Theoretical Economics 1 (2006), 3–17 1555-7561/20060003

A model of choice from lists

ARIEL RUBINSTEIN
School of Economics, Tel Aviv University, and Department of Economics, New York University

YUVAL SALANT
Graduate School of Business, Stanford University

The standard economic choice model assumes that the decision maker chooses from *sets* of alternatives. In contrast, we analyze a choice model in which the decision maker encounters the alternatives in the form of a *list*. We present two axioms similar in nature to the classical axioms of choice from sets. We show that they characterize all the choice functions from lists that involve the choice of either the first or the last optimal alternative in the list according to some preference relation. We then relate choice functions from lists to the classical notions of choice correspondences and random choice functions.

KEYWORDS. Choice from lists, rational choice, partition independence, weak axiom of revealed preference, satisficing.

JEL CLASSIFICATION. D00.

1. INTRODUCTION

The standard economic choice model assumes that decision makers choose from *sets*. However, it is often the case that the decision maker encounters the alternatives in the form of a *list*. The list may be physical, in the sense that the elements of the choice problem are presented to the decision maker sequentially along some dimension such as time or space. For example, when looking for a job, the decision maker receives offers one after the other or when purchasing a product online, the alternatives are listed from left to right or top to bottom. The list may also be virtual, in the sense that the alternatives come to the decision maker's mind in some order. For example, this might occur when a researcher chooses a research topic or when he chooses a journal to submit a paper to. The researcher typically does not have a comprehensive set of alternatives in front of him; rather, the alternatives come to his mind in some sequential manner. Whether the list is physical or virtual, it appears that the order in which the decision maker encounters the alternatives may be a substantive factor affecting his choice.

Ariel Rubinstein: rariel@post.tau.ac.il
Yuval Salant: salant@stanford.edu

We thank Bart Lipman and two anonymous referees of this journal for insightful comments. We also thank Bob Wilson for fruitful discussions, and Rafi Aviav and Ron Siegel for helpful comments.

Indeed, a variety of cognitive and procedural effects suggest that choice is order-dependent. A primacy effect gives advantage to the first few alternatives in a list since people examine them more attentively. A recency effect emerges due to the fact that people recall more vividly alternatives that appear at the end of a list, a factor that gives advantage to the last few alternatives in a list. In other cases, decision makers pay special attention to alternatives that stand out relative to their neighbors in the list. For example, a low-priced item will draw special attention if it is surrounded by high-priced items. In addition, the first element in a list may serve as a reference point to which subsequent alternatives are compared (see Tversky and Kahneman 1991) and thus choice may depend on the element that appears first. For example, if the items in a list differ in quality, then that of the first item may serve as a benchmark to which the quality of subsequent items is compared. Experimental scientists are aware of the sensitivity of choice to the order of presentation and hence the common practice of presenting alternatives to participants in different orderings. Some experimental and empirical findings that relate to order-of-presentation effects are surveyed in Section 7.

In what follows we present a model of choice from lists and conduct a basic investigation of the model. Our analysis will follow that of the model of choice from sets conducted in Choice Theory.

We explore choice functions from lists. A list is a sequence of distinct elements of a finite "grand" set X. A choice function from lists singles out one element from every list. We discuss two properties of choice functions from lists. The first property is Partition Independence (*PI*), which states that dividing a list arbitrarily into several sublists, choosing from each sublist and then choosing from the list of chosen elements, yields the same result as choosing from the original list. The second property is List Independence of Irrelevant Alternatives, which states that omitting unchosen elements from a list does not alter the choice. We show the equivalence of these two properties and prove that they characterize a particular class of choice functions from lists. In this class, each function is parameterized by a preference relation $\succsim$ over X and a labelling of every $\succsim$-indifference set by "First" or "Last". Given a list, the function identifies the set of $\succsim$-maximal elements within the list and chooses the first or the last element among them according to the label of the $\succsim$-indifference set they belong to. This class naturally generalizes the class of preference maximizing procedures in the context of choice functions from sets.

We then extend the discussion to cases where the order of the elements in the list is not directly observable, such as when the list is virtual. Under these circumstances, it seems reasonable to analyze choice correspondences, which attach to every set of alternatives all the elements that are chosen for *some* ordering of that set. We show that choice functions from lists that satisfy *PI* induce choice correspondences that satisfy the Weak Axiom of Revealed Preference (*WARP*). Conversely, if one starts with a choice correspondence that satisfies *WARP*, it can be "explained" by a choice function from lists that satisfies *PI*. Thus, the model of choice from lists provides a new interpretation of the notion of choice correspondences.

We also consider situations in which the decision maker deterministically chooses from lists generated from sets by a random process. If we do not observe the actual order of the elements, yet we have access to "many" of the decision maker's choices, we can summarize his choices by a random choice function. A random choice function assigns to every set of alternatives a probability measure over the set where the probability of an element is the likelihood that it will be chosen from the set. We show that a choice function from lists satisfies *PI* if and only if the induced random choice function is monotone in the sense that the probability of choosing an element from a set weakly increases when the set of available alternatives shrinks.

While our notion of a list does not allow duplication of elements in a given list, we will show toward the end of the paper that some of the results carry over to the case where repetition of elements in a list is allowed.

2. The model and examples

Let X be a finite set of N elements. A *list* is a non-empty finite sequence of distinct elements of X. Let $\mathscr{L}$ be the set of all lists. A *choice function from lists* $D : \mathscr{L} \to X$ is a function that assigns to every list $L = (a_1, \dots, a_K)$, a single element $D(L)$ from the set $\{a_1, \dots, a_K\}$. We abbreviate and write $D(a_1, \dots, a_K)$ instead of $D((a_1, \dots, a_K))$. Note that our formulation excludes lists in which an element appears more than once. In Section 6, we relax this assumption and investigate a model in which alternatives can appear multiple times in a list.

To demonstrate the richness of the framework and motivate the analysis to follow, we describe below several examples of families of choice functions from lists.

Example 1 (Rational choice). The decision maker has a strict preference relation (i.e., complete, asymmetric and transitive) $\succ$ over X and chooses the $\succ$-best element from every list. ◊

Example 2 (Satisficing (Simon 1955)). The decision maker has a strict preference relation $\succ$ over X and a satisfactory threshold $a^* \in X$. He chooses the first element in the list that is not inferior to a^*; if there is none he chooses the last element in the list. ◊

Example 3 (Place-dependent rationality). The decision maker has a preference relation $\succ$ over the set $X \times \{1, 2, \dots, N\}$ with (x, k) interpreted as the alternative x in the k-th place of a list. From the list $(a_1, \dots, a_K)$ the decision maker chooses the alternative a_k for which $(a_k, k) \succ (a_l, l)$ for all $1 \le l \le K$, $l \ne k$. Thus, the decision maker is "rational" but treats the alternative x at one position differently than at another position. This could occur, for example, when the position of an element in a list reflects the popularity or the relevance of the alternative (e.g., when the alternatives are the results returned by a search engine) and thus conveys information about its quality. ◊

Example 4 (The first element dictates the ordering). The first element in the list serves as a reference point (see Tversky and Kahneman 1991) according to which the decision maker evaluates the alternatives. Formally, for every element $a \in X$ there is a corresponding ordering $\succ_a$ of X such that $D(a_1, \dots, a_K)$ is the $\succ_{a_1}$-best element in the set

$\{a_1,\ldots,a_K\}$. For example, given a distance function on X, let $x \succ_a y$ if x is more distant from a than y. Then, D chooses from every list the element that is farthest from the first element in the list. $\lozenge$

EXAMPLE 5 (Successive choice (Salant 2003)). The primitive of this choice procedure is a binary relation R over X, where $x\,R\,y$ is interpreted as "x rejects y". Given a list $L = (a_1,\ldots,a_K)$, the decision maker stores a_1 in a "register"; at stage t of the computation, $1 \le t < K$, the decision maker replaces the register value y with a_{t+1} if $a_{t+1}\,R\,y$. When the list ends the decision maker chooses the alternative in the register. For example, let $u : X \to \mathbb{R}$ be a utility function and $s : X \to \mathbb{R}_+$ be a "bias" function. We say that $x\,R\,y$ if $u(x) > u(y) + s(y)$. In this case, when comparing a new alternative x to the one stored in the register y, the decision maker adds a "bonus" $s(y)$ to the alternative in the register. $\lozenge$

EXAMPLE 6 (Stop when you start to decline). The decision maker has an ordering $\succ$ on X. He chooses the element a_k where k is the maximal integer such that $a_1 \prec a_2 \prec \cdots \prec a_{k-1} \prec a_k$; that is, the decision maker goes through the alternatives according to the order of the list and stops the process on the first occasion that the value declines. $\lozenge$

EXAMPLE 7 (Contrast effect). The decision maker classifies the elements of X into two classes: "good" and "bad". A good element appears to be even better the greater the number of bad elements that surround it. The decision maker chooses a good element that is surrounded by the largest number of consecutive bad elements. $\lozenge$

EXAMPLE 8 (Knockout tournament). Let $\to$ be a tournament, that is, a complete asymmetric binary relation on X. The choice from the list $(a_1,\ldots,a_K)$ is calculated in rounds. In the first round, the decision maker matches a_{2k-1} and a_{2k} for $1 \le k \le \lfloor \frac{1}{2}K \rfloor$. The "winner" in each match is the alternative that beats the other according to the relation $\to$. For the case that K is odd, the unmatched alternative a_K is considered a winner as well. The decision maker continues to the next round with a list that contains the "winners" of the first round ordered like the original list. The process continues until all alternatives besides one are eliminated. $\lozenge$

EXAMPLE 9 (A "pseudo-random" function). Let $\succ$ be an ordering over X, let $(a_1,\ldots,a_K)$ be a list, and let $n_K = \lfloor \log_2 K \rfloor$. For $1 \le i \le n_K$, define $\varepsilon_i = 1$ if $a_i \succ a_{i+1}$ and $\varepsilon_i = 0$ if $a_i \prec a_{i+1}$. The decision maker chooses the element a_k where k is the n_K-digit binary number $\varepsilon_1\varepsilon_2\cdots\varepsilon_{n_K}$. We call this function pseudo-random because the chosen element is picked by what seems to be an arbitrary criterion and the choice is sensitive to any permutation of adjacent elements at the beginning of the list. $\lozenge$

3. AXIOMS

In this section we explore a few axioms regarding choice functions from lists. Denote by $S(L)$ the set of all the elements in a list L, that is, $S(a_1,\ldots,a_K) = \{a_1,\ldots,a_K\}$. We say that the lists L_1 and L_2 are *disjoint* if $S(L_1) \cap S(L_2) = \emptyset$. For any m lists L_1, …, L_m that are pairwise disjoint, define $\langle L_1, \dots, L_m \rangle$ to be the list that is the concatenation of the m lists.

Partition Independence (*PI*): We say that a choice function from lists D satisfies property *PI* if for every pair of disjoint lists, $L_1, L_2 \in \mathscr{L}$, we have

$$D(\langle L_1, L_2 \rangle) = D(D(L_1), D(L_2)).$$

The property *PI* requires that a decision maker chooses the same element from the list whether

(i) he chooses from the list as a whole, or

(ii) he partitions the list into two sublists, chooses from each sublist, and then makes a choice from the two-element list of chosen elements.

Note that if a choice function D satisfies *PI*, then for every list $(a_1, \dots, a_K)$ we have $D(a_1, \dots, a_K) = D(D(\dots D(D(a_1, a_2), a_3), \dots, a_{K-1}), a_K)$. That is, $D(a_1, \dots, a_K)$ can be computed by $K-1$ operations of the function D over pairs as follows. Start by computing $D(a_1, a_2)$, then compare the "winner" with a_3 to obtain $D(D(a_1, a_2), a_3)$ and so on, until $D(\dots, a_{K-1})$ is compared with a_K to obtain the chosen element. This means that every choice function that satisfies *PI* is a successive choice function (see Example 5 in Section 2). However, it will follow from the proof of Proposition 2 that not every successive choice function satisfies *PI*.

List Independence of Irrelevant Alternatives (*LIIA*): We say that a choice function from lists D satisfies property *LIIA* if for every list $(a_1, \dots, a_K)$,

$$D(a_1, \dots, a_K) = a_i \Rightarrow D(a_1, \dots, a_{j-1}, a_{j+1}, \dots, a_K) = a_i$$

for all $1 \le j \le K$, $j \ne i$. In other words, deleting an element that is not chosen from a list does not alter the choice.

The above two properties are close in spirit to Axiom 1 in Plott (1973) and the familiar Independence of Irrelevant Alternatives axiom in the context of choice from sets. Our first result establishes the equivalence between the two axioms.

PROPOSITION 1. *D satisfies LIIA if and only if it satisfies PI.*

PROOF. Assume D satisfies *LIIA*. Let $L = \langle L_1, L_2 \rangle$. Let $a = D(L)$ and assume w.l.o.g. that $a \in L_2$. Both L_2 and $(D(L_1), a)$ are sublists of L that include a and hence, by *LIIA*, $D(L_2) = a$ and $D(D(L_1), a) = a$. Consequently,

$$D(\langle L_1, L_2 \rangle) = D(L) = a = D(D(L_1), a) = D(D(L_1), D(L_2)).$$

In the other direction, assume D satisfies *PI* and let $D(a_1, \dots, a_K) = a_i$. We need to show that $D(a_1, \dots, a_{j-1}, a_{j+1}, \dots, a_K) = a_i$, where w.l.o.g. we assume that $j > i$. By

PI, for every $h \geq i$ we have $a_i = D(a_1, \dots, a_K) = D(D(a_1, \dots, a_h), D(a_{h+1}, \dots, a_K))$. Thus $D(a_1, \dots, a_h) = a_i$ and $D(a_i, D(a_{h+1}, \dots, a_K)) = a_i$. Therefore

$$\begin{aligned} D(a_1, \dots, a_{j-1}, a_{j+1}, \dots, a_K) &= D(D(a_1, \dots, a_{j-1}), D(a_{j+1}, \dots, a_K)) \\ &= D(a_i, D(a_{j+1}, \dots, a_K)) \\ &= a_i. \end{aligned}$$

□

We now derive the main result of this section. Recall that in the context of choice functions from sets, the requirement that $C(A) = C(\{C(A_1), C(A_2)\})$, for every choice set A and its partition into A_1 and A_2, characterizes all the choice functions that maximize some strict preference relation over X. We show that a choice function from lists that satisfies *PI* can be rationalized as the result of maximizing a weak preference relation over X, where indifferences are resolved on a list-position basis, i.e. either the first or the last maximizer is chosen. In the context of an employer who wishes to hire a worker from a list of candidates, this form of rationalization would fit a practice by which all candidates are graded and the employer chooses the candidate with the highest grade. In the case of multiple candidates with the same highest grade, the employer chooses either the first or the last among them according to list position (e.g. the candidate that applied to the job first or was interviewed last). Thus, to be consistent with *PI*, a choice function from lists can use the order in the list to break ties only in a very particular way.

Formally, let $\succsim$ be a preference relation over X and let $\delta : X \to \{1,2\}$ be a function satisfying $\delta(x) = \delta(y)$ whenever $x \sim y$. We refer to δ as a *priority indicator*. We denote by $D_{\succsim,\delta}$ the choice function that chooses from every list L the first (or the last) $\succsim$-maximal element according to whether the δ-value of the set of $\succsim$-maximal elements in $S(L)$ is 1 (or 2). Obviously, if $\succsim$ is a strict relation, then the resulting choice function is rational (Example 1). In addition, any satisficing choice function D (Example 2) can be represented as $D = D_{\succsim,\delta}$ by having $\succsim$ induce two indifference sets of satisfactory and unsatisfactory elements, and having the δ-value of the satisfactory and unsatisfactory elements be 1 and 2, respectively.

PROPOSITION 2. *A choice function from lists D satisfies PI if and only if there exists a unique preference relation $\succsim$ over X and a unique priority indicator δ such that $D = D_{\succsim,\delta}$.*

REMARK. In Section 6, we extend Proposition 2 to the case where alternatives can appear multiple times in a list.

PROOF. It is easy to verify that any choice function $D_{\succsim,\delta}$ satisfies *PI*.

In the other direction, let D be a choice function from lists that satisfies *PI*. By Proposition 1, D also satisfies *LIIA*. For every $a, b \in X$ we define:

(i) $a \succ b$ if $D(a,b) = D(b,a) = a$,

(ii) $a \sim_1 b$ if $D(a,b) = a$ and $D(b,a) = b$, and

(iii) $a \sim_2 b$ if $D(a,b) = b$ and $D(b,a) = a$.

For every $a, b \in X$ exactly one the following relations holds: $a \succ b$, $b \succ a$, $a \sim_1 b$ or $a \sim_2 b$. Therefore, $\succ$ is asymmetric. By definition, $\sim_1$ and $\sim_2$ are symmetric.

Let us go through the following series of simple claims.

CLAIM 1. *The relation $\succ$ is transitive.*

PROOF OF CLAIM. Assume $a \succ b$ and $b \succ c$. By *PI* we have $D(a,b,c) = D(a,D(b,c)) = a$ and by *LIIA* we have $D(a,c) = a$. Using the same reasoning, we also have $D(c,b,a) = a$ and $D(c,a) = a$. Thus, $a \succ c$.

CLAIM 2. *The relations $\sim_1$ and $\sim_2$ are transitive.*

PROOF OF CLAIM. We prove the claim for $\sim_1$ (a similar proof can be made for $\sim_2$). Assume that $a \sim_1 b$ and $b \sim_1 c$. By *PI*, $D(a,b,c) = a$ and $D(c,b,a) = c$. Therefore, by *LIIA* we have $D(a,c) = a$ and $D(c,a) = c$. Thus, $a \sim_1 c$.

CLAIM 3. *(i) If $a \succ b$ and $b \sim_1 c$, then $a \succ c$.*

(ii) If $a \sim_1 b$ and $b \succ c$, then $a \succ c$.

(iii) If $a \sim_2 b$ and $b \succ c$, then $a \succ c$.

(iv) If $a \succ b$ and $b \sim_2 c$, then $a \succ c$.

PROOF OF CLAIM. Let us prove (i). Assume $a \succ b$ and $b \sim_1 c$. Then, by *PI* we have $D(a,b,c) = D(a,D(b,c)) = a$ and $D(b,c,a) = D(D(b,c),a) = a$. Hence, by *LIIA*, $D(a,c) = D(c,a) = a$ and thus $a \succ c$. The other parts of the claim are proved in a similar way.

CLAIM 4. *It is impossible to have both $a \sim_1 b$ and $b \sim_2 c$.*

PROOF OF CLAIM. Assume to the contrary that $a \sim_1 b$ and $b \sim_2 c$. We cannot have $a \sim_1 c$ since the symmetry and transitivity of $\sim_1$ (see Claim 2) would then imply that $b \sim_1 c$, which is a contradiction to $b \sim_2 c$. Similarly, it is impossible that $a \sim_2 c$. By Claim 3, neither $a \succ c$ nor $c \succ a$ is true. Thus, a and c do not relate to each other by either $\succ$, $\sim_1$ or $\sim_2$, which is a contradiction.

Let us define $a \succsim b$ if $a \succ b$ or $a \sim_1 b$ or $a \sim_2 b$. By the above claims, the binary relation $\succsim$ is a preference relation (transitive and complete). If $a \sim b$ (namely, both $a \succsim b$ and $b \succsim a$), then either $a \sim_1 b$ or $a \sim_2 b$. By Claim 4, every $\succsim$-indifference set, $I \subseteq X$, is characterized by the fact that all its members relate to each other by either $\sim_1$ or $\sim_2$ (but not both). If all the members of I are related by $\sim_1$, we define $\delta(a) = 1$ for all $a \in I$; otherwise, we define $\delta(a) = 2$.

Let $L = (a_1,\dots,a_K)$ be a list and denote $D(L) = a_i$. By *LIIA* we have $a_i = D(a_i,a_j)$ for every $j > i$ and $a_i = D(a_j,a_i)$ for every $j < i$. Thus, a_i is a $\succsim$-maximal element of $S(L)$. If there is more than one $\succsim$-maximal element, then by *LIIA* the element $D(L)$ is also chosen from the list L^* that is obtained from L by eliminating all the non $\succsim$-maximal elements.

Since all the members in $S(L^*)$ have the same δ-value, by applying *PI* we obtain that $D(L^*)$ is either the first or the last element in L^* according to the δ-value of the elements in L^*. Thus, $D(L) = D(L^*)$ is also the first or the last $\succsim$-maximal element in L which means that $D = D_{\succsim,\delta}$.

Finally, the preference relation and the priority indicator are unique since $(\succsim,\delta) \neq (\succsim',\delta')$ implies that $D_{\succsim,\delta} \neq D_{\succsim',\delta'}$. □

A choice function from lists is *rationalizable* if there exists a strict preference relation $\succ$ over X such that for any list L, $D(L) \succ a'$ for all $a' \in S(L)\setminus\{D(L)\}$. Of course, a choice function characterized by Proposition 2 is rationalizable if and only if it satisfies the following property.

Order Invariance (*OI*): We say that a choice function D satisfies property *OI* if $D(a_1,a_2,\dots,a_K) = D(a_{\sigma(1)},a_{\sigma(2)},\dots,a_{\sigma(K)})$ for any list $(a_1,a_2,\dots,a_K)$ and any permutation σ of $\{1,\dots,K\}$.

4. Choice correspondences

The literature distinguishes between choice functions that assign a single element to every choice set and choice correspondences that assign to every choice set a non-empty subset (not necessarily a singleton) of elements. If choice is sensitive to order, one may be interested in choice correspondences that attach to every set of alternatives all the elements that are chosen for *some* ordering of that set. One might interpret such choice correspondences as a summary of the information available to a researcher who knows that the decision maker chooses from lists although he does not observe the order of the elements. For example, when a student chooses a graduate school, it may be the case that the set of available schools is known to an observer though the actual order in which the student evaluates the different schools is unobserved.

Formally, given a set A and an ordering O of X, let $L(A,O)$ be the list in which the elements of A are ordered according to O. Let D be a choice function from lists. Define the induced choice correspondence C_D by

$$C_D(A) = \{\, a \mid \text{there exists an ordering } O \text{ for which } D(L(A,O)) = a\}.$$

Our aim is to connect between the properties of a given choice function from lists D and the induced choice correspondence C_D. The following proposition links the property of Partition Independence in the context of choice from lists to the Weak Axiom of Revealed Preference (*WARP*) in the context of choice correspondences. (A choice correspondence C satisfies *WARP* if for every A, $B \subseteq X$, and a, $b \in A\cap B$, if $a \in C(A)$ and $b \in C(B)$, then $a \in C(B)$.)

Proposition 3. *(i) Let D be a choice function from lists. If D satisfies PI, then C_D satisfies WARP. (ii) Let C be a choice correspondence. If C satisfies WARP, then there exists a choice function from lists D that satisfies PI, such that $C_D = C$.*

PROOF. (i) Assume D satisfies *PI*. By Proposition 2, $D = D_{\succsim,\delta}$ for some preference relation $\succsim$ and a priority indicator δ. Fix two sets A, $B \subseteq X$ and two elements $a, b \in A \cap B$. Let $a \in C_D(A)$ and $b \in C_D(B)$. We need to show that $a \in C_D(B)$. Since $a \in C_D(A)$, it belongs to the set of $\succsim$-maximizers in A, which means that $a \succsim b$. Since $b \in C_D(B)$, b belongs to the set of $\succsim$-maximizers in B. Thus, $a \sim b$ and both are $\succsim$-maximizers in B. Assume w.l.o.g. that $\delta(a) = \delta(b) = 1$. Let O_a be an ordering that lists a first. Then $D(L(B, O_a)) = a$ and therefore $a \in C_D(B)$.

(ii) It is well established that if a choice correspondence C is defined over all subsets of X of up to three elements and satisfies *WARP*, then there exists a preference relation $\succsim$ over X such that for every set $A \subseteq X$, $C(A)$ is the set of $\succsim$-maximizers in A (see Mas-Colell et al. 1995, Chapter 1). Let $\succsim$ denote this preference relation and let $D = D_{\succsim,\delta}$, where $\delta(x) = 1$ for all $x \in X$. Then, D satisfies *PI*. Let $A \subseteq X$. We need to show that $C(A) = C_D(A)$. If $a \in C(A)$, then by definition $D(L(A, O_a)) = a$ and thus $a \in C_D(A)$. If $a \notin C(A)$, then a is not a $\succsim$-maximizer in A and there exists no ordering O for which $D(L(A, O)) = a$. Thus, $a \notin C_D(A)$. □

5. Random choice functions

In some cases, an individual's choice from a set lacks consistency, yet we observe that his choice has a systematic random description. For example, when choosing between two wines, it may be the case that the decision maker chooses each wine half of the time. In this case, we wish to assign probabilities to the different elements in the choice set, where the probability of an element is the likelihood that this element will be chosen from the given set. Formally, a *random choice function* C is a function that assigns to every set of alternatives $A \subseteq X$ a probability measure over A. We denote by $C(A)(a)$ the probability of choosing the alternative a from the set A.

An intuitive reason for randomness in choice from sets is that although the decision maker deterministically chooses from lists, there is an underlying random process that transforms sets into lists. For example, a consumer who wishes to purchase a camera might arbitrarily access one of several online retailers (e.g. according to the results of a search engine or the advice of a friend), who offer the same group of cameras, but list them in a different order.

Formally, let μ be a probability measure over the set of orderings of X. Given a choice function from lists D, we define the random choice function C_D^{μ} by

$$C_D^{\mu}(A)(a) = \mu(\{O \mid D(L(A, O)) = a\}).$$

Underlying the definition of C_D^{μ} is the assumption that the order by which alternatives are listed is independent of the set A. Of course, if D is invariant to the ordering of the alternatives, then $C_D^{\mu}(A)$ is concentrated on only one element of A.

In this section, we define two properties of random choice functions and connect them to the corresponding properties of choice functions from lists. The first can be traced back to Luce and Suppes (1965) (see also Tversky 1972 and Simonson and Tversky 1992). We say that a random choice function C is *monotone* if for every two sets A and A' such that $A' \subseteq A$ and for every element $a \in A'$, we have $C(A)(a) \leq C(A')(a)$.

That is, the probability of an element being chosen can only increase when the set of available alternatives shrinks. While the monotonicity property appears to be robust, our experimental results in Section 7 hint that it can be violated in quite a reasonable setting.

We first characterize all choice functions from lists for which the induced random choice function is monotone for every probability measure over the set of orderings.

PROPOSITION 4. *Let D be a choice function from lists. Then D satisfies PI if and only if for every probability measure μ, C_D^{μ} is monotone.*

PROOF. Assume D satisfies *PI* and hence $D = D_{\succsim,\delta}$ for some preference relation $\succsim$ and some priority indicator δ. Let A and A' be two sets such that $A' \subseteq A \subseteq X$, and let $a \in A'$. Let μ be a probability measure on the set of all orderings. For every ordering O, $L(A', O)$ is a sublist of $L(A, O)$ and hence if $D_{\succsim,\delta}(L(A, O)) = a$ then $D_{\succsim,\delta}(L(A', O)) = a$. Therefore, $\{O \mid D(L(A, O)) = a\} \subseteq \{O \mid D(L(A', O)) = a\}$ which implies that $C_D^{\mu}(A)(a) \le C_D^{\mu}(A')(a)$.

In the other direction, assume that for every degenerate probability measure μ (that is, a measure that assigns probability 1 to a single ordering) C_D^{μ} is monotone. By Proposition 1, we need to show that D satisfies *LIIA*. Let $L = (a_1, \ldots, a_K)$ be a list and L' be a sublist of L containing $a = D(L)$. Let μ be the probability measure that assigns probability 1 to an ordering O in which $a_i \, O \, a_{i+1}$ for all $1 \le i \le K-1$ (where $a_i \, O \, a_{i+1}$ means that a_i is ranked higher than a_{i+1} in the ordering O). It follows that $C_D^{\mu}(S(L))(a) = 1$. The monotonicity of C_D^{μ} implies that $C_D^{\mu}(S(L))(a) \le C_D^{\mu}(S(L'))(a)$ and thus $C_D^{\mu}(S(L'))(a) = 1$ which means that $D(L') = a$. □

The second property of random choice functions that we investigate is the preservation of inequalities. We say that a random choice function C *preserves inequalities* if for any set $A \subseteq X$ and for every $a, b \in A$, either

(i) $C(A)(a) = C(A)(b) = 0$ or

(ii) $C(A)(a) \ge C(A)(b)$ if and only if $C(\{a, b\})(a) \ge C(\{a, b\})(b)$.

For example, Luce (1959) presents a family of random choice functions $C_u(A)(a) = u(a)/\sum_{y \in A} u(y)$, each indexed by a function $u : X \to [0,1]$, that preserve inequalities. Tversky (1972), on the other hand, considers examples in which the preserving inequalities assumption is unlikely to hold. We now characterize all the choice functions from lists whose induced random choice functions preserve inequalities for every probability measure.

PROPOSITION 5. *Let D be a choice function from lists. Then $D = D_{\succsim,\delta}$ for some preference relation $\succsim$ with at most two elements in every indifference set and for some priority indicator δ if and only if C_D^{μ} preserves inequalities for all probability measures μ.*

PROOF. Let $D = D_{\succsim,\delta}$, where the preference relation $\succsim$ has at most two elements in every indifference set and let μ be a probability measure on the set of orderings of X. Let $A \subseteq X$ and $a, b \in A$ such that $a \succsim b$. If a is not a $\succsim$-maximizer in A, then $C_D^{\mu}(A)(a) =$

$C_D^\mu(A)(b)=0$ and inequalities are preserved. If a is a $\succsim$-maximizer in A and b is not, then $C_D^\mu(A)(b)=C_D^\mu(\{a,b\})(b)=0$ and inequalities are preserved. If both a and b are $\succsim$-maximizers in A and $\delta(a)=\delta(b)=1$ (and similarly for the case where $\delta(a)=\delta(b)=2$), then, since there are no other $\succsim$-maximizers in A, $D(L(A,O))=D(L(\{a,b\},O))=a$ iff O orders a before b. Thus, $C_D^\mu(A)(a)=C_D^\mu(\{a,b\})(a)$ and $C_D^\mu(A)(b)=C_D^\mu(\{a,b\})(b)$.

In the other direction, let D be a choice function such that C_D^μ preserves inequalities for all probability measures μ. We first show that D satisfies *LIIA*. Otherwise, there exists a list L and a sublist L', such that $D(L)\in S(L')$ but $D(L)\neq D(L')$. Then C_D^μ does not preserve inequalities for the probability measure that assigns probability 1 to an ordering O that satisfies $L(S(L),O)=L$ (that is, O orders the elements of $S(L)$ as they are listed in L). Thus, by Propositions 1 and 2, $D=D_{\succsim,\delta}$ for some preference relation $\succsim$ and some priority indicator δ. It remains to show that each indifference set of $\succsim$ contains at most two elements. Assume that there are three elements a, b and c such that $a\sim b\sim c$. Let μ be a probability measure that assigns a probability of $\frac{1}{2}$ to each of two orderings O_1 and O_2 that satisfy $a\,O_1\,b\,O_1\,c$ and $c\,O_2\,b\,O_2\,a$. For any two elements x and y in $\{a,b,c\}$, we have $C_D^\mu(\{x,y\})(x)=\frac{1}{2}$. Thus, preserving inequalities implies that $C_D^\mu(\{a,b,c\})(x)=\frac{1}{3}$ for every $x\in\{a,b,c\}$, but since μ has a support of only two orderings, it assigns positive probabilities to at most two elements, which is a contradiction. □

6. Duplication of alternatives

Throughout the paper we have focused on the possibility that the order in which alternatives appear might affect choice while assuming that all the elements in the list are non-identical. It is possible, however, that the number of times an element appears in the list also influences choice. For example, if the elements in the list represent advertisements for various products, then not only the order of the advertisements but also the number of times they are repeated might affect choice. In this section, we extend the model of choice from lists to allow for the possibility that elements appear more than once in a list.

Formally, we define a list to be any finite sequence of elements of X. We denote by $\mathscr{L}^*$ the set of all lists and explore choice functions defined over $\mathscr{L}^*$. The properties *PI* and *LIIA* can be naturally extended to the new setting as follows.

Partition Independence (*PI):** We say that a choice function from lists D (defined over $\mathscr{L}^*$) satisfies property *PI** if for every two lists, $L_1,L_2\in\mathscr{L}^*$, we have

$$D(\langle L_1,L_2\rangle)=D(D(L_1),D(L_2)).$$

List Independence of Irrelevant Alternatives (*LIIA):** We say that a choice function from lists D (defined over $\mathscr{L}^*$) satisfies property *LIIA** if for every list $L\in\mathscr{L}^*$, if a is chosen from L, then adding an instance of element b to L cannot lead to choosing an element other than a or b.

Note that whereas *PI* and *LIIA* are equivalent (see Proposition 1) , *PI** is not equivalent to *LIIA**. It will follow from Proposition 2* that *PI** implies *LIIA**; however, *LIIA** does not imply *PI**. For example, let $X = \{a, b, c\}$ and let D be a choice function that picks the most frequent element in a list (if there is more than one such element, D chooses the first among them). Then, D satisfies *LIIA** but

$$D(a, a, b, b, b, a, a, c, c, c) = a \neq D(b, c) = D(D(a, a, b, b, b), D(a, a, c, c, c)).$$

We now extend Proposition 2, which characterizes all the choice functions from lists that satisfy *PI*, to the setting of choice functions from lists with duplication of alternatives.

PROPOSITION 2*. *A choice function from lists D (defined over $\mathcal{L}^*$) satisfies PI^* if and only if there exists a preference relation $\succsim$ over X and a priority indicator δ such that $D = D_{\succsim,\delta}$.*

PROOF. We focus on the "only if" part of the proof. Let D be a choice function that satisfies PI^*. For a list $L = (a_1, \ldots, a_K)$ define $\text{dup}(L) = K - |S(L)|$. That is, $\text{dup}(L)$ is the maximal number of elements that can be deleted from L without changing the corresponding choice set $S(L)$. By Proposition 2, there exists a preference relation $\succsim$ and a priority indicator δ such that $D(L) = D_{\succsim,\delta}(L)$ for all lists L with $\text{dup}(L) = 0$. We will show by induction on $\text{dup}(L)$ that this continues to hold for any list. Indeed, assume that $D = D_{\succsim,\delta}$ for all lists L with $\text{dup}(L) \leq m$. Let us show that $D = D_{\succsim,\delta}$ for any list $L = (a_1, \ldots, a_K)$ with $\text{dup}(L) \leq m + 1$. Let $a_i = a_j$ for some $i < j$. By *PI** we have $D(L) = D(D(a_1, \ldots, a_i), D(a_{i+1}, \ldots, a_K))$. By the induction assumption, $x = D(a_1, \ldots, a_i)$ and $y = D(a_{i+1}, \ldots, a_K)$ are $\succsim$-maximal, $D(x, y)$ is a $\succsim$-maximal element in $\{x, y\}$, and therefore $D(L)$ is a $\succsim$-maximal element in $S(L)$. Assume $\delta(D(L)) = 1$ (an analogous argument can be made for $\delta(D(L)) = 2$). If $x \succsim y$, then x is the first $\succsim$-maximal element in L and hence $D(L) = D(x, y) = D_{\succsim,\delta}(x, y) = x = D_{\succsim,\delta}(L)$. If $y \succ x$, then y is the first $\succsim$-maximal element in L and therefore $D(L) = D(x, y) = D_{\succsim,\delta}(x, y) = y = D_{\succsim,\delta}(L)$. □

Proposition 2* implies that Proposition 3 continues to hold in the model with duplication of alternatives when we assume *PI**. Indeed, the proof of Proposition 3 uses only the fact that the choice functions in question are characterized by a weak preference relation and a priority indicator. This fact follows from *PI** in the current setting, and from *PI* in the previous setting. In addition, Proposition 2* implies that any choice function that satisfies *PI** and Order Invariance is rationalizable.

7. EXPERIMENTAL FINDINGS

The idea that choice is sensitive to the order of presentation is well established in the literature. Several empirical papers have reported on order effects in panel decisions in contests such as the World Figure Skating Competition (Bruine de Bruin 2005), the International Synchronized Swimming Competition (Wilson 1977), the Eurovision Song Contest (Bruine de Bruin 2005) and the Queen Elisabeth Contest for violin and piano (Glejser and Heyndels 2001). In these contests, the contestants appear sequentially and each judge awards each of them a numerical evaluation. The winner is the participant who receives the largest total number of points. It was found that the last few participants in the contest have an advantage since judges tend to increase the points they award over the course of the sequence.

The experimental literature also discusses order of presentation effects. For example, Houston et al. (1989) and Houston and Sherman (1995) showed that when participants have to choose between two sequentially presented options, they prefer the second option in pairs where the alternatives share several negative features but have unique positive features and prefer the first alternative when the alternatives share several positive features but have unique negative features.

Order effects are also common in strategic interactions. Rubinstein et al. (1996) investigated the behavior of players in a two-person game in which one player "hides" a treasure in one of four places laid out in a row and the other player "seeks" it. They found that both "hiders" and "seekers" favored middle positions over endpoints. Attali and Bar-Hillel (2003) investigated a similar question in the context of multiple choice tests. They found that "test makers and test takers have a strong and systematic tendency for hiding the correct answers—or, respectively, for seeking them—in middle positions". Christenfeld (1995) found that people tend to make a selection from the middle when choosing a product from a grocery shelf, deciding which bathroom stall to use or marking a box in a questionnaire.

Liu and Simonson (2005) examined the behavior of subjects who were asked to make a choice from a list of ten products according to the following procedure (which is similar to Example 5 in Section 2): First, a participant sees two products and has to decide which one to keep as a possible choice for the future. In each of the next eight rounds the participant is presented with an additional item and has to decide whether to keep his choice from the previous round or to replace it with the new option. After seeing ten options, participants are asked to decide whether they want to purchase the "winner". It turns out that among those that made a purchase, the first and the second options on the list were purchased by approximately 14% and 18% of the participants respectively.[1]

We conclude with our modest contribution to these experimental and empirical findings. In many instances, the decision maker lacks the ability (due to time or cognitive constraints) to evaluate all the alternatives and is forced to sample only a few of them and to choose one of the sampled alternatives. The choice of the sampled options is to a large extent an operation of arbitrary choice, but one that apparently has its own regularities.

In a short survey we conducted at Tel Aviv University, students were asked to respond online to the following question: "Imagine that you are the editor of a law journal. There is room for one more article in the next issue of the journal. Seven papers have already survived the editorial process and have been recommended for publication in your journal. The papers differ from each other in subject matter. Unfortunately, you have time to read only three of the papers. Once you have read them, you will have to immediately decide which paper to publish. The seven papers are numbered arbitrarily from 1 to 7. Which three papers will you read?"

[1]Personal communication with Wendy Liu.

The frequencies with which the 131 participants sampled the seven options are listed below:

1	2	3	4	5	6	7
79%	61%	54%	39%	14%	20%	34%

And surprisingly about 83% of the subjects chose one of following eight responses (which constitute 23% of the possible combinations):

1, 2, 3	1, 4, 7	2, 4, 6	2, 5, 7	1, 4, 6	1, 3, 7	1, 3, 6	2, 4, 7
39%	19%	6%	4%	4%	4%	4%	3%

Thus, approximately 40% of the participants in the experiment chose to sample the first three elements and almost 20% chose the triplet that consists of the first, middle and last elements. If this pattern extends to longer lists, one could interpret these results as suggesting that random choices violate the monotonicity property of random choice functions, according to which the addition of an element to a choice problem weakly decreases the probability that the other elements will be chosen. For example, imagine that the alternatives are candidates who are always listed in alphabetical order and that a decision maker samples three of them, evaluates them, and chooses one of them. Assume that the decision maker's choice following his evaluation is consistent with the maximization of a fixed ordering over the candidates, and that the fifth person in alphabetical order is the best candidate according to this ordering. If the list contains the first seven candidates in alphabetical order, candidate 5 will be sampled with very low probability and hence will probably not be chosen. On the other hand, if the set of candidates is extended to include nine candidates, it appears more likely (according to the results of our experiment) that candidate 5, who is now at the center of the list, will be picked for the sample and thus will have a higher probability of being chosen. This contradicts the monotonicity property.

References

Attali, Yigal and Maya Bar-Hillel (2003), "Guess where: the position of correct answers in multiple-choice test items as a psychometric variable." *Journal of Educational Measurement*, 40, 109–128. [15]

Bruine de Bruin, Wandi (2005), "Save the last dance for me: unwanted serial position effects in jury evaluations." *Acta Psychologica*, 118, 245–260. [14]

Christenfeld, Nicholas (1995), "Choices from identical options." *Psychological Science*, 6, 50–55. [15]

Glejser, Herbert and Bruno Heyndels (2001), "Efficiency and inefficiency in the ranking in competitions: the case of the Queen Elisabeth Music Contest." *Journal of Cultural Economics*, 25, 109–129. [14]

Houston, David A. and Steven J. Sherman (1995), "Cancellation and focus: the role of shared and unique features in the choice process." *Journal of Experimental Social Psychology*, 31, 357–378. [15]

Houston, David A., Steven J. Sherman, and Sara M. Baker (1989), "The influence of unique features and direction of comparison on preferences." *Journal of Experimental Social Psychology*, 25, 121–141. [15]

Liu, Wendy and Itamar Simonson (2005), "'Take it or leave it?'—The effect of explicit comparisons on commitment to purchase." *Advances in Consumer Research*, 32, 453–454. [15]

Luce, R. Duncan (1959), *Individual Choice Behavior: A Theoretical Analysis*. Wiley, New York. [12]

Luce, R. Duncan and Patrick Suppes (1965), "Preference, utility, and subjective probability." In *Handbook of Mathematical Psychology* (R. Duncan Luce, Robert R. Bush, and Eugene Galanter, eds.), volume 3, 249–410, Wiley, New York. [11]

Mas-Colell, Andreu, Michael D. Whinston, and Jerry R. Green (1995), *Microeconomic Theory*. Oxford University Press, New York. [11]

Plott, Charles R. (1973), "Path independence, rationality, and social choice." *Econometrica*, 41, 1075–1091. [7]

Rubinstein, Ariel, Amos Tversky, and Dana Heller (1996), "Naïve strategies in competitive games." In *Understanding Strategic Interaction: Essays in Honor of Reinhard Selten* (Wulf Albers, Werner Güth, Peter Hammerstein, Benny Moldovanu, and Eric van Damme, eds.), 394–402, Springer-Verlag, Berlin. [15]

Salant, Yuval (2003), "Limited computational resources favor rationality." Discussion Paper 320, Center for the Study of Rationality, Hebrew University of Jerusalem. [6]

Simon, Herbert A. (1955), "A behavioral model of rational choice." *Quarterly Journal of Economics*, 69, 99–118. [5]

Simonson, Itamar and Amos Tversky (1992), "Choice in context: tradeoff contrast and extremeness aversion." *Journal of Marketing Research*, 29, 281–295. [11]

Tversky, Amos (1972), "Elimination by aspects: a theory of choice." *Psychological Review*, 79, 281–299. [11, 12]

Tversky, Amos and Daniel Kahneman (1991), "Loss aversion in riskless choice: a reference-dependent model." *Quarterly Journal of Economics*, 106, 1039–1061. [4, 5]

Wilson, Vietta E. (1977), "Objectivity and effect of order of appearance in judging of synchronized swimming meets." *Perceptual and Motor Skills*, 44, 295–298. [14]

Submitted 2005-5-19. Final version accepted 2005-8-2. Available online 2005-12-2.