Proximity and Stochastic Choice*

Saptarshi Mukherjee
Department of Humanities and Social Sciences
Indian Institute of Technology - Delhi, India
Email: saptarshi@hss.iitd.ac.in

Ruhi Sonal
Department of Humanities and Social Sciences
Indian Institute of Technology - Delhi, India
Email: ruhisonal@gmail.com

September 15, 2016

Abstract

We model a decision-making environment where choice takes place in two stages- first, the decision-maker considers a *sub-list* from a list of alternatives and next, she decides to choose an alternative from the sub-list considered in the previous stage. Following the influential work by Manzini and Mariotti (2014), we consider the situation when the first stage decision-making exhibits variability, i.e. the decision-maker ends up paying attention to different sub-lists at different times. We relax a significant assumption made in Manzini and Mariotti (2014) called "menu independence" and still show that the choice rule can be fully characterized when it satisfies the behavioral axioms: proximity, list regularity and list assymetry. The proximity axiom introduces the notion that the influence of an alternative (a_j) on the choice probability of another alternative (a_k) in a list depends

*We thank Salvador Barberá, Miguel A. Ballester, Paola Manzini, Marco Mariotti, Tugce Cuhadaroglu, Jordi Massó, Arunava Sen, Debasis Mishra, Avinash Bora, Kemal Yildiz, and seminar/conference participants at 13th Meeting of Society for Social Choice and Welfare.

only on the "proximal" alternatives with respect to both these alternatives (i.e. alternatives lying between a_j and a_k) in that list. In our result, all behavioral parameters which include a complete and antisymmetric binary relation over the alternatives in any list, are fully identified.

Keywords: Attention parameter, choice function, list, stochastic choice, two-stage decision-making.
JEL Classification Numbers: C72, C78, D82.

1 Introduction

We model a decision-maker whose choice responses exhibit variability. Although many models in economic choice theory deal with deterministic behavior, market and experimental data often reveal stochastic nature of decision-making (McFadden (2000)). As Barberá and Pattanaik (1986) observe, development of psychology literature has also been inclined towards stochastic choice models. The problem of "imperfect" or "limited" attention has also been explored by Masatlioglu et al. (2012) and Lleras et al. (2015). Masatlioglu et al. (2012) provide an axiomatic characterization that enables revelation of preferences while relaxing the implicit assumption made in consumer theory that a decision-maker considers all feasible alternatives. Their model recognizes the possibility that a decision-maker may be unaware of the existence of some alternatives in the set from which she is making the choice. Lleras et al. (2015) employ a framework that views the problem as "choice overload" in which different products have to compete for the decision-maker's attention.

Manzini and Mariotti (2014) develop a model of a boundedly rational agent who maximises a preference relation but randomness in the choice data arises due to imperfect attention. They focus on the framework in which the decision-making is conceived as a two-stage process: the agent may not evaluate all possible options in a set A but only *considers* a subset (possibly strict) of it, called the *consideration set*, $C(A)$; once $C(A)$ is formed the final choice is made by maximising a preference relation over $C(A)$. The agent *considers* each feasible alternative $a \in A$ with a given (unobservable) probability (the attention parameter) and Manzini and Mariotti (2014) makes a specific assumption in their model: attention parameter or the probability for any alternative $a \in A$ to be included in the consideration set $C(A)$ is *menu independent*, i.e. it does not depend on the other alternatives in A. Given this assumption, they find the characterization of the choice rule when it satisfies two simple axioms.

Previously, two-stage decision making has been analyzed within deterministic frameworks in various contexts: Eliaz, Richter and Rubinstein (2011) axiomatize procedures using which a decision-maker may shortlist alternatives before making a final choice. They consider that examining each alternative may be "costly" for the decision-maker, and therefore it is convenient to formulate a consideration set in the first stage, from which a final choice is made. However, as Manzini and Mariotti (2014) point out, among the studies relating to consideration set model of choice, Masatlioglu et al is the pioneering work as it is the first to study how attention and preferences can be retrieved from a given choice data in this category of models. But in their model it is not possible to pin down the primitives by observing the choice data generated. Manzini and Mariotti (2014)'s contribution is significant in that regard: their characterization result uniquely identifies the primitives.

While Manzini and Mariotti (2014) point out that the "menu independence" is a significant assumption, they also show that the unrestricted menu dependence generates a model with no observable restrictions and "..it is not clear a priori what partial restrictions should be imposed on menu dependence". Manzini and Mariotti (2014) leave the characterization of such a decision problem as an open question.

In this paper we consider a richer structure of the set of alternatives- we assume that the agent encounters the alternatives in the form of a (finite) sequence or a list. Rubinstein and Salant (2006) introduces the notion of choice from lists. List is a natural restriction on the structure of the set of alternatives in many real life choice problems. For example, when the decision-maker buys a product online, the website displays the products in the form of a list; when a researcher decides on a journal to submit a paper, the journal names appear in a sequence to her mind etc. There are many more examples from daily life and we will mention some of them in the following discussion. We analyze two-stage decision-making in the framework of lists. In the first stage, the decision-maker pays attention to a sub-list of the entire list (say, ℓ), and in the second stage she chooses an alternative from the sub-list selected in the first stage. The randomness is attributed to the first stage of the decision-making procedure when a *consideration list* is formed. Once the *consideration list* is formed, we assume that the decision-making in the second stage is deterministic.

Note that the randomization over the sub-lists in the first stage is compatible with various behavioral assumptions about the decision-maker that result in "imperfect" or "limited" attention (à la Masatlioglu et. al. (2012); Lleras et al. (2015); Manzini and Mariotti (2014), as discussed above) in this stage. For example, the decision-maker may pre-commit to concentrate on different sub-lists in different occasions. This could be based on several factors as exemplified by the following examples:

(1) consider an individual visiting a supermarket to buy a pet-bottle of

soft drink. If time-pressured, she may consider only the products in specific positions on the shelf of the soft drinks (say, the first and the middle and the last item of the list, to get a feel of the variety; or the first few products are considered and the rest is ignored);

(2) suppose that an individual plans to buy a cosmetic product and she prefers buying products online. On an occasion when the internet speed is slow, she may avoid browsing through all the pages to see the entire list of products- she may just plan to consider the products listed on the first page and ignore clicking the "next" button (this example is attributed to Masatlioglu et. al. (2012)).

Such mechanisms may be adopted by the decision-maker due to cognitive constraints that may make it costly to pay attention to every alternative (Eliaz, Richter and Rubinstein (2011)).

When the decision-maker does not pre-commit to paying attention to alternatives that appear in specific positions, she may observe the alternatives appearing in a list along the way and decide which ones to consider. In this case also several factors may lead to limited attention or randomness in the consideration lists generated in the first stage. We discuss two examples below:

(1) often representatives from various companies are present near the product-shelves in the supermarkets to promote their products and explain their attributes. A consumer visiting the series of products may get influenced and consider the products accordingly. This effect could be attributed to choosing different "reference points" (Tversky and Kahneman 1991; Rubinstein and Salant 2006) to form the consideration lists.

One could construct a similar example for online shopping as well, where various discounts (or "combination offers") which appear while going through the list influence the consumer to form her consideration list.

(2) Another important factor that could cause attention paid to an alternative to differ is the possibility that the decision-maker is "unaware" about the existence or attributes of a particular alternative in the list. For example, a new brand of wrist watches may often be ignored by consumers until they become "aware" of its quality. Goeree (2008) explores this problem arising out of imperfect information and stresses the role of advertising as an important influence in determining the set from which the consumers ultimately choose from. Consumers may "consider" only those alternatives about which they have some prior information. A new product may therefore attract less attention even if it is of superior quality in the absence of awareness. A regular shopper in a supermarket may gradually become aware of a new product kept on a shelf of the market and the probability of paying attention to it may increase.

We provide an axiomatic characterization of the class of random choice rules from lists, i.e. the rules which specify the probability with which an alternative is chosen from any list. Unlike Manzini and Mariotti (2014), where the characterization requires the attention paramter to be menu independent, our characterization allows the attention parameter to be menu/list dependent. We adapt two axioms employed by Manzini and Mariotti (2014) to the framework of lists, namely, *list regularity* and *list asymmetry* to develop a model with identifiable parameters for attention as well as revelation of the decision-maker's preferences. We introduce the notion of "proximity" of any two alternatives in a list and relate it to measuring the influence of one alternative on the other. Specifically, our axiom, *"proximity"* can be seen as an independence requirement stating that the influence of removing an alternative a_j on the choice probability of another alternative a_k in any list

$l = (a_1, \dots, a_K)$, where $K \geq k > j$ (without loss of generality) for the corresponding list, should be independent of the alternatives before a_j and after a_k in that list. Thus the axiom requires that the influence of an alternative (a_j) on the choice probability of another alternative (a_k) in a list depends only on the "proximal" alternatives with respect to both these alternatives i.e. on the alternatives lying between a_j and a_k in that list.

Apesteguia and Ballester (2016) model preference heterogeneity in a random utility model. They introduce the notion that the presence of an alternative in between two alternatives plays a crucial role in determining choice probabilities of alternatives placed at the two extremes. Their axiom-"centrality" requires that given a particular order, in every triplet when one of the two extreme alternatives is removed, the choice probability of the other extreme alternative remains unchanged. Our axiom of "proximity" differs from it as we require that *only* those alternatives that lie in between two alternatives in a list influences their choice probabilities. More crucially, in Apesteguia and Ballester (2016) the order of the alternatives is endogenous. On the other hand, in our model the order in which the alternatives appear is exogenously given to the decision-maker, as in the examples illustrated earlier.

The *proximity* axiom characterizes a class of random choice rules, called "dominated sub-lists rules". There is a specific form of *random consideration function* (which determines probabilities for selecting different sub-lists in the first stage) with respect to which a dominated sub-list rule (say, P) is defined as follows: consider any list l. There is a complete and antisymmetric binary relation T^l over the set of alternatives appearing in l. Probability of selecting an alternative x from l, as given by $P(x, l)$ is then computed as follows: add the probabilities of the sub-lists (as given by the random consideration function) in which x appears and x beats all other alternatives in the sub-lists in a pairwise comparison according to T^l. As a decision-making process, the rule appears to be natural: the event of choosing any alternative x from a list l is the union of sub-events in which attention to paid to a sub-list l' in which the other alternatives in that list are *dominated* by x according to a complete and antisymmetric binary relation. The name of the rule: "dominated sub-lists rules" follows from the notion of dominance. Note that the binary relation reflects the "true" taste of the decision-maker over the alternatives and this binary relation is subject to change with the list from which the choice is made.

In this regard it may be pertinent to compare our rule with the rationalizable stochastic choice rules as discussed in Falmagne (1978) or Barberà and Pattanaik (1986): a stochastic choice rule C defined over the subsets of a finite set X is rationalizable if there exists a probability assignment over the set of all possible linear orderings over X such that $C(x, A)$, $A \subseteq X$, is the union of probability measures over all those linear orderings over X, in each of which x is the best alternative among all the alternatives in A. In our model, if P is a dominated sub-lists choice rule from lists, there exists a complete and antisymmetric binary relation for every list and a probability assignment over all the sub-lists such that $P(x, l)$ is the union of probability measures over all those sub-lists containing x, in each of which x *beats* the other alternatives in the sub-list as per the binary relation. Thus the randomness in the stochastic choice rules in Falmagne (1978) is induced by the probability assignment over the linear orderings and in the dominated sub-lists choice rules (from lists) the origin of the randomness is the probability assignment over the possible sub-lists, i.e. consideration lists.

We know that the observed choice data may not reflect the true taste of the decision-maker (for a detailed discussion, see Masatlioglu et. al. (2012)). The binary relation in any dominated sub-lists choice rule reflects the true taste of the decision-maker. In our model this underlying binary relation is identified for any dominated sub-lists choice rule. Also the attention parameters are identified for singleton and binary lists (i.e. the lists with two alternatives). Interestingly, we show that identifying the attention parameters for binary lists is sufficient to characterize the rule in general. Hence our model fits well with the strand of literature in revealed preference theory. Thus we address the concerns as raised by Manzini and Mariotti (2014) in their paper and fully characterize the decision-making procedure from lists and also identify the underlying parameters.

2 Random choice from lists

Let X be a non-empty finite set of alternatives. A *list* ℓ is a finite sequence of alternatives drawn from X. We assume that each alternative from X appears only once in a list, and denote the set of all possible lists from X by Λ. For $\ell \in \Lambda$, $X(\ell)$ is the set of alternatives appearing in ℓ, while the *length* $le(\ell)$ of ℓ is defined by $le(\ell) := |X(\ell)|$. That is, $\ell = (a_1, \ldots, a_{le(\ell)})$. Given any list $l \in \Lambda$, a sub-list l' of l (denoted by $l' \subseteq l$) is a list where $X(l') \subseteq X(l)$ and the order in which the alternatives appear in l' is the same as they appear in l. For any $x \in X$, we let (x) denote the (singleton) list containing only x. The set $\Lambda(k)$ contains all lists of length at most k.

The decision-maker observes a list and chooses an alternative from it. The decision-maker, however has an option not to pick any alternative from a list- in this case we assume that a default alternative x^* is chosen by the decision-maker. Let $X^* = X \cup \{x^*\}$.

Definition 1 A *random choice rule* (from lists) is a mapping $P : X^* \times \Lambda \to [0,1]$ such that: $\sum_{a \in X(\ell) \cup \{x^*\}} P(a,\ell) = 1$ for each $\ell \in \Lambda$; $P(a,\ell) = 0$ for each $a \in X \setminus X(\ell)$; and $P(a,\ell) \in (0,1)$ for each $\ell \in \Lambda$ and $a \in X(\ell)$

In our model the DM does not necessarily *consider* the whole list, but may pay attention only to a sub-list (called "consideration list") and chooses an element from it. Thus it is a two-stage decision making process as in Manzini and Mariotti (2014)- in the first stage the decision-maker (randomly) decides on a sub-list to consider and in the final stage she chooses an element from that sub-list. The randomness in the choice data arises from the first stage as in the second stage, the choice from the consideration list is assumed to be deterministic. Note that x^* is chosen from any list l only if no sub-list draws attention in the first stage. We denote by $c_j(\ell)$, a *consideration sub list* drawn from a list ℓ, containing those alternatives in ℓ that are paid attention to until the decision maker observes the j^{th} alternative. When the j^{th} alternative is the last alternative in the list ℓ, the consideration sub list is the decision-maker's *consideration list,* $c(\ell)$.

Definition 2 A *random consideration function* (from lists) is a mapping $\pi : \Lambda \times \Lambda \to [0,1]$ such that: $\pi(\ell',\ell) \in [0,1]$ for each $\ell' \subseteq \ell \in \Lambda$ and $\sum_{\ell' \subseteq \ell} \pi(\ell',\ell) = 1$.

A random consideration function from lists (henceforth, *rcf*) tells us the probability of considering (or paying attention to) a particular sub-list from any given list. We note that the consideration list from any list $l \in \Lambda$ can be empty and in this case the default alternative x^* is chosen. Thus $P(x^*,l) = \pi(\phi,l)$ for any $l \in \Lambda$. Since the decision-maker does not necessarily pay attention to the *entire* list, $P(x,l)$, $x \in X(l)$ may not reflect the genuine taste or preference for x compared to the other alternatives in l. As discussed in Manzini and Mariotti (2014), if $P(x,l) > P(y,l)$, $x,y \in X(l)$, it may be due to the fact that the decision-maker pays attention to x more frequently than to y even though she prefers y more to x (perhaps in a stochastic sense). Let $\delta : X \times \Lambda \rightarrow [0,1]$ denote *attention parameter from lists*, i.e. $\delta(x,l)$, $x \in X(l)$, $l \in \Lambda$ is the probability of drawing attention to x in any list l. When a decision-maker observes alternatives in a list, the sequence in which the alternatives appear can influence the probability that attention is paid to an alternative. The probability that attention is paid to an alternative b in the lists (a,b) and (b,a) may differ. In the list (a,b), the probability that attention is paid to b is conditioned by the event that a has been considered (or has not been considered). When a is considered by the decision-maker, the probability that attention is paid to b is $\delta(b,(a,b)|a \in c((a,b)))$. When a is not considered, the corresponding attention parameter for b is $\delta(b,(a,b)|a \notin c((a,b)))$. Thus, for any list $\ell \equiv (\ell_1, \ldots, \ell_K)$ and any sub-list $\ell' \equiv (\ell'_1, \ldots, \ell'_k)$, $\ell'_j \in X(\ell)$, $\forall 1 \leq j \leq k$, we have the following:

$$\pi(\ell', \ell) = \Pi_{j=1}^{j=k} \Pi_{i=1}^{j=k} \delta(\ell'_j, \ell | c_j(\ell))(1 - \delta(\ell''_i, \ell | c_j(\ell))$$

where $\ell''_i \in X(\ell) \setminus \{\ell'_1, \ldots, \ell'_k\}$

In this model the decision-making follows a two-stage process: given any list, a consideration list is formed (with some probability based on δ) in the first stage and then an alternative is chosen from the consideration list formed in the first stage. Given this process we can express the probability of choosing an alternative (say, x) from any list l as follows:

$$P(x,l) = \Sigma_{\{l' \subseteq l | x \in X(l')\}} \pi(l', l).s(x, l').$$

Here $s : X \times \Lambda \rightarrow [0,1]$, such that $s(a,l') = 0$ if $a \notin X(l')$, $s(x,l') \in \{0,1\}$ for any $x \in X(l')$ and $\Sigma_{\{x \in X(l')\}} s(x,l') = 1$, $l' \in \Lambda$. s can be viewed as a degenerate probability distribution which takes either of the two values 0 or 1. This is because the choice from the consideration list in the second stage is deterministic. The *true* preference of a decision maker is represented by s, as it tells us the probability of an alternative being chosen when all the alternatives in the list are considered.

We note that in Manzini and Mariotti (2014), the "attention parameters" measure the probability with which an alternative can be shortlisted from any set. However, Manzini and Mariotti (2014) assume that the attention parameter of any alternative remains the same *irrespective* of the set from which the alternative is being considered. **They call this assumption "menu independence". Thus in Manzini and Mariotti (2014) the attention parameters are "menu independent" where menu is the set of feasible alternatives. As Manzini and Mariotti (2014) mention in their paper, this assumption is a significant one.** This is because, the probability with which an alternative grabs attention in a set is likely to depend on what other alternatives are available in the set, i.e. the entire feasible set of alternatives. For example, more colorful objects are likely to attract more attention when some shabby objects are also present in the feasible set ("Contrast effect" (Rubinstein and Salant, 2006)). But Manzini and Mariotti (2014) show that relaxing this assumption in their setup would lead to a highly permissive result and also the underlying preferences remain unidentified: "Unrestricted menu dependence yields a model with no observable restrictions (Theorem 2), while it is not clear a priori what partial restrictions should be imposed on menu dependence." (Manzini and Mariotti, 2015). In this paper, we characterize the problem while allowing attention parameters to be list dependent.

Let P be a random choice rule and let π and δ denote the random consideration function and the attention parameter. For any $a, b \in X$, we have the following:

$$P(a,(a)) = \pi(a,(a)) = \delta(a,(a)) = \delta(a,(b,a)|x \notin c((b,a)))$$

$$P(a,(a,b)) = \pi((a,b),(a,b))s(a,(a,b)) + \pi((a),(a,b))s(a,(a)) = \delta(a,(a,b))\delta(b,(a,b)|a \in c((a,b))).s(a,(a,b)) + \delta(a,(a,b))(1-\delta(b,(a,b)|a \notin c((a,b))).$$

$$P(b,(a,b)) = \delta(a,(a,b))\delta(b,(a,b)|a \in c((a,b))).s(b,(a,b)) + (1-\delta(a,(a,b)))\delta(b,(a,b)|a \notin c((a,b)))$$

Next we will define a particular random consideration function- *"reference dependent random consideration function"*. Informally the probability of "considering" (or "paying attention to") any sub-list from a list (say, ℓ' from ℓ) in this rule can be described as follows: the decision-maker picks an alternative as a reference point, and then compares all other alternatives to it, one at a time. The probability of paying attention to a particular sub-list ℓ' with a_j as the reference is the joint probability of paying attention to all those alternatives a_i that occur in ℓ' in binary list (a_i, a_j) or (a_j, a_i) (depending on whether a_i precedes or succeeds a_j in the list ℓ), and the probability of not paying attention to those alternatives a_l in binary lists (a_l, a_j) (or (a_j, a_l)) which do not occur in ℓ'. We define it formally.

Definition 2 Let $\delta : X \times \Lambda \to [0,1]$ be an attention parameter for lists. A random consideration function π is a **reference dependent rcf**, denoted π_j, if for any $\ell = (a_1, \ldots, a_k) \in \Lambda$; and $a_j \in \ell' \subseteq \ell$

$$\pi_j(\ell', \ell) = \Pi_{i=1,\ldots j-1; a_i \in X(l')}\delta\left(a_i,(a_i,a_j)\right).\delta(a_j,(a_i,a_j)|a_i \in c((a_i,a_j))) \\ .\Pi_{l=j+1,\ldots k; a_l \in X(\ell')}\delta(a_j,(a_j,a_l)).\delta(a_l,(a_j,a_l)|a_j \in c((a_j,a_l)))$$

$$.\Pi_{i=1,\dots j-1; a_i \in X(\ell)\setminus X(\ell')}(1-\delta(a_i,(a_i,a_j)))\delta(a_j,(a_i,a_j)|a_i \notin c((a_i,a_j)))$$
$$.\Pi_{l=j+1,\dots k; a_l \in X(\ell)\setminus X(\ell')}\delta(a_j,(a_j,a_l)).(1-\delta(a_l,(a_j,a_l)|a_j \in c((a_j,a_l)))$$

Thus a **reference based rcf** computes the probability of any sublist ℓ' from ℓ as the probability of a joint event where any element in ℓ' (say, y) draws attention in binary list $(x,y) \subseteq \ell'$ (or in $(y,x) \subseteq \ell'$ but any element (say, z) in $\ell \setminus \ell'$ fails to draw attention in binary lists $(x,z) \subseteq \ell \setminus \ell'$ (or $(z,x) \subseteq \ell \setminus \ell'$, when x is the reference alternative.

3 Characterization of "dominated sub-lists rule"

3.1 Axioms and characterization

We borrow the idea of "impact of an alternative on the other" from Manzini and Mariotti (2014) and adapt it in our setting as follows: consider a list $\ell \in \Lambda$, $\ell = (a_1,\dots,a_k,\dots,a_K)$. Suppose that $x \in X(\ell)$, $x \neq a_k$. We define the "impact" of x on the choice probability of a_k in ℓ as follows: $\frac{P(a_k,\ell\setminus x)}{P(a_k,\ell)}$. Note that x could be an alternative that appears before or after a_k in ℓ. From lemma 1, it is clear that in any binary list, the impact of an alternative on the other is greater than or equal to unity. Thus the removal of an alternative from a binary list weakly increases the probability of the other alternative.

We adapt Manzini and Mariotti's (2014) axioms to the framework of lists:

Axiom 1: List regularity (LR) For any $a, b \in X(\ell)$

$$P(a,(a)) \geq P(a,(b,a))$$

Axiom 2: List assymetry (LA) For any $a, b \in X(\ell)$

$$P(b,(b))>P(b,(a,b)) \iff P(a,(a)) = P(a,(b,a))$$

LR states that when an alternative is removed from a binary list, the choice probability of the remaining alternative may increase or may remain the same. This rules out situations in which the removal of an alternative reduces the chances of another alternative being chosen. For example, a decision-maker may be more likely to choose *coffee* from the list *(tea,coffee)* than from the singleton *(coffee)*. Such situations do not conform to list regularity. LA requires that for any two alternatives, only one may impact the other. When the decision-maker observes the list (a, b), and observing a prior to b influences her probability of choosing b, then if b were observed prior to a, $a's$ choice probability would remain unaffected. Notice that s represents the probability of choosing an alternative when all the alternatives in the list are considered, i.e., there is no problem of imperfect attention. As an alternative may or may not be chosen, $s(x, \ell)$ can take only one of the two values: 0 or 1. In a binary list (a, b), $s(a,(a, b)) = 1$ indicates that when both a and b are considered, the decision maker chooses a. As $s(a,(a, b) + s(b,(a, b)) = 1$, $s(b,(a, b))$ must be 0. Hence, s in this example reveals that the decision-maker has a preference for a over b.With the above two axioms, we can identify the genuine preference s in our model:

Suppose for some $a, b \in X$, $P(a,(a)) = P(a,(b,a))$. Then,
$P(a,(a)) = P(a,(b,a)) = \delta(b,(b,a)).\delta(a,(b,a)|b \in c((a,b))).s(a,(b,a)) + [1 - \delta(b,(b,a))].\delta(a,(b,a)|b \notin c((a,b)))$ (1)
As $P(a,(a)) = \delta(a,(b,a)|b \notin c((a,b)))$; (1) can be written as:

$P(a,(a)) - [1-\delta(b,(b,a))].P(a,(a)) = \delta(b,(b,a)).\delta(a,(b,a)|b \in c((a,b))).s(a,(b,a))$

Which implies

$$\delta(b,(b,a)).P(a,(a)) = \delta(b,(b,a)).\delta(a,(b,a)|b \in c((a,b))).s(a,(b,a)) \qquad (2)$$

As $P : X \to (0,1)$, $\delta(b,(b,a)) \neq 0$. Also, $P(a,(a)) \neq 0$. The left hand side in (2) is therefore non zero.

It follows that in the right hand side of (2), $s(a,(b,a)) = 1$. Thus, s is identified.

Further, $s(a,(b,a)) = 1$ implies that $s(b,(b,a)) = 0$. Therefore, we know that:

$P(b,(b,a)) = \delta(b,(b,a)) - \delta(b,(b,a)).\delta(a,(b,a)|b \in c((b,a)))$

Which can be re-written as:

$$\delta(b,(b,a)).\delta(a,(b,a)|b \in c((b,a))) = \delta(b,(b,a)) - P(b,(b,a)) \qquad (3)$$

and

$$P(a,(b,a)) = \delta(b,(b,a)).\delta(a,(b,a)|b \in c((b,a))) + P(a,(a))[1-\delta(b,(b,a))] \qquad (4)$$

From (3) and (4):

$P(a,(b,a)) - P(a,(a)) = \delta(b,(b,a)) - P(b,(b,a)) - P(a,(a)).\delta(b,(b,a))$

Which implies

$\delta(b,(b,a)) = \frac{P(a,(b,a))+P(b,(b,a))-P(a,(a))}{1-P(a,(a))}$

Thus, the list dependent attention parameter $\delta(b,(b,a))$ is identified.

Using the expression derived for $\delta(b,(b,a)$ in $P(b,(b,a))$ with $s(b,(b,a)) = 0$, we get:

$\delta(a,(b,a)|b \in c((b,a))) = \frac{P(a,(b,a))-P(a,(a))[1-P(b,(b,a))]}{P(a,(b,a))+P(b,(b,a))-P(a,(a))}$

Thus, the attention parameter $\delta(a,(b,a)|b \in c((b,a)))$, which represents the conditional probability that attention is paid to a in the list (b,a), given that the decision-maker pays attention to b is also identified.

We introduce our final axiom which is based on the influence of proximity on the impact of an alternative on the choice probability of another alternative.

Axiom 3: Proximity Let P be a random choice rule. For any $\ell \in \Lambda$, $\ell = (a_1, \ldots, a_k, \ldots, a_K)$, we have that for each $t \in \{-k+1, -k+2, \ldots, -k+K\} \setminus \{0\}$,

$$\frac{P(a_k, \ell)}{P(a_k, \ell \setminus (a_{k+t}))} = \frac{P(a_k, \ell')}{P(a_k, \ell' \setminus (a_{k+t}))}$$

holds for each $\ell' = \ell \setminus (x)$, $x \in \{a_1, \ldots, a_{k+t-1}, a_{k+t+1}, \ldots, a_K\}$.

The idea behind this axiom is simple: impact of an alternative (say, x) on another (say, y) is *independent* of the alternatives not appearing in between x and y. We explain with an example. Suppose that the decision-maker is choosing a product from a shelf in a store and observes the product as displayed on the shelf, i.e. in the form of a list. The impact of a product x on the choice probability of another product y in this case depends only on the products the decision-maker encounters *between* x and y. This axiom promotes the idea that only the alternatives that are in proximity to *both* x and y affect the measure of the impact.

A very useful implication for a random choice rule from lists satisfying the above requirement is captured by the following lemma.

Lemma 1 *A random choice rule P satisfies Proximity if and only if for any $\ell \in \Lambda$, $\ell = (a_1, \ldots, a_K)$,*

$$P(a_j, \ell) = \frac{\Pi_{i=1,\ldots,j-1} P(a_j, (a_i, a_j)) \Pi_{i=j+1,\ldots,K} P(a_j, (a_j, a_i))}{[P(a_j, (a_j))]^{K-2}}$$

Proof. Let P satisfy Proximity. Take $\ell = (a_1, \ldots, a_K) \in \Lambda$ and $a_j \in X(\ell)$. Consider first the list $\ell' = (a_1, \ldots, a_j)$ and notice that by the repeated use of

Proximity we have

$$\frac{P(a_j, \ell')}{P(a_j, \ell' \setminus (a_{j-1}))} = \frac{P(a_j, (a_{j-1}, a_j))}{P(a_j, (a_j))}$$

and thus

$$P(a_j, \ell') = \frac{P(a_j, \ell' \setminus (a_{j-1}))P(a_j, (a_{j-1}, a_j))}{P(a_j, (a_j))}$$

holds.

We have further that

$$\frac{P(a_j, \ell' \setminus (a_{j-1}))}{P(a_j, \ell' \setminus (a_{j-2}, a_{j-1}))} = \frac{P(a_j, (a_{j-2}, a_j))}{P(a_j, (a_j))}$$

resulting in

$$P(a_j, \ell' \setminus (a_{j-1})) = \frac{P(a_j, \ell' \setminus (a_{j-2}, a_{j-1}))P(a_j, (a_{j-2}, a_j))}{P(a_j, (a_j))}.$$

Thus,

$$P(a_j, \ell') = \frac{P(a_j, \ell' \setminus (a_{j-2}, a_{j-1}))P(a_j, (a_{j-2}, a_j))P(a_j, (a_{j-1}, a_j))}{[P(a_j, (a_j))]^2}.$$

Continuing in the same way we get

$$P(a_j, \ell') = \frac{\Pi_{i=1,\ldots,j-1} P(a_j, (a_i, a_j))}{[P(a_j, (a_j))]^{j-1}}.$$

Consider then the list $\ell'' = (\ell', a_{j+1})$ and notice that

$$\frac{P(a_j, \ell'')}{P(a_j, \ell'' \setminus (a_{j+1}))} = \frac{P(a_j, (a_j, a_{j+1}))}{P(a_j, (a_j))}$$

and thus

$$\begin{aligned} P(a_j, \ell'') &= \frac{P(a_j, (a_j, a_{j+1}))P(a_j, \ell'' \setminus (a_{j+1}))}{P(a_j, (a_j))} \\ &= \frac{P(a_j, (a_j, a_{j+1}))P(a_j, \ell')}{P(a_j, (a_j))} \\ &= \frac{P(a_j, (a_j, a_{j+1}))\Pi_{i=1,\ldots,j-1} P(a_j, (a_i, a_j))}{[P(a_j, (a_j))]^j}. \end{aligned}$$

Consequentially adding all alternatives from ℓ to ℓ' results then in

$$P(a_j,\ell)=\frac{\Pi_{i=1,\ldots,j-1}P(a_j,(a_i,a_j))\Pi_{i=j+1,\ldots,K}P(a_j,(a_j,a_i))}{[P(a_j,(a_j))]^{K-2}}$$

as required to be shown.

Suppose now that $P(a_j,\ell)=\frac{\Pi_{i=1,\ldots,j-1}P(a_j,(a_i,a_j))\Pi_{i=j+1,\ldots,K}P(a_j,(a_j,a_i))}{[P(a_j,(a_j))]^{K-2}}$ holds for $\ell=(a_1,\ldots,a_K)\in\Lambda$ and $a_j\in X(\ell)$. Take $a_g,a_h\in X(\ell)$ and suppose that w.l.o.g. $g<h<j$. Notice that

$$\begin{aligned}
P(a_j,\ell) &= \frac{P(a_j,\ell\setminus(a_g,a_h)).P(a_j,(a_g,a_j)).P(a_j,(a_h,a_j))}{[P(a_j,(a_j))]^2}\\
&= \frac{P(a_j,\ell\setminus(a_h)).P(a_j,(a_g,a_j))}{P(a_j,(a_j))}\\
&= \frac{P(a_j,\ell\setminus(a_g)).P(a_j,(a_h,a_j))}{P(a_j,(a_j))}
\end{aligned}$$

and thus,

$$\frac{P(a_j,\ell)}{P(a_j,\ell\setminus(a_h))}=\frac{P(a_j,\ell\setminus(a_g))}{P(a_j,\ell\setminus(a_g,a_h))}$$

immediately follows as required to conclude that Proximity is satisfied. ■

Let us state the random choice rule which we characterize in this section.

Definition 3: Let π_j be a *reference dependent rcf.* A random choice rule from lists P is called **"dominated sub-lists rule"** if for any $l\in\Lambda$, there is a complete and antisymmetric binary relation T^l over $X(l)$ such that $x\in X(l)$,

$$P(x,l)=\Sigma_{\{l'\subseteq l|x\in X(l'),xT^ly\forall y\in X(l')\}}\pi_x(l',l)$$

Consider any $x\in X(l)$, $l\in\Lambda$. Let π^δ be a reference dependent rcf. For every $l^*\in\Lambda$ there exists a complete and antisymmetric relation T^{l^*} over $X(l^*)$. The rule defined as above computes the probability of choosing x from l as follows: consider *all* sub-lists of l where x appears and the other alternatives in those sub-lists are *dominated* by x according to T^{l^*}. Next, add the probabilities of shortlisting these sub-lists from l as provided by π_x. This value is $P(x,l)$. One can argue that the rule is a natural one- the randomness is generated by the shortlisting stage (i.e. the first stage) of the decision-making process and the event of selecting any alternative x from any list is just the union of all events in which a sub-list is considered where x is the best alternative.

Let $P_{\delta,T}$ denote a representative dominated sub-lists rule with δ as the underlying attention parameter of the rule and and $T = \{T^l\}_{l\in\Lambda}$ is the set of complete and antisymmetric binary relations- one for each admissible list $l \in \Lambda$.

We are now ready to present our main result.

Theorem 1 *A random choice rule P satisfies Proximity if and only if it is "dominated sub-lists" rule.*

Proof. In view of the proof of Lemma 2, a dominated sub-lists rule satisfies Proximity.

Suppose P is a random choice rule satisfying Proximity. Let $l \in \Lambda$.

Define the following binary relations for any $\ell_i, \ell_j \in X(l)$, $i < j$:

$l_i T_1^l l_j$ if $s(l_i,(l_i,l_j)) = 1$,

$l_j T_2^l l_i$ if $s(l_j,(l_i,l_j)) = 1$.

For any $x, y \in X(l)$, we say that $xT^l y$ if either $xT_1^l y$ or $xT_2^l y$.

T^l is a complete and antisymmetric relation over $X(l)$.

We know from Lemma 1 that for any $\ell = (a_1,\ldots,a_K) \in \Lambda$, any $j : 1 \leq j \leq K$, we have that

$$P(a_j,\ell).\{[P(a_j,(a_j))]^{K-2}\} = \\ \Pi_{i=1,\ldots,j-1}P(a_j,(a_i,a_j))\Pi_{k=j+1,\ldots,K}P(a_j,(a_j,a_k)) \qquad (1)$$

Notice that

$$P(a_j, (a_i, a_j)) =$$
$$\pi_{a_j}((a_i, a_j), (a_i, a_j)).s(a_j, (a_i, a_j)) + \pi_{a_j}((a_j), (a_i, a_j)).s(a_j, (a_j))$$

Here π_{a_j} is the reference based rcf as follows from the previous discussion. Expressing π_{a_j} in terms of attention parameters and setting $s(a_j, (a_j)) = 1$, we get:

$$\begin{aligned}
&P(a_j, (a_i, a_j)) = \\
&\delta(a_i, (a_i, a_j)).\delta(a_j, (a_i, a_j)|a_i \in c((a_i, a_j))).s(a_j, (a_i, a_j)) \\
&\quad + (1 - \delta(a_i, (a_i, a_j))).\delta(a_j, (a_i, a_j)|a_i \notin c((a_i, a_j)))
\end{aligned}$$

Similarly,

$$\begin{aligned}
&P(a_j, (a_j, a_k)) = \\
&\delta(a_j, (a_j, a_k)).\delta(a_k, (a_j, a_k)|a_j \in c((a_j, a_k))).s(a_j, (a_j, a_k)) \\
&\quad + \delta(a_j, (a_j, a_k))(1 - \delta(a_k, (a_j, a_k)|a_j \in c((a_j, a_k))).
\end{aligned}$$

Therefore,

$$\begin{aligned}
&\Pi_{i=1,\ldots,j-1} P(a_j, (a_i, a_j)) \Pi_{k=j+1,\ldots,K} P(a_j, (a_j, a_k)) = \\
&\Pi_{i=1,\ldots,j-1}\{\delta(a_i, (a_i, a_j)).\delta(a_j, (a_i, a_j)|a_i \in c((a_i, a_j))).s(a_j, (a_i, a_j)) \\
&\quad + (1 - \delta(a_i, (a_i, a_j))).\delta(a_j, (a_i, a_j)|a_i \notin \\
&\quad c((a_i, a_j)))\}.\Pi_{k=j+1,\ldots,K}\{\delta(a_j, (a_j, a_k)).\delta(a_k, (a_j, a_k)|a_j \in \\
&\quad c((a_j, a_k))).s(a_j, (a_j, a_k)) \\
&\quad + \delta(a_j, (a_j, a_k))(1 - \delta(a_k, (a_j, a_k)|a_j \in c((a_j, a_k)))\}
\end{aligned}$$

Notice that for all a_i, a_k such that $s(a_j, (a_i, a_j)) = 0$ and $s(a_j, (a_j, a_k)) = 0$, the first term of the expression for $P(a_j, (a_i, a_j))$ (or $P(a_j, (a_j, a_k))$) becomes 0. Expanding the above expression yields the following:

$$\Pi_{i=1,\ldots,j-1}P(a_j,(a_i,a_j))\Pi_{k=j+1,\ldots,K}P(a_j,(a_j,a_k)) =$$
$$\Sigma_{\ell'\subseteq\ell|\forall a'_i,a'_k\in X(\ell'),a_i,a_k\in X(\ell)\setminus X(\ell');a_jT^la'_i,a_jT^la'_k}\Pi_{i=1}^{i=j-1}\delta(a_j,(a'_i,a_j)|a'_i\in$$
$$c((a'_i,a_j))).\delta(a'_i,(a'_i,a_j)).\delta(a_j,(a_i,a_j)|a_i\notin$$
$$c((a_i,a_j))).[1-\delta(a_i,(a_i,a_j))].\Pi_{k=j+1}^{k=K}\delta(a_j,(a_j,a'_k)).\delta(a'_k,(a_j,a'_k)|a_j\in$$
$$c((a_j,(a_j,a'_k)).\delta(a_j,(a_j,a_k)).[1-\delta(a_k,(a_j,a_k)|a_j\in c((a_j,a_k))]$$

Using the definition of π_{a_j} in the above expression, we get the following:

$$\Pi_{i=1,\ldots,j-1}P(a_j,(a_i,a_j))\Pi_{k=j+1,\ldots,K}P(a_j,(a_j,a_k)) =$$
$$\left[\Sigma_{\ell'\subseteq\ell|a_j\in X(l'),\forall x\neq a_j\in X(\ell');a_jT^lx}\pi_{a_j}(\ell',\ell)\right]$$

Using (1) in the right hand side of the above equation, we get:

$$P(a_j,\ell).\{P(a_j,(a_j))\}^{k-2} =$$
$$\left[\Sigma_{\ell'\subseteq\ell|a_j\in X(l'),\forall x\neq a_j\in X(\ell');a_jT^lx}\pi_{a_j}(\ell',\ell)\right]$$

Thus, $P(a_j,\ell)$ is simply the probability of considering those sublists in which a_j is the dominant alternative according to preference relation T^l, normalized for the probability of choosing a_j in the singleton list (a_j).

4 Concluding remarks

We have characterized two-stage decision making with randomness in the first stage within the framework of lists, based on the notion that choice probabilities are influenced by proximal alternatives. Our characterization introduces a new class of stochastic choice rules and does not require the stringent menu independence assumption of Manzini and Mariotti (2014). Within the list based probabilistic framework, we allow the attention to be menu or list dependent. Also, the genuine preferences of the decision maker and the attention parameters are fully identified.

References

[1] Aguiar, V. (2015): Random consideration sets: inferring preferences from psychological biases, Working Paper, Brown University.

[2] Barberá, S. and P. Pattanaik (1986): Falmagne and the rationalizability of stochastic choices in terms of random orderings, Econometrica 54(3), 707-715.

[3] Barberá, S. and B. Grodal (2011) Preference for flexibility and the opportunitites of choice, Journal of Mathematical Economics 47, 272-278.

[4] Barberá, S. and A. Neme (2015): Ordinal relative satisficing behavior, Working Paper, Barcelona GSE.

[5] Brady, R. and John Rehbeck (2015): Menu-dependent stochastic feasibility, Working Paper, University of California - San Diego.

[6] Manzini, P., Mariotti, M. and C.J. Tyson (2013): Two-stage threshold representations, Theoretical Economics 8, 875-882.

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

[8] Rubinstein, A. and Y. Salant (2006): A model of choice from lists, Theoretical Economics 1, 3-17.

[9] Eliaz, K., Richter, M. and A. Rubinstein (2011): Choosing the two finalists, Economic Theory 46(2), 211-219.

[10] Masatlioglu, Y., Nakajima, D. and E.Y. Ozbay (2012): Revealed attention, American Economic Review 102(5), 2183-2205.

[11] Bajraj, G. and L. Ülkü (2015): Choosing two finalists *and* the winner, Social Choice and Welfare 45, 729-744.

[12] Apesteguia, J. and M.A. Ballester (2016): Single crossing random utility models, Working Paper, Barcelona GSE