masatlioglu, Debasis Mishra and Levent Ülkü for several helpful comments and discussions.

Email addresses: `thierry.marchant@ugent.be` (Thierry Marchant), `asen@isid.ac.in` (Arunava Sen)

standard model of rational choice. In this paper, we propose and axiomatize a particularly simple form of such a procedure.

We assume that the decision-maker has a strict ordering $\succ$ over the set of all alternatives X which she uses to choose alternatives. Choice requires the evaluation and comparison of alternatives and she can process at most k alternatives in a menu where $k > 2$. Consequently, she chooses a $\succ$-maximal alternative for menus of size less than or equal to k. For menus of larger size, she randomly chooses a sub-menu of size k using the uniform distribution and picks the $\succ$-maximal alternative from the sub-menu. We believe that the assumption of the uniform distribution is justified since the decision-maker makes the first-step selection without any reference to the properties of the alternatives. We call this stochastic decision function a *bounded processing capacity rule* or BPCR.

Our main result is a characterization of BPCRs in terms of six axioms on the choice function. The capacity k and the ordering $\succ$ can be uniquely identified from choice functions satisfying the axioms. Virtually all the axioms are variants of axioms that are familiar in the literature on stochastic choice functions. We show that our axioms are independent when there are at least four alternatives.

One interpretation of our choice function is that the decision-maker makes "mistakes" while choosing from large menus.[1] These mistakes are in the form

[1]Mistakes could also be the result of "choice overload" (see Iyengar and Lepper [13], Scheibehenne et al. [21], Aguiar et al. [2]). Indeed, some aspects of the "jam study" or Study 1 in Iyengar and Lepper [13] are consistent with a BPCR. In particular, consumers in both the small and large jam menus sampled approximately the same number of jams.

of departures from the decision-maker's "true" preference. In a BPCR some alternatives that are not $\succ$-maximal are picked with strictly positive probability in large menus. Moreover a BPCR offers an additional, potentially testable insight that is not available in a deterministic model. It chooses a higher-ranked alternative according to $\succ$ with higher probability than a lower-ranked alternative in a menu. Mistakes are made less frequently for less preferred alternatives. Ortoleva [18] and Frick [10] develop the "mistakes approach" in the deterministic context. Frick [10] for example,[2] characterizes a class of rules where the decision-maker has a true preference represented by a utility function v. In every menu, the decision-maker chooses alternatives that are within $\delta \geq 0$ of the v-maximal alternative in the menu. The factor δ increases monotonically in the size of the menu, so that the decision-maker makes larger mistakes in terms of choosing non v-maximal alternatives, the larger is the menu. The rule, however makes no predictions regarding the relative frequency with which various mistakes are made.

Aguiar et al. [1] develop a model of stochastic choice where a boundedly-rational decision-maker makes mistakes. As in Frick [10], the decision-maker's true preferences are represented by a utility function v. She randomly selects a sequence of alternatives in a menu and picks the first "satisficing" alternative in the sequence, i.e. an alternative whose utility exceeds a threshold level v^*.

We note that choice overload is sometimes referred to "as the case where the propensity of not choosing (or the probability of picking a default alternative) increases in larger choice sets" (Aguiar et al. [2]). Our model does not explain choice overload in this sense.

[2] The model in Ortoleva [18] is different and is concerned with choices over lotteries over menus. The underlying idea is however, similar to that in Frick [10].

If no satisficing alternative exists, she selects the v-maximal alternative in the menu. Mistakes occur because she may choose an alternative x where $v(x) \geq v^*$ even though there exists y in the menu and $v(y) > v(x)$ if x precedes y in the chosen sequence. This procedure is not compatible with our notion of bounded capacity. Suppose x and y are the only alternatives that are satisficing in the set X which contains many other alternatives. The decision-maker makes mistakes in the "small" menu $\{x, y\}$; on the other hand, she does not make any mistakes in the "large" menu $X \setminus \{x, y\}$.

Geng and Özbay [11] follow an approach similar to ours. It considers a model where the decision-maker has a capacity of k-alternatives and follows the two-step procedure outlined earlier. However, the decision-maker uses a deterministic heuristic in the first step. One of the goals of the paper is to identify k from choice data for various (deterministic) heuristics. In contrast, the decision-maker in our model uses a simple but random heuristic in the first stage.

Dutta [7] proposes a model where the decision-maker evaluates pairs of alternatives in sequence according to a preference ordering. Inferior alternatives are eliminated at each stage. Continuing with this procedure will eventually lead to the best alternative according to the preference ordering. However the decision-maker may stop the process at any stage with some probability and choose from the remaining alternatives with equal probability. This procedure is called a Gradual Pairwise Comparison (GPC) procedure. In a sense, GPC procedures are duals of BPCRs. In the former, alternatives are eliminated on the basis of the decision-maker's preference ordering; alternatives are then chosen using a uniform probability distribution over the remaining alternatives. Exactly the opposite occurs in a BPCR. Alternatives are eliminated by means of a uniform rule over subsets of alternatives of fixed size. An alternative is then chosen over the remaining alternatives using the decision-maker's preference ordering.

Our paper is also a contribution to the literature on consideration sets in the context of stochastic decision functions inspired by Luce [15]. This literature takes one of two approaches. The first (for example, Ahumada and Ulku [3] and Echenique and Saito [8]) considers a deterministic consideration function that selects a subset of every menu from which an outcome is then picked by a (random) Luce procedure. The second approach (for example, Manzini and Mariotti [16], Brady and Rehbeck [4], Kovach and Suleymanov [14], Cattaneo et al. [5] and Cheung and Masatlioglu [6]) to which our paper belongs is one where consideration sets are picked randomly from a menu. An alternative is then picked from each chosen consideration set by maximizing a deterministic preference ordering.

Our rule does not belong to the class of rules analyzed in Manzini and Mariotti [16] and Brady and Rehbeck [4]. Among other reasons, BPCRs are deterministic for some menus and probabilistic for others. At a methodological level, we do not assume the existence of a default alternative on which both Manzini and Mariotti [16] and Brady and Rehbeck [4] rely heavily (see also Horan [12] and Kovach and Suleymanov [14]). Cattaneo et al. [5], Aguiar et al. [2] and Cheung and Masatlioglu [6] consider general models of random attention. They characterize (the large) classes of stochastic decision functions that satisfy mild monotonicity requirements on attention. Our rule belongs to the classes characterized by these papers [3] but is, of course, one of several that do so. A relative strength of our result resides in its ability to characterize a fairly narrow class of stochastic decision functions. This class uses far fewer parameters than, for instance, the rules in Cattaneo et al. [5] or Cheung and Masatlioglu [6]: a preference ordering and one positive integer compared to a preference ordering and 2^X real numbers. This makes it much more manageable if one needs, for instance, to elicit the parameters from a set of empirical observations.

Theorem 3.3.1 by Rafaï [19] characterizes a large class of stochastic decision function resembling ours: from each menu (not only the large ones), it draws a sub-menu using some probability distribution defined over a family D of sub-menus of X, with D closed under inclusion. When a sub-menu is drawn, the decision-maker chooses according to a linear ordering $\succ$ (like in our rule). The parameter space for this rule is huge. When the probability distribution is restricted to be uniform, Rafaï's rule does not coincide with ours because (1) it draws sub-menus even from small menus and (2) it does not draw sub-menus of fixed size.

The rest of the paper is organized as follows. Section 2 sets out the model and the axioms. Section 3 contains the main result of the paper. Section 4 shows the independence of our axioms. The Appendix contains the proof of an elementary combinatorial identity used in the proof of our main result.

[3] We would like to thank Yusufcan Masatlioglu for pointing out that a BPCR is a special case of attention overload.

2. The Model and Axioms

There is a finite set of alternatives $X = \{x, y, \ldots\}$ of size n. Let $\mathcal{P}$ denote the set of all probability distributions over X.

Definition **1.** *A stochastic choice function p is a mapping from $2^X \setminus \emptyset$ into $\mathcal{P}$ such that $p(x, Y) = 0$ whenever $x \notin Y$.*

A menu Y is a non-empty subset of X. A stochastic choice function p assigns probability $p(x, Y)$ to every $x \in Y$. Thus $p(x, Y) \geq 0$ and $\sum_{x \in Y} p(x, Y) = 1$. Fishburn [9] also uses the primitive p, but, unlike him, we do not define a deterministic choice function: we try to explain p by means of an unobserved preference relation $\succ$.

Let $\succ$ be a strict ordering[4] over X and let $k \geq 2$ be an integer. For all menus Y such that $|Y| > k$, let $Y(k) = \{Z \subset Y : |Z| = k\}$. For all menus Y such that $|Y| > k$ and $x \in Y$, let $Y_x(k) = \{Z \in Y(k) : x \succ z$ for all $z \in Z \setminus \{x\}\}$. The set $Y(k)$ denotes the set of subsets of Y that have exactly k alternatives. The set $Y_x(k)$ is the set of subsets of $Y(k)$ that have x as their $\succ$-maximal alternative. Note that $Y_x(k) = \emptyset$ is a possibility.

We are interested in the class of stochastic choice functions defined below.

Definition **2.** *A Bounded Processing Capacity Rule (or BPCR) is a stochastic choice function $p_k^{\succ}$ where $\succ$ is a strict ordering over X and k is an integer ($n \geq k \geq 2$) such that, for all menus Y,*

$$p_k^{\succ}(x, Y) = \begin{cases} 1 & \text{if } |Y| \leq k \text{ and } x \succ z \text{ for all } z \in Y \setminus \{x\} \\ \frac{|Y_x(k)|}{|Y(k)|} & \text{if } |Y| > k. \end{cases}$$

[4] A strict ordering over X is a complete, transitive and antisymmetric binary relation on X.

The decision-maker has a strict ordering $\succ$ over X and is able to process $k \geq 2$ alternatives at a time. When presented with a menu with no more than k alternatives, the decision-maker chooses the best alternative in the menu according to $\succ$; otherwise he selects a subset of exactly k alternatives with uniform probability and chooses the best alternative in the subset according to $\succ$.

It is possible to compute the probability of choosing an alternative in a menu of size greater than k in terms of its rank in the menu according to $\succ$. Let Y be a menu with $|Y| > k$. Let $x \in Y$. The rank of x in Y (according to $\succ$) is denoted by $r_Y^{\succ}(x) = |\{z \in Y : x \succ z\}|$. Note that there are $\binom{|Y|}{k}$ subsets of size k and $\binom{i}{k-1}$ of those subsets have as maximal alternative the alternative of rank i in Y.[5] Using this observation, Definition 2 can be rewritten as follows:

Definition **3.** *A Bounded Processing Capacity Rule (or BPCR) is a stochastic choice function $p_k^{\succ}$ where $\succ$ is a strict ordering over X and k is an integer, $n \geq k \geq 2$ such that, for all menus Y,*

$$p_k^{\succ}(x, Y) = \begin{cases} 1 & \text{if } |Y| \leq k \text{ and } r_Y^{\succ}(x) = |Y| - 1, \\ \frac{\binom{r_Y^{\succ}(x)}{k-1}}{\binom{|Y|}{k}} & \text{if } |Y| > k. \end{cases}$$

Remark **1.** Let $p_k^{\succ}$ be a BPCR. Suppose the decision-maker has to choose from a menu Y where $|Y| = l$. If $l > k$, an alternative x in Y is chosen with strictly positive probability if and only if it "beats" at least $k-1$ alternatives in Y according to $\succ$, i.e $|\{y \in Y : x \succ y\}| \geq k-1$. Equivalently, there must be exactly $k-1$ alternatives in Y that are chosen with zero probability. On the other hand, if $l \leq k$, an alternative x is chosen if and only if $|\{y \in Y : x \succ y\}| = l - 1$.

[5] We adopt the convention that $\binom{i}{j} = 0$ if $i < j$.

Remark **2.** There is another feature of a BPCR which we have alluded to earlier, which deserves special mention. Consider an arbitrary BPCR $p_k^{\succ}$. Let Y be a menu with $|Y| > k$ and suppose $x, y \in Y$ with $x \succ y$. The rank of x in Y is therefore larger than the rank of y in Y. Consequently, $p_k^{\succ}(x, Y) > p_k^{\succ}(y, Y)$ according to Definition 3. Choice overload causes the decision-maker to commit "mistakes" while choosing her $\succ$-best alternative in Y. However, she is still more likely to choose a better alternative (according to $\succ$) than a worse one, i.e. less likely to make a bigger mistake.

We illustrate a BPCR with an example below.

Example **1.** Let $X = \{x_1, x_2, x_3, x_4\}$ with $x_1 \succ x_2 \succ x_3 \succ x_4$. For the purposes of this example, we shall introduce some special notation. A menu Y will be understood to be of the form $\{x_{i_1}, x_{i_2}, \ldots, x_{i_{|Y|}}\}$ where $i_1 < i_2 \ldots < i_{|Y|}$. For any capacity k, BPCR $p_k^{\succ}$ and menu Y, $p_k^{\succ}(\cdot, Y)$ will denote the $|Y|$-dimensional vector whose first element is the probability assigned to alternative x_{i_1}, the second element is the probability assigned to alternative x_{i_2} and so on.

Then,

$$p_2^{\succ}(\cdot, Y) = \begin{cases} (1, 0) & \text{if } |Y| = 2 \\ (\frac{2}{3}, \frac{1}{3}, 0) & \text{if } |Y| = 3 \\ (\frac{1}{2}, \frac{1}{3}, \frac{1}{6}, 0) & \text{if } |Y| = 4 \end{cases}$$

$$p_3^{\succ}(\cdot, Y) = \begin{cases} (1, 0) & \text{if } |Y| = 2 \\ (1, 0, 0) & \text{if } |Y| = 3 \\ (\frac{3}{4}, \frac{1}{4}, 0, 0) & \text{if } |Y| = 4 \end{cases}$$

and

$$p_4^{\succ}(\cdot, Y) = \begin{cases} (1, 0) & \text{if } |Y| = 2 \\ (1, 0, 0) & \text{if } |Y| = 3 \\ (1, 0, 0, 0) & \text{if } |Y| = 4 \end{cases}$$

Binary choices are deterministic when capacity is two. Binary and ternary choices are deterministic when capacity is three and all choices are deterministic when capacity is four. Note also that, for any capacity and menu, the decision-maker chooses a higher ranked alternative according to $\succ$ with higher probability. $\diamond$

We now introduce the axioms that characterize BPCRs. The first axiom is Sen's Condition α (see Rubinstein [20], Lecture 3) or contraction consistency, adapted to our context. If the alternative x is chosen with the highest probability in a menu, it must be chosen with the highest probability in every sub-menu of the original menu.

A 1. *Contraction (CONT): If $x \in Y \subset Z$ and $p(x, Z) > p(x', Z)$ for all $x' \in Z \setminus \{x\}$, then $p(x, Y) > p(x', Y)$ for all $x' \in Y \setminus \{x\}$.*

Fishburn [9] uses a similar condition (named P4') but with weak inequalities instead of strict ones.

The next axiom is a counterpart of Sen's Condition β or expansion consistency (Rubinstein [20], Lecture 3). If an alternative y is never chosen from a menu and the menu is enlarged with a "better alternative" x, then y continues to be never chosen. The alternative x is better than y in the sense that the former is chosen with a strictly higher probability than the latter from the binary menu $\{x, y\}$.

A 2. *Expansion (EXP): Suppose $y \in Y$ and $p(y, Y) = 0$. Let $x \notin Y$ and suppose $p(x, \{x, y\}) > p(y, \{x, y\})$. Then $p(y, Y \cup \{x\}) = 0$.*

Fishburn [9] uses a similar condition (named P4) but without the restriction that $p(x, \{x, y\}) > p(y, \{x, y\})$. Our condition is therefore weaker. Conditions similar to CONT and EXP are also found in Ok and Tserenjigmid [17].

The next axiom requires the existence of an alternative in every menu that is chosen with a strictly higher probability than any other alternative in the menu. The same axiom is used in Dutta [7].

A 3. *Unique Best (UB): For all $Y \subseteq X$, there is $y \in Y$ such that $p(y, Y) > p(x, Y)$ for all $x \in Y \setminus \{x\}$.*

Let $p^*(Y)$ be the largest choice probability observed in menu Y. In other words, $p^*(Y) = \max_{z \in Y} p(z, Y)$. Let Y^- be the set Y without the alternatives with maximal probability. Formally, $Y^- = Y \setminus \{z \in Y : p(z, Y) = p^*(Y)\}$.

The next axiom is a weak version of the well-known Luce independence axiom (Luce [15]). Let y and z be alternatives in Y^-, both chosen with non-zero probabilities. Then, the ratio of the probabilities of choosing y and z in Y and Y^- are the same, i.e. this ratio does not depend on whether the 'best' alternatives are available or not. The Luce Independence axiom requires this ratio to be independent for all pairs of alternatives and for all menus that include this pair.

A 4. *Limited Luce Independence (LLI): Let $y, z \in Y^-$ with $p(y, Y^-), p(z, Y^-) > 0$. Then*

$$\frac{p(y, Y^-)}{p(z, Y^-)} = \frac{p(y, Y)}{p(z, Y)}.$$

The LLI axiom imposes restrictions on the manner in which probabilities of non-maximal alternatives are updated when an alternative chosen with maximal probability is removed from a menu. The next axiom is also about updating, but relates to alternatives chosen with maximal probability. If an alternative is removed from a menu, it is natural to expect the probabilities in the smaller menu to increase since the same probability mass is now distributed over fewer alternatives. The following axiom requires the ratio of maximal probabilities in two menus that differ only by the removal of a maximal alternative in one, to be inversely proportional to the $s$ize of the two menus.

A 5. *Ratio of Maximals (RM): Let Y be a menu such that $p^*(Y) < 1$. Then,*

$$\frac{p^*(Y)}{p^*(Y^-)} = \frac{|Y^-|}{|Y|}.$$

The RM condition can be stated in another way. Noting that $|Y| = |Y^-| + 1$, rearrangement of terms yields $p^*(Y^-) = p^*(Y) + \frac{p^*(Y)}{|Y^-|}$. In other words, the maximal probability in the smaller menu Y^- is obtained by adding a fraction $\frac{1}{|Y^-|}$ of the maximal probability in the larger menu Y to itself.

The final axiom is the axiom of Neutrality. For any menu Y and stochastic choice function p, let $p(Y,Y) = \{p(Y,x) : x \in Y\}$ i.e. $p(Y,Y)$ is the set of probabilities with which various alternatives in Y are chosen by p.

A 6. *Neutrality (NEU): For any permutation σ of X and any $Y \subseteq X$, $p(Y,Y) = p(\sigma(Y),\sigma(Y))$.*

The NEU axiom is the familiar requirement that a stochastic choice function should not discriminate between alternatives. An important consequence of this axiom is that the set of probabilities for alternatives in different menus are the same if the menus are of the same size.

3. Characterization

Our result is the following:

Theorem 1. *A BPCR satisfies CONT, EXP, UB, LLI, RM and NEU. Conversely, let p be a stochastic choice function satisfying the axioms. Then there exists a unique strict ordering $\succ$ over X and a unique integer $k \geq 2$ such that the BPCR $p_k^{\succ}$ coincides with p.*

Proof. Necessity: Let $p_k^{\succ}$ be a BPCR where $\succ$ is a strict ordering over X and k is an integer such that $n \geq k \geq 2$. We show that $p_k^{\succ}$ satisfies all the axioms in the statement of the Theorem.

CONT: Suppose $x \in Y \subset Z$ and $p_k^{\succ}(x, Z) > p_k^{\succ}(x', Z)$ for all $x' \in Z \setminus \{x\}$. If $|Z| \leq k$, then x is maximal in the restriction of $\succ$ to Z. It is also maximal in the restriction of $\succ$ to Y. Hence $p_k^{\succ}(x, Y) > p_k^{\succ}(x', Y)$ for all $x' \in Y \setminus \{x\}$. If $|Z| > k$, x has rank $|Z| - 1$ in the restriction of $\succ$ to Z according to Definition 3. It has also rank $|Z| - 1$ in the restriction of $\succ$ to Y. Hence $p_k^{\succ}(x, Y) > p_k^{\succ}(x', Y)$ for all $x' \in Y \setminus \{x\}$.

EXP: Suppose $y \in Y$ and $p_k^{\succ}(y, Y) = 0$. Let $x \notin Y$ and suppose $p_k^{\succ}(x, \{x, y\}) > p_k^{\succ}(y, \{x, y\})$. Since $x \succ y$, we have $|\{z \in Y : y \succ z\}| = |\{z \in Y \cup \{x\} : y \succ z\}|$. Since $p_k^{\succ}(y, Y) = 0$, we have two cases: (i) if $|Y| > k$, we have $|\{z \in Y : y \succ z\}| < k - 1$ and (ii) if $|Y| \leq k$, then $|\{z \in Y : y \succ z\}| < |Y| - 1$. Both conclusions follow from Remark 1. Therefore, the following hold: (iii) if $|Y \cup \{x\}| > k$, we have $|\{z \in Y \cup \{x\} : y \succ z\}| < k - 1$ and (iv) if $|Y| \leq k$, then $|\{z \in Y \cup \{x\} : y \succ z\}| < |Y| - 1$. Again Remark 1 implies $p_k^{\succ}(y, Y \cup \{x\}) = 0$.

UB: This follows immediately by inspection of Definition 3.

LLI: Let Y be a menu such that $|Y| = l + 1$. Thus $|Y^-| = l$. Pick $y, z \in Y^-$ such that $p_k^{\succ}(y, Y^-), p_k^{\succ}(z, Y^-) > 0$. Let i (resp. j) be the rank of y (resp. z) in the restriction of $\succ$ to Y^-. Then the rank of y (resp. z) in the restriction of $\succ$ to Y is also i (resp. j). It follows from Definition 3 that

$$\frac{p_k^{\succ}(y, Y^-)}{p_k^{\succ}(z, Y^-)} = \frac{\binom{i}{k-1}\binom{l}{k}}{\binom{l}{k}\binom{j}{k-1}} = \frac{\binom{i}{k-1}}{\binom{j}{k-1}}.$$

Similarly,

$$\frac{p_k^{\succ}(y,Y)}{p_k^{\succ}(z,Y)} = \frac{\binom{i}{k-1}\binom{l+1}{k}}{\binom{l+1}{k}\binom{j}{k-1}} = \frac{\binom{i}{k-1}}{\binom{j}{k-1}}.$$

Clearly $\frac{p_k^{\succ}(y,Y^-)}{p_k^{\succ}(z,Y^-)} = \frac{p_k^{\succ}(y,Y)}{p_k^{\succ}(z,Y)}$ so that LLI is satisfied.

RM: Let Y be a menu such that $|Y| = l+1$. Thus $|Y^-| = l$. According to Definition 3, $p_k^{\succ *}(Y) = \frac{\binom{l}{k-1}}{\binom{l+1}{k}}$. Similarly, $p_k^{\succ *}(Y^-) = \frac{\binom{l-1}{k-1}}{\binom{l}{k}}$. After simplification, $p_k^{\succ *}(Y) = \frac{k}{l+1}$ and $p_k^{\succ *}(Y^-) = \frac{k}{l}$. Thus, $\frac{p_k^{\succ *}(Y)}{p_k^{\succ *}(Y^-)} = \frac{l}{l+1}$ as required.

NEU: This follows immediately by inspection of Definition 3.

Sufficiency: Let p be a stochastic choice function satisfying CONT, EXP, UB, LLI, RM and NEU. We begin by identifying a strict ordering $\succ$ over X.

Let $\succ$ be a binary relation defined on X by $x \succ y$ iff $p(x,\{x,y\}) = 1$, for $x \neq y$. We show that $\succ$ is (i) complete, (ii) anti-symmetric and (iii) transitive.

Suppose there exists $x, y \in X$ such that $0 < p(x,\{x,y\}) < 1$. Assume without loss of generality that $p(x,\{x,y\}) \geq p(y,\{x,y\})$. UB implies that the previous inequality must be strict. Applying RM, we have $\frac{p(x,\{x,y\})}{p(x,\{x\})} = \frac{1}{2}$. Since $p(x,\{x\}) = 1$, we conclude that $p(x,\{x,y\}) = \frac{1}{2}$. Therefore $p(x,\{x,y\}) = p(y,\{x,y\})$ contradicting UB. Consequently $0 < p(x,\{x,y\}) < 1$ cannot hold, i.e. $p(x,\{x,y\})$ is either zero or one and either $x \succ y$ or $y \succ x$ must hold. Moreover both cannot hold so that $\succ$ is anti-symmetric.

Suppose $\succ$ is not transitive, i.e. $x \succ y$, $y \succ z$ and $x \not\succ z$. Clearly $x \neq z$. Since $\succ$ is complete, we must have $z \succ x$. So, $p(x,\{x,y\}) = 1$, $p(y,\{y,z\}) = 1$ and $p(z,\{x,z\}) = 1$. According to UB, one of $p(x,\{x,y,z\})$,

$p(y, \{x, y, z\})$ and $p(z, \{x, y, z\})$ is strictly larger than the two others. Let us consider the three possible cases.

- $p(x, \{x, y, z\})$ is strictly larger than both $p(y, \{x, y, z\})$ and $p(z, \{x, y, z\}$. Then CONT implies $p(x, \{x, z\}) > p(z, \{x, z\}) = 1$.
- $p(y, \{x, y, z\})$ is strictly larger than both $p(x, \{x, y, z\})$ and $p(z, \{x, y, z\})$. Then CONT implies $p(y, \{x, y\}) > p(x, \{x, y\}) = 1$.
- $p(z, \{x, y, z\})$ is strictly larger than both $p(x, \{x, y, z\})$ and $p(y, \{x, y, z\})$. Then CONT implies $p(z, \{y, z\}) > p(y, \{y, z\}) = 1$.

The three cases imply a probability strictly larger than 1, which is impossible. This proves that $\succ$ is transitive. In view of the fact that $\succ$ is complete, anti-symmetric and transitive, all alternatives in X can be relabelled $1, 2, \ldots, n$ in a manner such that $i \succ j$ iff $i > j$. Thus the smallest and largest alternatives in X according to $\succ$ are 1 and n respectively. For any $j \in \{1, \ldots, n\}$, let $[j]$ denote the set $\{1, \ldots, j\}$.

Let k be the smallest integer in the set $[n-1]$ such that $p(j, [j]) = 1$ for all $j \leq k$ and $p(k+1, [k+1]) < 1$. If $p(j, [j]) = 1$ for all $j \in [n]$, we let $k = n$. We claim that $k \geq 2$. To see this, observe that $p(1, [1]) = 1$. If $k = 1$, then RM and $p(1, [1]) = 1$ imply $p(2, [2]) = p(1, [2]) = 1/2$, thereby violating UB. Therefore $k \geq 2$.

Consider the case where $k = n$. We will show that $p = p_n^{\succ}$. Pick an arbitrary menu Y. Suppose $|Y| = l$ and j^* is the alternative with the highest index present in Y, i.e. $j^* \succ x$ for all $x \in Y \backslash \{j^*\}$. By assumption, $p(l, [l]) = 1$ so that $p(r, [l]) = 0$ for all $r < l$. The NEU axiom implies the following: for any menu Z such that $|Y| = |Z|$, $p(Y, Y) = p(Z, Z)$ holds. By picking $Z = [l]$

and noting that $|Y| = |[l]|$, we can apply NEU to infer that $p(Y, Y) = \{1, 0\}$. Thus, there exists an alternative in Y which is chosen with probability one while the remaining alternatives are never chosen. We complete the proof by showing that the alternative chosen with probability one in Y is j^*. By assumption, $p(j^*, [j]) = 1$. Since $Y \subset [j^*]$, CONT implies $p(j^*, Y) > p(x, Y)$ for all $x \in Y \setminus \{j^*\}$. Therefore $p(j^*, Y) = 1$ as required.

We now consider the remaining case, i.e. there exists $k \in \{2, \ldots, n-1\}$ such that $p(j, [j]) = 1$, for all $j \leq k$ and $p(k+1, [k+1]) < 1$. Our goal is to show the following: for all $l \in \{k+1, \ldots, n\}$,

$$p(r, [l]) = \begin{cases} \frac{\binom{r-1}{k-1}}{\binom{l}{k}} \text{ if } k \leq r \leq l \\ 0 \text{ if } r < k. \end{cases} \tag{1}$$

We establish (1) by induction on l. We first show that (1) holds for $l = k+1$. Then we show that if (1) holds for some $l \geq k+1$, it also holds for $l+1$.

By assumption, $p(k, [k]) = 1$, so that $p(r, [k]) = 0$ for all $r < k$. Since $k \succ r$ for all $r < k$, EXP implies $p(r, [k+1]) = 0$ for all $r < k$. According to RM, $\frac{p(k+1,[k+1])}{p(k,[k])} = \frac{k}{k+1}$. Since $p(k, [k]) = 1$ by assumption, $p(k+1, [k+1]) = \frac{k}{k+1}$. Also $p(k, [k+1]) + p(k+1, [k+1]) = 1$, so that $p(k, [k+1]) = \frac{1}{k+1}$. Noting that $\frac{k}{k+1} = \frac{\binom{k}{k-1}}{\binom{k+1}{k}}$ and $\frac{1}{k+1} = \frac{\binom{k-1}{k-1}}{\binom{k+1}{k}}$, we can confirm that (1) is satisfied for the case of $l = k+1$.

Suppose that (1) holds for some integer l where $l \geq k+1$. We claim that (1) holds for $l+1$, i.e.

$$p(r,[l+1]) = \begin{cases} \frac{\binom{r-1}{k-1}}{\binom{l+1}{k}} \text{ if } k \leq r \leq l+1 \\ 0 \text{ if } r < k. \end{cases} \tag{2}$$

Since $l > k$, we have $l+1 \succ r$ for all $r < k$. Since $p(r,[l]) = 0$ for all $r < k$ by the induction hypothesis, we can apply EXP to conclude that $p(r,[l+1]) = 0)$ for all $r < k$.

We now consider $p(l+1,[l+1])$. We have

$$\begin{aligned} p(l+1,[l+1]) &= p(l,[l]).\frac{l}{l+1} \\ &= \frac{\binom{l-1}{k-1}}{\binom{l}{k}}.\frac{l}{l+1} \\ &= \frac{\binom{l}{k-1}}{\binom{l+1}{k}} \end{aligned} \tag{3}$$

The first equality follows from RM, the second from the induction hypothesis and the last from some routine manipulation. Observe that Equation 3 establishes Equation 2 for the case $r = l+1$.

We claim that $p(l+1,[l+1]) > p(r,[l+1])$ for all $r < l+1$. Suppose this is false. Since p satisfies UB, there exists $r \neq l+1$ such that $p(r,[l+1]) > p(r+1,[l+1])$. By CONT, $p(r,\{r,l+1\}) > p(l+1,\{r,l+1\})$ which contradicts our assumption that $l+1 \succ r$.

We now consider choices from the menu $[l+1]$. The argument in the previous paragraph implies that $l+1$ is the alternative which gets maximal probability in $[l+1]$. Pick r such that $l+1 > r > k$. According to the induction hypothesis, $p(r,[l]) = \frac{\binom{r-1}{k-1}}{\binom{l}{k}} > 0$ and $p(k,[l]) = \frac{\binom{k-1}{k-1}}{\binom{l}{k}} > 0$. Axiom LLI can therefore be applied to infer that

$$\frac{p(r,[l+1])}{p(k,[l+1])} = \frac{p(r,[l])}{p(k,[l])} = \binom{r-1}{k-1}. \tag{4}$$

Therefore,

$$p(r,[l+1]) = p(k,[l+1])\binom{r-1}{k-1}. \tag{5}$$

Adding over all r from k to l and noting that $p(k,[l+1]) = \binom{k-1}{k-1}p(k,[l+1])$, we have

$$\sum_{r=k}^{l} p(r,[l+1]) = p(k,[l+1])\sum_{r=k}^{l}\binom{r-1}{k-1}. \tag{6}$$

Since $\sum_{r=k}^{l} p(r,[l+1]) + p(l+1,[l+1]) = 1$ and $p(l+1,[l+1]) = \frac{\binom{l}{k-1}}{\binom{l+1}{k}}$, Equation 6 can be rewritten as follows:

$$\frac{\binom{l+1}{k} - \binom{l}{k-1}}{\binom{l+1}{k}} = p(k,[l+1])\sum_{r=k}^{l}\binom{r-1}{k-1}. \tag{7}$$

However,

$$\begin{aligned} & \sum_{r=k}^{l}\binom{r-1}{k-1} + \binom{l}{k-1} \\ = & \sum_{r=k}^{l+1}\binom{r-1}{k-1} \\ = & \binom{l+1}{k}. \end{aligned} \tag{8}$$

The last step in Equation 8 is an identity which is proved in the Appendix. Using (8) in Equation 7 and cancelling terms, we have

$$p(k,[l+1]) = \frac{1}{\binom{l+1}{k}} = \frac{\binom{k-1}{k-1}}{\binom{l+1}{k}}. \tag{9}$$

Observe that Equation 9 establishes Equation 2 for the case $r = k$. Finally, consider $r = k+1, \ldots, l$. Observe that Equation 5 implies

$$p(r,[l+1]) = \frac{\binom{r-1}{k-1}}{\binom{l+1}{k}}. \tag{10}$$

Thus, Equation 10 establishes Equation 2 for $r = k+1, \ldots, l$ and Equation 2 is verified generally. Since $p(j,[j]) = 1$ for all $j \leq k$ by assumption, comparison with the probabilities specified in Definition 2 confirms that p agrees with $p_k^{\succ}$ for the sets $[l]$, $l \in [n]$. The final step in the proof consists in showing that p and $p_k^{\succ}$ agree over all sets.

Recall that we have already proved the result for the case $k = n$. Assume therefore that $k < n$. Pick an arbitrary menu Y with $|Y| = l$. Assume without loss of generality that $Y = \{y_l, y_{l-1}, \ldots, y_1\}$ where $y_l \succ y_{l-1} \succ \ldots \succ y_1$. According to NEU, the set of choice probabilities of alternatives in Y is equal to the set of choice probabilities of alternatives in $[l]$. In order to complete the proof, we shall show that $p(y_j, Y) = p(j,[l])$ for all $j \in [l]$.

We first argue that $m_Y = y_l$, where m_Y denotes the alternative that has the largest probability in Y. Let y_l be the alternative j^* in the list of alternatives $1, 2, \ldots, n$. We can verify by observation that $j^* = m_Z$ where $Z = [j^*]$. Since $y_l \succ y_{l-1} \succ \ldots \succ y_1$ by assumption, we must have $Y \subset Z$. Applying CONT, we have $m_Y = j^* \equiv y_l$ as claimed.

We shall show that $p(y_j, Y) = p(j,[l])$ for all $j \in [l]$ by induction on l. We first claim that this is true for all $l \leq k$. In this case, $p(l,[l]) = 1$ and $p(j,[l]) = 0$ for all $j < l$. By applying NEU, it follows that $p(y_j, Y) = 1$ for some $y_j \in Y$. Since $m_Y = y_l$, it must be the case that $p(y_l, Y) = 1$ proving our claim.

Suppose $p(y_j, Y) = p(j, [r])$ for all $j \in [r]$ holds for all Y such that $|Y| = r$. In view of the argument in the previous paragraph, we can assume $r \geq k$. We will show that $p(y_j, Y) = p(j, [r+1])$ for all $j \in [r+1]$ holds for all Y such that $|Y| = r+1$. Let $Y = \{y_{r+1}, y_r, \dots, y_1\}$ where $y_{r+1} \succ y_r \succ \cdots \succ y_1$. We know that $m_Y = y_{r+1}$. Since $|Y^-| = r$, the induction hypothesis applies, so that $p(y_j, Y^-) = p(j, [r])$ for all $j \in [r]$. Observe that $p(j, [r]) = 0$ for $j \in [k-1]$ so that $p(y_j, Y^-) = 0$ for $j \in [k-1]$. Since $y_{r+1} \succ y_j$ for all $j \in [k-1]$, EXP implies $p(y_j, Y) = 0$ for all $j \in [k-1]$. Summarizing, we have shown the following thus far: $p(y_j, Y) = p(j, [r+1])$ for $j \in [k-1]$ and $j = r+1$.

Clearly $p(y_j Y^-) > 0$ for $j = k, \dots, r$. Consider the case where $r = k$. We know $p(y_{r+1}, Y) = p(r+1, [r+1])$ and $p(y_j, Y) = 0$ for all $j \in [k-1]$. Therefore NEU implies $p(y_r, Y) = p(r, [r+1])$ completing the proof of this case. Suppose $r > k$. Let $y_i, y_j \in Y \setminus \{y_{r+1}\}$ be such that $y_i \succ y_j$ and $p(y_i, Y^-), p(y_j, Y^-) > 0$, i.e. $i, j \in \{k, \dots, r\}$ and $i > j$. Applying LLI and the induction hypothesis, we have

$$\frac{p(y_i, Y)}{p(y_j, Y)} = \frac{p(y_i, Y^-)}{p(y_j, Y^-)} = \frac{p(i, [r])}{p(j, [r])}.$$

It follows from inspection that $p(i, [r]) > p(j, [r])$. Hence, we have $p(y_i, Y) > p(y_j, Y)$. Note that NEU and our earlier results imply that the numbers $p(y_r, Y), p(y_{r-1}, Y), \dots, p(y_k, Y)$ are a permutation of the numbers $p(r, [r+1]), p(r-1, [r+1]), \dots, p(k, [r+1])$. Since $p(r, [r+1]) > \dots p(r-1, [r+1]) > \dots > p(k, [r+1])$ and $p(y_r, Y) > p(y_{r-1}, Y) > \dots > p(y_k, Y)$, it follows that $p(y_j, Y) = p(j, [r+1])$ for $j = k, \dots, r$ and this completes the proof of the result. □

REMARK **3.** In the case where $n = 3$, the LLI and NEU axioms are redundant. Let $X = \{x_1, x_2, x_3\}$. Using CONT and UB, we can infer the ordering $\succ$ over X. Assume without loss of generality that $x_3 \succ x_2 \succ x_1$. There are two cases to consider. The first is $p(x_3, \{x_1, x_2, x_3\}) = 1$. Here p is the BPRC $p_3^{\succ}$. The second case is when $p(x_3, \{x_1, x_2, x_3\}) < 1$. By EXP, we have $p(\{x_1, x_2, x_3\}, x_1) = 0$ and RM implies $p(x_3, \{x_1, x_2, x_3\})/p(x_3, \{x_2, x_3\}) = 2/3$. Since $p(x_3, \{x_2, x_3\}) = 1$, $p(x_3, \{x_1, x_2, x_3\}) = \frac{2}{3}$ and $p(x_2, \{x_1, x_2, x_3\}) = \frac{1}{3}$. Thus p is the BPRC $p_2^{\succ}$. Note that the LLI and NEU axioms were not used in the argument.

4. INDEPENDENCE OF THE AXIOMS

.

In this section, we show that the axioms in Theorem 1 are independent. For each condition, we provide an example satisfying all conditions but one, indicated between parentheses. According to Remark 3, the axioms can only be independent when $n \geq 4$. Some of the examples below only consider the case $n = 4$. This is for expositional convenience—the examples can be extended generally.

EXAMPLE **2.** (CONT) Let $X = [4]$ and define the stochastic choice function p as follows:

- $p(1, [4]) = 1/2$, $p(2, [4]) = 1/3$, $p(3, [4]) = 1/6$,
- $p(1, [3]) = \frac{2}{3}$, $p(2, [3]) = \frac{1}{3}$,
- $p(2, \{1, 2\}) = 1$, $p(1, \{1, 3\}) = 1$, $p(1, \{1, 4\}) = 1$, $p(2, \{2, 3\}) = 1$, $p(2, \{2, 4\}) = 1$, $p(3, \{3, 4\}) = 1$.

For any menu Y of size 3 that is not equal to $[3]$, $p(\cdot, Y)$ is equal to $p(\cdot, [3])$ up to a bijection $\sigma : [l] \to Y$ respecting the natural ordering.

1. CONT is violated because $p(1,[3]) > p(2,[3])$, but $p(1,\{1,2\}) < p(2,\{1,2\})$.

2. Let us first consider the menus of size 3. One of them (namely [3]) has zero probability for alternative 3. But there is no $x \in X \setminus [3]$ where $p(x,\{x,3\}) > p(3,\{x,3\})$. Hence, EXP does not apply. The three other menus of size 3 have zero probability for alternative 4. For each such menu Y, there is $x \in X \setminus Y$ such that $p(x,\{x,4\}) > p(4,\{x,4\})$. In all cases, we have $p(4,[4]) = 0$ in accordance with EXP.

   Let us now consider the menus of size 2.

   - $p(1,\{1,2\}) = 0$, but there is no $x \in X \setminus \{1,2\}$ such that $p(x,\{x,1\}) > p(1,\{x,1\})$. Hence, EXP does not apply.
   - $p(3,\{1,3\}) = 0$, $p(2,\{2,3\}) > p(3,\{2,3\})$ and $p(3,\{1,2,3\}) = 0$ in accordance with EXP. The situation is similar for the remaining menus of size 2.

3. UB can be verified directly.

4. $Y = X$ is the only menu where two alternatives are chosen with strictly positive probability from Y^-. We have

$$\frac{p(2,\{2,3,4\})}{p(3,\{2,3,4\})} = 2 = \frac{p(2,[4])}{p(3,[4])}$$

   and LLI holds.

5. We have to check RM for all menus of size 3 or 4. For the menu [4], we have

$$\frac{p(1,[4])}{p(2,\{2,3,4\})} = 3/4 = \frac{|\{2,3,4\}|}{|[4]|}.$$

   For the menu $\{1,2,3\}$, we have

$$\frac{p(1,\{1,2,3\})}{p(2,\{2,3\})} = 2/3 = \frac{|\{2,3\}|}{|\{1,2,3\}|}.$$

   The situation is similar for the menus $\{1,2,4\}$, $\{1,3,4\}$ and $\{2,3,4\}$. Hence RM holds.

6. For menus of size 3, NEU holds by construction. For menus of size 2, NEU holds because there is always a probability equal to 1 and another equal to zero. ◇

EXAMPLE **3.** (EXP) Let $X = [4]$. Define the stochastic choice function p as follows:

- $p(2, [2]) = 1$, $p(1, [2]) = 0$,
- $p(3, [3]) = \frac{2}{3}$, $p(2, [3]) = \frac{1}{6}$, $p(1, [3]) = \frac{1}{6}$,
- $p(4, [4]) = \frac{1}{2}$, $p(3, [4]) = \frac{1}{3}$, $p(2, [4]) = \frac{1}{12}$, $p(1, [4]) = \frac{1}{12}$.

For any menu Y of size l that is not equal to $[l]$, $p(\cdot, Y)$ is equal to $p(\cdot, [l])$ up to a bijection $\sigma : [l] \to Y$ respecting the natural ordering.

1. Since $p(1, [2]) = p(1, \{1, 3\}) = 0$ and $p(1, [3]) > 0$, we have a violation of EXP.

2. CONT and UB can be verified directly by inspection. NEU holds by construction.

3. Note that if $Y = [3]$, then Y^- does not have two non-zero choice probabilities. Therefore the only candidate menu for verifying LLI is $Y = X$. Note that $\frac{p(3,[4])}{p(2,[4])} = \frac{p(3,[3])}{p(2,[3])} = 4$, $\frac{p(3,[4])}{p(1,[4])} = \frac{p(3,[3])}{p(1,[3])} = 4$ while $\frac{p(2,[4])}{p(1,[4])} = \frac{p(2,[3])}{p(1,[3])} = 1$. Therefore LLI is satisfied.

5. It suffices to verify RM only for menus of the type $[l]$. Since $\frac{p(4,[4])}{p(3,[3])} = \frac{3}{4}$ and $\frac{p(3,[3])}{p(2,[2])} = \frac{2}{3}$, RM is verified. $\diamond$

EXAMPLE **4.** (UB) Let $X = \{x_1, x_2, \ldots, x_n\}$ with $n \geq 3$. Define the stochastic choice function p as follows: for any menu Y and $y \in Y$, $p(y, Y) = \frac{1}{|Y|}$.

1. UB is violated because all probabilities in a menu are equal.

2. CONT is satisfied vacuously because no alternative in a menu is chosen with a strictly higher probability than the other alternatives.

3. EXP holds trivially because all alternatives in a menu are chosen with strictly positive probability.

4. LLI holds vacuously because the alternative m_Y is not well-defined for any menu Y.

5. RM holds vacuously for the same reason as LLI.

6. Let Y and Z be menus such that $|Y| = |Z|$. NEU holds with any bijection $\sigma : Y \to Z$. $\diamond$

Example **5.** (LLI) Let $X = [4]$. Define the stochastic choice function p as follows:

- $p(2,[2]) = 1$, $p(1,[2]) = 0$,
- $p(3,[3]) = \frac{2}{3}$, $p(2,[3]) = \frac{1}{3}$, $p(1,[3]) = 0$,
- $p(4,[4]) = \frac{1}{2}$, $p(3,[4]) = \frac{5}{12}$, $p(2,[4]) = \frac{1}{12}$, $p(1,[4]) = 0$.

For any menu Y of size l that is not equal to $[l]$, $p(\cdot, Y)$ is equal to $p(\cdot,[l])$ up to a bijection $\sigma : [l] \to Y$ respecting the natural ordering.

1. LLI is violated because $\frac{p(3,[4])}{p(2,[4])} = 5$ while $\frac{p(3,[3])}{p(2,[3])} = 2$.

2. CONT, EXP, UB can be verified directly. NEU is satisfied by construction.

3. RM is satisfied because $\frac{p(4,[4])}{p(3,[3])} = \frac{3}{4}$ while $\frac{p(3,[3])}{p(2,[2])} = \frac{2}{3}$. $\diamond$

Example **6.** (RM) Let $X = [4]$. Define the stochastic choice function p as follows:

- $p(2,[2]) = 1$,
- $p(3,[3]) = \frac{2}{3}$, $p(2,[3]) = \frac{1}{3}$,
- $p(4,[4]) = \frac{2}{3}$, $p(3,[4]) = \frac{2}{9}$, $p(2,[4]) = \frac{1}{9}$.

For any menu Y of size l that is not equal to $[l]$, $p(\cdot, Y)$ is equal to $p(\cdot,[l])$ up to a bijection $\sigma : [l] \to Y$ respecting the natural ordering.

1. RM is violated because $\frac{p(4,[4])}{p(3,[3])} = 1$ while RM requires this ratio to be $\frac{3}{4}$.

2. CONT, EXP, UB can be verified directly. NEU is satisfied by construction.

3. LLI is satisfied because $\frac{p(3,[4])}{p(2,[4])} = \frac{p(3,[3])}{p(2,[3])} = 2$. $\diamond$

Example 7. (NEU) Let $X = [4]$. Define the stochastic choice function p as follows:

- $p(i, \{i, j\}) = 1$ whenever $i > j$,
- $p(3, \{1, 2, 3\}) = \frac{2}{3}$ and $p(2, \{1, 2, 3\}) = \frac{1}{3}$,
- $p(4, \{1, 2, 4\}) = \frac{2}{3}$ and $p(2, \{1, 2, 4\}) = \frac{1}{3}$,
- $p(4, \{2, 3, 4\}) = 1$,
- $p(4, \{1, 3, 4\}) = 1$,
- $p(4, [4]) = \frac{1}{2}$, $p(3, [4]) = \frac{1}{3}$ and $p(2, [4]) = \frac{1}{6}$.

1. Let $Y = \{2, 3, 4\}$ and $Z = \{1, 2, 3\}$. There are two alternatives in Y that have zero probability of being chosen while there is only one such alternative in Z. Clearly there cannot be a bijection between $p(\cdot, Y)$ and $p(\cdot, Z)$. Hence NEU is violated.

2. UB can be easily verified by inspection. For any menu Y, $i = m_Y$ if $i > j$ for all $j \in Y$. Therefore p satisfies CONT.

3. Observe that 1 is never chosen in any menu while 4 is always chosen with positive probability. In order to verify EXP, we therefore need to only consider 2 and 3. Consider alternative 2. It is chosen with probability 0 in the menus $\{2, 3\}, \{2, 4\}$ and $\{2, 3, 4\}$. Enlarging $\{2, 3\}$ or $\{2, 4\}$ by adding a better alternative than 2 yields $\{2, 3, 4\}$ in which 2 is chosen with probability 0. Enlarging $\{2, 3, 4\}$ by adding a better alternative than 2 is not possible. Consider alternative 3. It is chosen with probability 0 only in the menus $\{3, 4\}$, $\{1, 3, 4\}$ and $\{2, 3, 4\}$. None of these menus can be enlarged by adding a better alternative than 3. Hence p satisfies EXP.

4. In order to verify LLI, we are required to identify menus Y such that there are at least two alternatives chosen with strictly positive probabilities in Y^{-}. The only such menu is X and the only two alternatives chosen with strictly positive probability in X^{-} are 3 and 2. Since $\frac{p(3,\{1,2,3\})}{p(2,\{1,2,3\})} = 2 = \frac{p(3,X)}{p(2,X)}$, LLI is satisfied.

5. In order to verify RM, we are required to identify menus Y such that $p(m_Y, Y) < 1$. The only such menus are X, $\{1,2,3\}$ and $\{1,2,4\}$. Since $\frac{p(4,X)}{p(3,\{1,2,3\})} = \frac{3}{4}$ and $\frac{p(3,\{1,2,3\})}{p(2,\{1,2\})} = \frac{p(4,\{1,2,4\})}{p(2,\{1,2\})} = \frac{2}{3}$, RM is satisfied. $\diamond$

5. Conclusion and discussion

We presented a new stochastic choice function modelling the behaviour of a decision-maker with limited processing capacity. When presented with a large menu, she randomly chooses a sub-menu of fixed size with uniform probability and selects the best alternative according to a strict ordering $\succ$. For smal menus, she just maximizes $\succ$. This rule is characterized by six independent axioms. In this section, we discuss some noteworthy features of our model and axioms.

5.1. The Uniform Distribution

A salient feature of the BPCR is that sub-menus of size k are drawn according to a uniform distribution. It it is tempting to consider other distributions, but if we do so without imposing any restrictions on the distribution, then none of our six axioms is guaranteed to hold. A sensible restriction would be to assume that sub-menus of size k containing more preferred alternatives (according to $\succ$) have a strictly greater chance of being sampled. In that case, only CONT, EXP and UB are guaranteed to hold.

We note that from a computational complexity viewpoint, drawing uniformly is "easier" than according to a distribution giving greater chance to more preferred alternatives. The uniform distribution assumption is therefore consistent with the notion of bounded processing capacity.

5.2. Processing capacity k

None of the axioms in Theorem 1 makes any reference to a threshold k and yet such an integer k appears in the definition of the family of stochastic choice functions characterized by the theorem. The following loose reasoning shows how the threshold emerges from the interaction between the axioms. Let t be the number of alternatives chosen with zero probability out of the menu X. By EXP, the number of alternatives chosen with zero probability out of smaller menus of the form $[l]$ cannot exceed t. By NEU, the same holds for any menu of size l. Consider the smallest menu $[l]$ such that the number of alternatives chosen with zero probability out of $[l]$ is t. By LLI, for all menus $[m]$ with $l < m < |X|$, the number of alternatives chosen with zero probability is at least t. But we have seen earlier that it cannot exceed t. It is therefore equal to t. If we set $k = t + 1$, we can see the threshold appearing. For other details, one needs to follow the proof of Theorem 1.

We would like to make a remark regarding the processing capacity k in the definition of a BPCR. We have assumed it to be deterministic but one could consider a model where it is an integer-valued random variable for instance, having a unimodal distribution centred around k. Such a stochastic choice function would satisfy CONT, EXP, UB and NEU but would violate LLI and

RM.

5.3. Unique Best

This condition can be seen as unduly restrictive, but it allows us to discuss the essential features of the BPCR without delving into cumbersome details. Indeed, if UB does not hold, then it is necessary to consider a preference relation $\succsim$ allowing for indifference and then adding a tie-breaking mechanism. This clearly makes the axiomatic analysis more complex whilst the more general model is essentially identical to the one we analyze in this paper.

5.4. Neutrality

Neutrality has consequences that are not immediately obvious: the choice probabilities associated with the elements of a menu depend only on the size of the menu and the rank of the elements in the restriction of $\succ$ to the menu. Other attributes of the elements in the menu do not play any role. For instance, any kind of cardinal information about the alternatives must be discarded (another name for Neutrality could be Ordinality). A virtue of Neutrality is therefore to constrain the BPCR to be simple from a computational/cognitive viewpoint.

Neutrality may seem to be the axiom that implies a uniform distribution but it does not, as shown by Example 6. Indeed, the stochastic choice function of Example 6 is neutral and can be seen as the result of a BPCR with $k = 2$ and $4 \succ 3 \succ 2 \succ 1$ except that the sub-menus are not drawn according to a uniform distribution but according to this distribution:

$$P(\{3,4\}|[4]) = P(\{2,4\}|[4]) = P(\{1,4\}|[4]) = 2/9,$$
$$P(\{2,3\}|[4]) = P(\{1,3\}|[4]) = P(\{1,2\}|[4]) = 1/9,$$
$$P(\{x,y\}|\{x,y,z\}) = 1/3 \quad \forall x,y,z \in [4],$$

where $P(Z|Y)$ represents the probability of drawing the sub-menu Z out of the menu Y.

It would be interesting to characterize stochastic choice functions that satisfy all the axioms in Theorem 1 except NEU. A close look at the proof of Theorem 1 shows that the resulting family of stochastic choice functions does not have a simple structure: the choice probabilities for all menus of the form $[l]$ are identical to those predicted by the BPCR, but the choice probabilities for other menus are less restricted and can deviate from the BPCR predictions, as in Example 7.

6. Appendix

Proposition: Pick integers k and l such that $l \geq k \geq 2$. Then $\sum_{r=k}^{l+1} \binom{r-1}{k-1} = \binom{l+1}{k}$.

Proof. Fix k. We prove the proposition by induction on l. We first show that it holds when $l = k$. This is true because

$$\binom{k-1}{k-1} + \binom{k}{k-1} = 1 + k = \binom{k+1}{k}.$$

Suppose the Proposition holds for l equal to some $s \geq k$ We show that it holds when $l = s+1$. We have

$$
\begin{aligned}
\sum_{r=k}^{l+1} \binom{r-1}{k-1} &= \sum_{r=k}^{s+2} \binom{r-1}{k-1} \\
&= \sum_{r=k}^{s+1} \binom{r-1}{k-1} + \binom{s+1}{k-1} \\
&= \binom{s+1}{k} + \binom{s+1}{k-1} \\
&= \frac{(s+1)!}{k!(s-k+1)!} + \frac{(s+1)!}{(k-1)!(s-k+2)!} \\
&= \frac{(s+1)!((s-k+2)+k)!}{k!(s-k+2)!} \\
&= \frac{(s+2)!}{k!(s-k+2)!} \\
&= \binom{s+2}{k} \\
&= \binom{l+1}{k}.
\end{aligned}
$$

□

[1] Aguiar, V. H., Boccardi, M. J., Dean, M., 2016. Satisficing and stochastic choice. Journal of Economic Theory 166, 445–482.

[2] Aguiar, V. H., Boccardi, M. J., Kashaev, N., Kim, J., 2018. Does random consideration explain behavior when choice is hard? evidence from a large-scale experiment. arXiv preprint arXiv:1812.09619.

[3] Ahumada, A., Ulku, L., 2018. Luce rule with limited consideration. Mathematical Social Sciences 93, 52–56.

[4] Brady, R., Rehbeck, J., 2016. Menu-dependent stochastic feasibility. Econometrica 84 (3), 1203–1223.

[5] Cattaneo, M., Ma, X., Masatlioglu, Y., Suleymanov, E., 2019. A random attention model. Journal of Political Economy forthcoming.

[6] Cheung, P. H., Masatlioglu, Y., 2020. Attention overload, available at http://econweb.umd.edu/~masatlioglu/attentionoverload.pdf (2020/03/06).

[7] Dutta, R., 2020. Gradual pairwise comparison and stochastic choice. Theoretical Economics 15, 1335–1364.

[8] Echenique, F., Saito, K., 2019. General luce rule. Economic Theory 68 (4), 811–826.

[9] Fishburn, P. C., 1978. Choice probabilities and choice functions. Journal of Mathematical Psychology 18 (3), 205–219.

[10] Frick, M., 2016. Monotone threshold representations. Theoretical Economics 11, 757–772.

[11] Geng, S., Özbay, E. Y., 2019. How many is too many, available at http://econweb.umd.edu/~ozbay/capacity.pdf (2019/03/12).

[12] Horan, S., 2019. Random consideration and choice: A case study of default options. Mathematical Social Sciences 102, 73–84.

[13] Iyengar, S. S., Lepper, M. R., 2000. When choice is demotivating: Can one desire too much of a good thing? Journal of personality and social psychology 79 (6), 995.

[14] Kovach, M., Suleymanov, E., 2021. Reference dependence and random attention. arXiv preprint arXiv:2106.13350.

[15] Luce, R., 1959. Individual Choice Behavior: A Theoretical Analysis. New York: John Wiley and Sons.

[16] Manzini, P., Mariotti, M., 2014. Stochastic choice and consideration sets. Econometrica 82 (3), 1153–1176.

[17] Ok, E. A., Tserenjigmid, G., 2022. Indifference, indecisiveness, experimentation, and stochastic choice. Theoretical Economics 163, 651–686.

[18] Ortoleva, P., 2013. The price of flexibility: Towards a theory of thinking aversion. Journal of Economic Theory 148, 903–934.

[19] Rafaï, I., 2019. Integration of limited attention in economic theory. Ph.D. thesis, COMUE Université Côte d'Azur (2015-2019).

[20] Rubinstein, A., 2012. Lecture Notes in Microeconomic Theory: The Economic Agent. Princeton University Press, Princeton and Oxford.

[21] Scheibehenne, B., Greifeneder, R., Todd, P. M., 2010. Can there ever be too many options? a meta-analytic review of choice overload. Journal of consumer research 37 (3), 409–425.