Choice by lexicographic semiorders

Paola Manzini

University of St. Andrews and IZA

Marco Mariotti*

University of St. Andrews

July 2011

Abstract

In Tversky's [33] model of a lexicographic semiorder, a preference is generated by the sequential application of numerical criteria, by declaring an alternative x better than an alternative y if the first criterion that distinguishes between x and y ranks x higher than y by an amount exceeding a *fixed* threshold. We generalise this idea to a fully-fledged model of boundedly rational choice. We explore the connection with sequential rationalisability of choice ([1], [21]), and we provide axiomatic characterisations of both models in terms of observable choice data.

J.E.L. codes: D0.

Keywords: Lexicographic semiorders, bounded rationality, revealed preference, choice.

1

1   Introduction

Lexicographic heuristics have gained much attention in the study of decision making, in several fields: in psychology (e.g. Tversky [33], [34]; Gigerenzer and Todd [9]); in positive economics (e.g. Rubinstein [28]; Leland [14]; Manzini and Mariotti [21]; Apesteguia and Ballester [1]); in normative economics (e.g. Tadenuma [31], [32]; Houy and Tadenuma [11]); in marketing science (e.g. Yee, Dahan, Hauser and Orlin [36]; Kohli and Jedidi [13]). Medina, Naeh and Segal [26] note that the Talmud contains arguments in favour of a lexicographic ranking of the rationales used to adjudicate between pairs of alternatives.

The explanation for this success is obvious: lexicographic procedures look appealingly simple and realistic since they eschew the complex trade-offs between several criteria of classical decision makers. On the other hand, the lack of trade-offs may also seem to constitute a disadvantage (especially among economists). Price may be the most important criterion in the purchase of a house from a set of suitable ones. Yet who would be prevented by a difference of a few bucks from selecting a house in a much more desirable neighbourhood? Arguably, very few people would be so uncompromising as to ignore any significant improvement in one dimension because of an arbitrarily small loss in the most important dimension. When modelling boundedly rational behaviour, the rigid application of simple 'rules of thumb' (such as 'buy the cheapest house among the acceptable ones') may look even less realistic than the trade-offs of textbook utility maximisation.

In other words, it is reasonable that, even in a boundedly rational heuristic, criteria that detect significant differences between the alternatives under consideration should over-ride criteria that do not. In this paper we study a model of choice that formalises this intuition. Note that a number of 'basic criteria' could be aggregated into a single, more complex criterion, to which our observations on the house buyer above would nevertheless still apply: if the agent constructs an index which trades off price and location, that index constitutes a new criterion, for which it may be unwise not to ignore small differences in favour, say, of house size, and so on.[1] Only a fully rational decision maker would be able

---

[1] As another example, in Manzini and Mariotti [20] we have proposed a multi-criterion model of choice over time in which the first criterion is the exponentially discounted value, which trades off the time and size of a monetary reward.

to pack together all possible trade-offs in a single criterion. However, in a more realistic model of decision making, there is a limit to the number of simultaneous trade-offs the decision maker is able to carry out. Thus, it seems plausible to expect the decision maker to rely on a lexicographic list of 'slack' criteria. The choice procedure we propose can explain observed 'anomalies', while at the same time preserving a convincing flexibility.

Considerations of this kind have already led some of the researchers mentioned above[2] to build models of *preference or binary choice* based on the application of numerical criteria where small differences in the values of criteria are ignored.[3] However, such models leave unanswered the issue of *choice* from more complex sets (e.g., budget sets). They do not study choice functions. If binary preferences are derived from a boundedly rational procedure, the issue of associating such preferences with higher order choices is far from trivial: on the one hand it may be impossible to maximise the preference (when it is cyclical); and on the other hand it may be inappropriate to even consider maximisation when the issue is one of bounded rationality.

We focus on Tversky's [33] fruitful notion of *lexicographic semiorder*, in which a preference is generated by the sequential application of numerical criteria, by declaring an alternative x better than an alternative y if the first criterion that distinguishes between x and y ranks x higher than y by an amount exceeding a *fixed* threshold. Our first contribution is to define a choice procedure (choice by lexicographic semiorder) based on Tversky's idea.

Tversky himself considered lexicographic semiorders appealing but restrictive as a model of preference.[4] In fact, this judgement is shown to be somewhat pessimistic. Even when the agent is endowed with very rudimental discriminatory abilities (being only able to classify criteria values in 'good', 'neutral' and 'bad', where just 'good' and 'bad' are rankable), the model can account for a very rich variety of behaviours. In particular, when only binary choices are involved (as for example in several voting models[5]) the model is

---

[2]Tversky [33], Rubinstein [28], Leland [14]

[3]A difference being small is often interpreted as 'similarity'.

[4]See Section 3.

[5]See e.g. Kalandrakis [12] and the references therein for a recent example. We discuss Kalandrakis' work more fully in section 3.1.

shown to be *completely* unrestrictive, provided that the set of alternatives is not too large (proposition 1). More in general, the model can explain any set of choice data satisfying WARP (proposition 2) - since such choices may not satisfy SARP, they may be highly 'irrational' in that they exhibit strict revealed preference cycles.

The model turns out to be connected with another, much more general-looking, notion of boundedly rational choice, namely 'sequentially rationalisable choice' (Manzini and Mariotti [21]): an arbitrary number of arbitrary asymmetric binary relations ('rationales') is applied sequentially to single out an alternative. On any finite domain,[6] the lexicographic semiorder model restricts choice data the same way as the sequential rationalisability model under the additional assumption that the rationales used in the latter are acyclic (proposition 3).

We note, however, that the clause 'on any finite domain' is key. When this clause is relaxed even marginally, by allowing a countably infinite number of *finite* choice sets, the equivalence breaks down in a major way: even the use of only two rationales may produce behaviours that cannot be generated by any number of semiorders and any number of discriminations (proposition 4). So, the two models are in general clearly distinct.

Next, we characterise choice by lexicographic semiorders in terms of a new contraction consistency condition (Reducibility), at the same time providing an algorithm to construct the semiorders (theorem 1).

Our technique leads straightforwardly to a relaxation of Reducibility which characterises sequential rationalisability (theorem 2). This result, while quite tangential to the main line of enquiry of this paper, is of independent interest, since the characterisation of sequential rationalisability has proved to be a hard problem which we left open in [21]. Our results in this respect build on and complement those by Apesteguia and Ballester [1], who were the first to draw attention to the restriction of sequential rationalisability to acyclic rationales and to provide a characterisation for it on finite domains. In the Appendix we work out one of their examples of sequentially rationalisable choices to construct the rationales with our algorithm. Our work can also be fruitfully seen as an extension of the approach in Mandler, Manzini and Mariotti [19]: we discuss this relation in the

---

[6]That is a domain including a finite number of finite sets.

concluding section.

2 Lexicographic semiorders: preferences and choice

Fix a nonempty set X. A *semiorder* (Luce [16]) is an irreflexive[7] binary relation P on X which satisfies two additional properties:

1. $(x, y), (w, z) \in P$ imply $(x, z) \in P$ or $(w, y) \in P$;

2. $(x, y) \in P$ and $(y, z) \in P$ imply $(x, w) \in P$ or $(w, z) \in P$.

 Given the irreflexivity of P, each of (1) or (2) imply that P is also transitive.[8] So a semiorder is a very special type of strict partial order. The interest of semiorders is that they can be interpreted as a simple threshold model of (partial) rankings: on suitable domains, P is a semiorder if and only if there exists a real valued function f on X and a number $\sigma \geq 0$ such that $(x, y) \in P$ if and only if $f(x) > f(y) + \sigma$. Here $f(x)$ is the 'value' of the alternative x and σ is the amount by which the value of one alternative x must exceed the value of another alternative y for x to be declared superior to y. The fact that σ is fixed makes this a very parsimonious model of binary preferences.[9]

 Tversky [33] essentially proposed a lexicographic procedure, which extends the use of semiorders, to make binary comparisons between alternatives in a set X. There exists an ordered sequence $f = (f_1, ..., f_n)$ of real valued functions on X and a $\sigma > 0$ such that x is declared better than y iff, for the first i for which $|f_i(x) - f_i(y)| > \sigma$, we have $f_i(x) > f_i(y) + \sigma$. The idea is that the agent compares alternatives along several dimensions. As in our opening example, dimensions are ranked in order of importance, and a later dimension is only considered if all previous dimensions failed to discriminate between the two alternatives under consideration. In other words, the agent examines the dimensions *lexicographically*: as soon as a dimension i is found for which one alternative x

---

[7]Irreflexivity: for all $x \in X$, $(x, x) \notin P$.

[8]Transitivity: for all $x, y, z \in X$, $(x, y) \in P$, $(y, z) \in P \Rightarrow (x, z) \in P$.

[9]In an *interval order* (Fishburn [7]), characterised by condition 1 alone, the threshold σ is allowed to vary with the alternatives being compared, being a function $\sigma : X \to \mathbb{R}_+$. This makes for a much richer structure. See e.g. Fishburn [8].

is superior to another alternative y by an amount exceeding the threshold σ, x is declared better than y. When such an i is found, no dimension j that comes later in the order has any bearing, no matter the size of the differences between the alternatives in these subsequent dimensions. That σ is chosen to be the same for all f_i is not a relevant issue, since even if we had different σ_i, the f_i and σ_i can always be rescaled so as to choose $\sigma_i = 1$. Given f and σ, this procedure can be used to generate a revealed preference relation $\succ_{(f,\sigma)}$ on pairs of alternatives.[10]

Suppose now that the agent wants to apply the procedure to produce a selection out of choice sets S larger than the binary ones. There are several ways to do so, some of which are however problematic. One could for example start from the binary revealed preference relation and use either of the following two plausible methods:

- the choice from S is the set of the maximal elements of $\succ_{(f,\sigma)}$

- the choice from S is the top cycle (or the uncovered set) of $\succ_{(f,\sigma)}$ restricted to each S.[11]

Unfortunately, the preference relation $\succ_{(f,\sigma)}$ may be cyclic - this 'anomalous' feature was indeed the very point of Tversky introducing the procedure. So the first method above may not be well-defined if a nonempty-valued choice function is desired. The second method above borrows the ideas of authors such as Ehlers and Sprumont [5] and Lombardi [15], who use weaker notions of maximization to produce choices out of non-standard preferences formed of asymmetric and complete binary relations (tournaments). These methods would for example select the entire set $S = \{x_1, x_2, ..., x_n\}$ whenever $x_1 \succ_{(f,\sigma)} x_2 \succ_{(f,\sigma)} \cdots \succ_{(f,\sigma)} x_n \succ_{(f,\sigma)} x_1$.

Here we pursue a different natural way of extending Tversky's idea. The method we suggest is, on the one hand, more in line with the procedural (as opposed to maximising)

---

[10]Rubinstein [28] proposes a related but distinct procedure. This procedure has recently been studied experimentally by Binmore, Voorhoeve and Wallace [2].

[11]More precisely, let $P|S$ denote the restriction to S of a complete asymmetric binary relation P defined on X. (Completeness: for all $x, y \in X$ either $(x, y) \in P$ or $(y, x) \in P$. Asymmetry: for all $x, y \in X$, $(x, y) \in P \Rightarrow (y, x) \notin P$). Let $(P|S)^t$ denote the transitive closure of $P|S$. The *top cycle* of P in S is the set of maximal elements of $(P|S)^t$ in S. Define the covering relation $C(P, S)$ of P in S by: $(x, y) \in C(P, S)$ iff $x, y \in S$ and either $(x, y) \in P$ or there exists $z \in S$ such that $(x, z) \in P$ and $(z, y) \in P$. The *uncovered set* of P in S is the set of maximal elements of $C(P, S)$ in S.

nature of Tversky's approach; and, on the other hand, it can produce a unique selection even from the awkward cycles discussed above. The reason for these two features is that the method, unlike the others suggested, preserves and uses the information on the order in which the dimensions are considered.

We impose no arbitrary uniform bound on the number of dimensions that the agent is allowed to consider. Nevertheless, we insist that the procedure always halts in a finite number of steps in any choice situation.

Our proposed procedure works via a process of sequential elimination. Formally, let Σ be a domain of choice sets, where each S in Σ is a nonempty subset of X. A *choice function* on Σ is a function $c : \Sigma \to X$ such that $c(S) \in S$ for all $S \in \Sigma$. A choice set S which has the form $S = \{x\}$ for some $x \in X$ will be called *trivial.* A collection $\mathcal{C} \subseteq \Sigma$ of choice sets is trivial if each $S \in \mathcal{C}$ is trivial.

An ordered sequence $f = (f_i)_{i \in I}$, where I is either an interval of numbers $\{1, ..., n\}$ or the entire set of natural numbers $\mathbb{N}$, together with a $\sigma > 0$ is a *lexicographic semiorder on* X, denoted $(f_1, f_2, ..., \sigma) = (f_i, \sigma)_{i \in I}$. We abuse terminology and call each f_i directly a semiorder although strictly speaking f_i is a numerical representation of it.

Given a choice set $S \subseteq X$ and a lexicographic semiorder $(f_i, \sigma)_{i \in I}$, define inductively the following 'survivor sets' $M_i(S)$, for all $i > 0$:

$$M_0(S) = S$$
$$M_i(S) = \{s \in M_{i-1}(S) \,|\, \forall s' \in M_{i-1}(S) \ \ f_i(s) + \sigma \geq f_i(s')\}$$

This sequence of sets captures the procedure the agent follows in order to arrive at a final selection from the choice set S: at every round i he looks for alternatives in the current survivor set $M_{i-1}(S)$ which are judged 'worse' than some other alternative in $M_{i-1}(S)$ according to the Tversky procedure described before. He discards all such inferior alternatives (if any), generating the next survivor set $M_i(S)$, and so on.

Definition 1 *A choice function c is a **choice by lexicographic semiorder (cles)** iff there exists a lexicographic semiorder $(f_i, \sigma)_{i \in I}$ such that, for all $S \in \Sigma$, there is a $j \in I$ for which $\{c(S)\} = M_j(S) = M_k(S)$ for all $k \geq j$.*

*In this case we say that $(f_i, \sigma)_{i \in I}$ **induces** c.*

That is, for a cles c, the iterative elimination procedure described before stops on any choice set S after a finite number of steps, yielding precisely the alternative that c picks in S. Note that, in spite of this property of 'finite termination', there might not exist any fixed j that works for all S. When such a j exists, which means that I can be chosen to be finite, we say that c is a choice by finite lexicographic semiorder.[12]

Basic Semiorders

A semiorder f_i is *basic* if it ranges only in $\{-1, 0, 1\}$ and $\sigma = 1$. A lexicographic semiorder $(f_i, \sigma)_{i \in I}$ is basic if each f_i is basic. So, with a basic lexicographic semiorder the agent has only a very limited power of discrimination. Essentially, on each dimension he can only perform a rough classification of alternatives into 'good' ones (those x for which $f_i(x) = 1$), 'bad' ones ($f_i(x) = -1$), and 'neutral ones' ($f_i(x) = 0$): a good alternative 'beats' a bad one (on the given dimension), and a neutral alternative neither beats a bad one nor is beaten by a good one.

A basic lexicographic semiorder can be denoted simply as $f = (f_i)_{i \in I}$. To emphasise that the survivor sets $M_i(S)$ are obtained from the basic lexicographic semiorder f we write them as $M_i^f(S)$.

Example: Let $X = \{x, y, z\}$ and let $\Sigma = \{\{x, y\}, \{y, z\}, \{z, x\}, X\}$. Let $c(\{x, y\}) = c(X) = x$, $c(\{y, z\}) = y$ and $c(\{x, z\}) = z$. This is a choice function by basic lexicographic semiorder. To see this, let $f_1(x) = 0$, $f_1(y) = 1$, $f_1(z) = -1$, $f_2(x) = 1$, $f_2(y) = -1$, $f_2(z) = 0$, $f_3(x) = -1$, $f_3(y) = 1$, $f_3(z) = 1$. Observe how different (unique) choices from X can be obtained by permuting the order of the f_i.

3   Characterisation

Tversky thought that the model of binary choice by lexicographic semiorders, while useful to explain the anomaly of cyclical preferences, had a narrow scope otherwise. He writes:

---

[12] Aside from this twist, we could also have called the model 'semiorder sequentially rationalisable choice', following the terminology we initiated in Manzini and Mariotti [21]. However, we prefer to use the Tversky terminology, in recognition of the priority of his idea.

" ... despite its intuitive appeal, it is based on a noncompensatory principle that is likely to be too restrictive in many contexts." (Tversky [33], p. 40).

Following this logic, one might conjecture that the version with basic semiorders, with its minimal concession to discriminatory powers, is even more restrictive. We study this issue, highlighting the role of the domain of c.

We begin by observing that, when restricted to binary choices (interpretable as possibly incomplete preferences) as in the original application, the cles model is in fact completely unrestrictive provided that the set X is not too large. We state this result separately because of its interest, although it is a particular case of the more general proposition 2 below.

Proposition 1 *Let c be defined on a domain Σ such that $S \in \Sigma$ implies $S = \{x, y\}$ for some distinct $x, y \in X$. Let X be countable. Then there exists a basic lexicographic semiorder which induces c.*

As noted, more generally all choice data satisfying a classical revealed preference axiom could be generated by a cles:

WARP: If $x = c(S)$, $y \in S$ and $y = c(T)$ for some $S, T \in \Sigma$ then $x \notin T$.

Proposition 2 *Let c satisfy WARP. Let X be countable. Then there exists a basic lexicographic semiorder which induces c.*

Proof. Enumerate the elements in X with a bijection b from $\mathbb{N}$ (or an interval of $\mathbb{N}$ of cardinality equal to $|X|$ if X is finite) to X. Define a basic lexicographic semiorder as follows. For all $x, y \in X$, let

$$f_{b^{-1}(y)}(x) = \begin{cases} 1 & \text{if} & x = y \\ -1 & \text{if} & \exists S \in \Sigma \text{ such that } x \neq y = c(S) \text{ and } x \in S \\ 0 & \text{otherwise} \end{cases}$$

Let $c(S) = x \neq y \in S$. WARP and the definition of f imply that, for all $z \in X$ with $b^{-1}(z) < b^{-1}(x)$, $f_{b^{-1}(z)}(x) = 0$ if $z = y$, and that $f_{b^{-1}(z)}(y) \leq 0$ whenever $f_{b^{-1}(z)}(x) = -1$. Therefore $x \in M_i(S)$ for all $i < b^{-1}(x)$. And since $f_{b^{-1}(x)}(x) = 1$ and $f_{b^{-1}(x)}(y) = -1$ for all $y \in S$ with $y \neq x$, $\{x\} = M_{b^{-1}(x)}(S) = M_k(S)$ for all $k \geq b^{-1}(x)$. ∎

Because WARP does not imply SARP on general domains,[13] a cles can explain cyclical patterns of strict revealed preference, generalising example 2.

In order to pinpoint the restrictions on behavior implied by the cles, we recall some definitions.

Definition 2 *A choice function c is **sequentially rationalisable** whenever there exists an ordered list $\{P_i\}_{i \in I}$ of asymmetric relations, with $P_i \subseteq X \times X$ for $i \in I$, such that, defining recursively*

$$M_0^*(S) = S$$
$$M_i^*(S) = \left\{ x \in M_{i-1}^*(S) \,|\, \forall y \in M_{i-1}^*(S) \ \ (y,x) \notin P_i \right\}$$

for all $S \in \Sigma$ there is a $j \in I$ such that

$$\{c(S)\} = M_j^*(S) = M_k^*(S) \ \text{ for all } k \geq j$$

*In that case we say that $\{P_i\}_{i \in I}$ sequentially rationalise c. Each P_i is a **rationale**.*[14]

Two specialisations of sequential rationalisability are the following:

Definition 3 *A choice function is a **Rational Shortlist Method** (RSM) iff it is sequentially rationalisable with two rationales. A choice function is **acyclic sequentially rationalisable** iff it is sequentially rationalisable by rationales that are acyclic.*

Sequentially rationalisable choice functions and RSMs were defined in Manzini and Mariotti [21]. The restriction to acylic rationales for the finite case has been studied by Apesteguia and Ballester [1]. Evidently, the cles model we are considering in this paper is a restriction of sequential rationalisability by constraining the rationales to be semiorders. Both acyclic and standard sequential rationalisability constitute at first sight a much more

---

[13]SARP says that the revealed preference relation P_c, given by $xP_cy \Leftrightarrow \exists S \in \Sigma : x = c(S), y \in S$, is acyclic.

[14]This definition slightly extends the one we originally gave in [21], and has the same format of the definition of a lexicographic semiorder. In the original definition we considered a finite ordered list $P_1, ... P_K$ of asymmetric relation with the $\{c(S)\} = M_K^*(S)$ for all $S \in \Sigma$. While still imposing finite termination on each choice set, the current definition disposes with the assumption that there exists a *uniform* bound K on the number of rationales needed to rationalise a given choice function.

general model, because the rationales are not required to have any threshold structure and can thus apparently accommodate more sophisticated discriminations. But in fact, for arbitrary finite domains, the behaviours that can be generated by the lexicographic semiorder model and those that can be generated by the acyclic sequential rationalisability model are just the same. And, we need look no further than *basic* semiorders to yield this equivalence.

On the other side of the coin, the restriction to finite domains is not merely a convenience for the inductive argument used in the proof, but it is necessary for the equivalence to hold. When the restriction is relaxed even marginally (by retaining the finiteness of each choice set but allowing for a countable number of choice sets), the model of acyclic sequential rationalisability suddenly appears far more general than the lexicographic semiorder model: even only two acyclic rationales suffice to produce behaviours that cannot be induced by any basic lexicographic semiorder. And increasing the discriminatory ability of the agent is to no avail: the 'basic' restriction is inessential for this result.

These assertions are made precise in the next two results. In the case of the domain consisting of all finite subsets, the first result can also be derived from theorem B.1 in [1], as we explain below. We present here a different method of proof which highlights the importance of the domain and is instructive in this respect.

Proposition 3 *Let X be finite. Then a choice function c is acyclic sequentially rationalisable if and only if it is induced by a basic lexicographic semiorder.*

Proof. A semiorder is an acyclic rationale, so it suffices to prove the 'only if' part of the statement. Given acyclic rationales $(P_1, ..., P_K)$, recall the definition 2 of survivor sets $M_i^*(S)$. We will show that, for any domain Σ, there exists a basic lexicographic semiorder $f = (f_i)_{i \in I}$ such that, for all $S \in \Sigma$, there is a $j \in I$ such that $M_K^*(S) = M_j^f(S) = M_k^f(S)$ for all $k \geq j$. This proves the assertion in the statement.

The proof is by induction on the sum of the cardinalities of the sets S in Σ, which we denote by $n(\Sigma) = \sum_{S \in \Sigma} |S|$. If $n(\Sigma) = 1$ the claim is obviously true. Take now $n(\Sigma) > 1$. If Σ is trivial, then the claim is also obviously true, so assume Σ is not trivial, and w.l.o.g. assume in addition that P_1 is nonempty on some $S \in \Sigma$ (otherwise just exclude P_1 and renumber the remaining P_i). By the acyclicity of P_1 and the finiteness of

11

X there exist $S \in \Sigma$ and $x, y \in S$ such that $(x, y) \in P_1$ and $(y, z) \notin P_1$ for all $z \in \bigcup_{S \in \Sigma} S$ with $y, z \in T$ for some $T \in \Sigma$ (in words, y is P_1-dominated in some choice set and it does not P_1-dominate any element which appears together with y in any choice set). Fix those x and y, and define

$$\Sigma' = \{S : \{x, y\} \nsubseteq S \in \Sigma\} \cup \{S : S = T \backslash \{y\} \text{ for some } T \in \Sigma \text{ s.t. } \{x, y\} \subseteq T\}$$

Because a T as in the right-hand member of the union above exists by construction, $n(\Sigma') < n(\Sigma)$. So by the inductive hypothesis there exists a basic lexicographic semiorder $f = (f_i)_{i \in I}$ such that, for all $S \in \Sigma'$, there is a $j \in I$ such that $M_K^*(S) = M_j^f(S) = M_k^f(S)$ for all $k \geq j$. Now consider the basic lexicographic semiorder $g = (g_i)_{i \in I'}$ defined by

$$g_i = f_{i-1} \text{ for all } i > 1$$

$$g_1(x) = 1, g_1(y) = -1 \text{ and } g_1(z) = 0 \text{ for all } z \neq x, y$$

Thus, for all $S \in \Sigma$ such that $\{x, y\} \subseteq S$, $M_1^g(S) = S \backslash \{y\} \in \Sigma'$ and consequently $M_K^*(S \backslash \{y\}) = M_{j+1}^g(S) = M_k^g(S)$ for all $k \geq j+1$ (this follows by the second line of the displayed definition of g and the fact that $M_K^*(S \backslash \{y\}) = M_j^f(S \backslash \{y\}) = M_k^f(S \backslash \{y\})$ for all $k \geq j$). Moreover, clearly for all $S \in \Sigma$ such that $\{x, y\} \subseteq S$, $M_K^*(S) = M_K^*(S \backslash \{y\})$. Therefore, for all $S \in \Sigma$, $M_K^*(S) = M_K^*(S \backslash \{y\}) = M_{j+1}^g(S) = M_k^g(S)$ for all $k \geq j+1$.∎

Proposition 4 *There exist Rational Shortlist Methods using acyclic rationales which are not induced by any lexicographic semiorder.*

Proof. Let $X = \{1, 2...\}$, let Σ be the collection of finite subsets of X, and let c be uniquely defined as the RSM rationalised by the following two acyclic rationales P_1 and P_2:

$$P_1 = \{(i, i+1) : i \in X\}$$

and

$$P_2 = \{(j, i) : j > i+1\}$$

We show that c is not induced by any lexicographic semiorder. By contradiction, suppose that $(f_\alpha, \sigma)_{\alpha \in I}$ is a lexicographic semiorder which induces c. Let $i, j \in X$ be such

12

that $f_1(j) > f_1(i) + \sigma$. Such an i and j exists w.l.o.g., possibly by renumbering the f_α so that f_1 is the first f_α for which $f_1(k') > f_1(k) + \sigma$ for some $k, k' \in X$. Also, note that $i \neq 1$ since the application of the rationales yields $c(\{1, 2, .., l\}) = 1$ for all $l \in X$. It must be $j = i - 1$ (that is, i is eliminated by $i - 1$ in the first step in any set that contains both of them). Otherwise suppose first that $j > i$. Then $c(\{i, i+1, i+2, ..., j\}) = i$ would be contradicted by $i \notin M_1(\{i, i+1, i+2, .., j\})$. Alternatively, suppose that $j < i - 1$. Then $c(\{j, i\}) = i$ would be contradicted by $i \notin M_1(\{j, i\})$.

Thus, $f_1(i-1) > f_1(i) + \sigma$. Since $c(\{i-1, i+1\}) = i+1$, it must be that, letting n be the first α for which $M_\alpha(\{i-1, i+1\}) \neq \{i-1, i+1\}$, we have $f_n(i+1) > f_n(i-1) + \sigma$. Applying this fact to $S = \{i-1, i, i+1\}$, we have that if $n = 1$ then $M_1(S) = \{i+1\}$, contradicting $c(S) = i - 1$. If instead $n > 1$, then either $f_1(i) > f_1(i+1) + \sigma$, in which case we also have $f_1(i-1) > f_1(i+1) + \sigma$, contradicting $c(\{i-1, i+1\}) = i+1$; or this is not the case, so that $c(S) = c(M_1(S)) = c(\{i-1, i+1\}) = i+1$. In both cases we have a contradiction with $c(S) = i - 1$.[15] ∎

Some observations are in order. Apesteguia and Ballester [1] define a *simple rationale* P as a relation of the type $P = \{(x, y)\}$ for some x and y in X. That is, a simple rationale relates only one pair of alternatives. Our notion of 'basic' refers instead to the number of discriminations the agent is able to make, rather than to the number of pairs ranked by the relation (which may be high). However, [1] show that, for the case of the domain consisting of all subsets of a finite set X, sequential rationalisability with acyclic rationales is equivalent to sequential rationalisability with simple rationales. The ranking made by a simple rationale $P = \{(x, y)\}$ can be expressed with a basic semiorder (though not vice-versa), by setting $f(x) = 1$, $f(y) = -1$ and $f(z) = 0$ for all other z. Therefore, as observed above, proposition 3 can be derived by their result in the case of full domain.

While a simple rationale can be expressed by means of a single semiorder, there is no upper bound to the number of simple rationales needed to express a basic semiorder. For example, the rationale $P = \{(x, y) : y \in X \setminus \{x\}\}$, for a fixed x, is a single basic semiorder

---

[15] Observe that it is at this step of the proof that the domain assumption bites. For $\{i-1, i, i+1\}$ and $\{i-1, i+1\}$ might not be well-defined if we did not have the entire integer set at our disposal.

for any n, which is nevertheless decomposed into $(n-1)$ distinct simple rationales.[16]

Proposition 4 shows that the domain restriction $|X| < \infty$ of theorem B.1 of [1] is necessary. Their result establishes that, on the domain of all nonempty subsets of a finite set X, the only crucial distinction is between the asymmetry and the acyclicity (a strengthening of asymmetry) of the rationales: further strengthening acyclicity to transitivity, for example, produces no further behavioural restriction. Proposition 4 shows that on larger domains the move from acyclicity to transitivity (semiorders) crosses another important threshold: the transitivity of the agent's discriminatory power alone suffices to rule out behaviours allowed by acyclic rationales. This remains true no matter how limited that power is.

3.1 'Revealed preference' characterisation

How could an external observer establish whether a set of choice data (i.e. a choice function c) could have been generated by the procedure we have proposed? The key to answering this question is to consider the behaviour of c over *restricted domains* of choice, as well as on the domain Σ of definition of c. The method we shall suggest can be viewed an extension of techniques used in standard analysis of rationalisability of choice functions on special domains. For example, in recent work, Kalandrakis [12] studies the rationalisability of a set of *binary* voting choices on Euclidean policy space by means of a quasiconcave utility function. He identifies rationalisability conditions with the following format: for every subdomain of choice $\mathcal{C}$, there exists an 'extreme alternative' (i.e. not obtainable as a convex combination of other alternatives) x such that x is never chosen from choice sets in the collection $\mathcal{C}$. The interpretation is that x is a least preferred alternative among those appearing in the choice problems in $\mathcal{C}$. This permits the ultimate construction of a (quasiconcave) utility function.

---

[16]In recent work, Mandler [18] has studied in detail the general issue of the minimum number of rationales needed to express a given arbitrary preference relation (interpretable as the base relation of a choice function) using the procedure of sequential rationalisability. His main result is that a 'rational agent' (an agent with complete and transitive preferences) never needs more, and sometimes needs fewer, rationales than a non-rational agent.

Let us see how analogous ideas can work in our setting. Because our model does not involve the simple maximisation of preferences, we cannot hope to identify 'least preferred' alternatives. But if the agent were really using our lexicographic procedure, in any subdomain $\mathcal{C}$ we should at least be able to identify alternatives x and y such that x makes y '$\mathcal{C}-irrelevant$': namely, if x and y belong to some S in $\mathcal{C}$, removing y from S has no effect on the final choice from S (so that, in particular y is never chosen if x is available). This alternative y is simply one of the alternatives the agent would eliminate with the first semiorder which he applies on $\mathcal{C}$, say $f_{\mathcal{C}}$, and x is an alternative with $f_{\mathcal{C}}(x) > f_{\mathcal{C}}(y) + \sigma$. In other words, given any $\mathcal{C}$ an agent following our procedure should always implicitly indicate at least one pair (x, y) where x makes y $\mathcal{C}-$irrelevant.

To illustrate, take any choice function c (defined on a possibly large domain Σ) for which $c(\{x, y\}) = c(\{x, z\}) = x$, $c(\{x, y, z\}) = y$. Consider the subdomain $\mathcal{C} = \{\{x, y\}, \{x, z\}, \{x, y, z\}\}$. Because the agent has chosen x both from $\{x, y\}$ and $\{x, z\}$, so that x is not made $\mathcal{C}-$irrelevant by either of the other alternatives, the agent is 'indicating' that, even if he were using a lexicographic heuristic, the first rationale which is active on $\mathcal{C}$ would not eliminate x. Similarly, because $c(\{x, y, z\}) \neq c(\{x, z\})$ and $c(\{x, y, z\}) \neq c(\{x, y\})$, so that neither y nor z are $\mathcal{C}-$irrelevant, the agent is also indicating that the first active rationale would not eliminate y or z. Thus, no alternative can be eliminated, and we can conclude that the agent cannot possibly be choosing according to our lexicographic procedure.

The remarkable thing is that the following axiom, which formalises this intuition, is not only necessary, but turns out to embody *all* the observable implications of the model:

Reducibility: For every nonempty $\mathcal{C} \subseteq \Sigma$, there exists $S \in \mathcal{C}$ and $x, y \in S$ such that, for all $T \in \mathcal{C}$:

$$(T \setminus \{y\}) \in \mathcal{C}, \ x \in T \Rightarrow c(T) = c(T \setminus \{y\})$$

A choice function which satisfies Reducibility is called *reducible*.

If x makes y $\mathcal{C}-$irrelevant, we cannot identify from choice data alone *why* this is the case. It could simply be that x is 'better' than y (e.g. superior by any criterion). But it could also be that x is pizza, y is steak tartare, and you simply ignore steak tartare in any

restaurant which also offers pizza (though you may or may not choose pizza). Here, pizza might be a negative signal about the kitchen's sophistication, so that you are induced to ignore sophisticated items on the menu, even if you may end up not choosing the signal item itself.[17]

Obviously, one extreme way of satisfying Reducibility is the existence of a 'best' alternative. If c is a choice function that maximizes an ordinary strict preference relation, an alternative which is chosen from an S in $\mathcal{C}$ trivially makes $\mathcal{C}-$irrelevant any alternative which is not chosen from S. In fact in standard theory 'irrelevant' is essentially synonymous with 'unchosen'. Therefore c is reducible in the standard case.

Reducibility relaxes the standard requirement that all rejected alternatives need to be made $\mathcal{C}$-irrelevant on all $\mathcal{C}$ (via the single preference relation) by the 'best' (chosen) alternative, and it does so in two ways. First, some rejected alternatives may not be made $\mathcal{C}$-irrelevant. And, second, an alternative may be made $\mathcal{C}$-irrelevant by some other alternative which is itself not chosen. In other words, Reducibility requires just a bare skeleton of preference to survive.

An example of a reducible non-standard choice function is the three-cycle of choice: $X = \{x, y, z\}$, $c(X) = c(\{x, y\}) = x$, $c(\{y, z\}) = y$, $c(\{x, z\}) = z$. Here y makes z $\mathcal{C}-$irrelevant when either X or $\{y, z\}$ are in $\mathcal{C}$, and Reducibility is satisfied vacuously otherwise. Observe that the choice from the grand set does not make either y or z $\mathcal{C}-$irrelevant for $\mathcal{C}$ coinciding with the full domain.

On the contrary, the choice function c in the proof of proposition 4 (where c is sequentially rationalisable but not cles) is not reducible. In that example Reducibility fails on the collection $\mathcal{C} = \Sigma$. To see this, observe that no i can make $i+1$ $\mathcal{C}-$irrelevant, since we would have the contradiction $c(\{i, i+1, i+2\}) = i \neq i+2 = c(\{i, i+2\})$. Also, $i+1$ cannot make i $\mathcal{C}-$irrelevant, for $c(\{i, i+1\}) = i$. Moreover, no two non adjacent alternatives i and j with $j > i+1$ are suitable either. It cannot be that j makes i $\mathcal{C}-$irrelevant

---

[17]In this example pizza plays a symmetric role that of frog legs in the celebrated example by Luce and Raiffa [17] (a decision maker chooses steak when frog legs are on the menu and salmon when they are not). In Luce and Raiffa's example, frog legs are a positive signal about the quality of the restaurant, so that the decision maker is induced by the presence of frog legs on the menu to choose a high quality item, even if not frog legs themselves.

since $c(\{j, j-1, ..., i+2, i\}) = i$, and it cannot be that i makes j $\mathcal{C}-$irrelevant since $c(\{i, j\}) = j$. This reasoning also highlights the role that infinite domains play in separating lexicographic semiorders from sequentially rationalizable choice.

Reducibility is easily seen to be a weakening of a standard contraction consistency axiom. Consider the following formulation of Independence of Irrelevant Alternatives:

Independence of Irrelevant Alternatives (IIA): Let $\mathcal{C} \subseteq \Sigma$. Then $c(S) = c(S \setminus \{y\})$ for all $y \in S \setminus \{c(S)\}$ for all $S \in \mathcal{C}$ such that $S \setminus \{y\} \in \Sigma$.

Now consider the following weakening (where we highlight in boldface the additional conditions):

Reducibility (restated): Let $\mathcal{C} \subseteq \Sigma$. Then **for some** $x \in X$, $c(S) = c(S \setminus \{y\})$ for **some** $y \in S \setminus \{c(S)\}$ for all $S \in \mathcal{C}$ such that $S \setminus \{y\} \in \Sigma$ **and** $S \ni x$.

While standard IIA requires the choice to be unchanged if *any* unchosen alternative is removed from *any* set, Reducibility requires this to hold only for *some* alternative and for *some* sets (those containing x). Because IIA is so strong, the fact that if it holds, it must hold on the entire domain Σ as well as on any subcollection $\mathcal{C}$, usually does not need to be made explicit.

Below we establish that Reducibility identifies all the observable implications of the lexicographic semiorder procedure, and that basic lexicographic semiorders cover exactly the same ground as general lexicographic semiorders.

Theorem 1 *Let X be finite. Let c be a choice function defined on the domain Σ of all finite subsets of X. Then the following statements are equivalent:*

(i) c is a choice by lexicographic semiorder;

(ii) c is reducible;

(iii) c is a choice by basic lexicographic semiorder.

Proof. *(i)* $\Rightarrow$ *(ii).* Let c be induced by the lexicographic semiorder $(f_i, \sigma)_{i \in I}$, and let $\mathcal{C} \subseteq \Sigma$ be any non-trivial collection of choice sets. Let

$$j = \min \{i : M_i(S) \neq S \text{ for some } S \in \mathcal{C}\}$$

17

(j is well-defined because of the single valuedness of c).[18]

Let $T \in \mathcal{C}$ be such that $M_j(T) \neq T$. Fix $x, y \in T$ such that $f_j(x) > f_j(y) + \sigma$. For any $S \in \mathcal{C}$ either $\{x, y\} \nsubseteq S$, in which case Reducibility holds vacuously; or $\{x, y\} \subseteq S$. In this latter case (which holds at least for $S = T$), for any $z \in S$, if $f_j(y) > f_j(z) + \sigma$ then also $f_j(x) > f_j(z) + \sigma$. Therefore $M_j(S) = M_j(S \backslash \{y\})$, implying $c(S) = c(S \backslash \{y\})$.

(ii) $\Rightarrow$ *(iii)*. Let c be a reducible choice function on Σ. We first provide an algorithm to construct a basic lexicographic semiorder for any choice function, then show that this semiorder induces c.

The algorithm proceeds by recursively defining a sequence of collections $\{\mathcal{C}_i\}_{i \in I}$ and an associated sequence of pairs $\{x_i, y_i\}_{i \in I}$, where I is the interval $\{0, 1, ..., n\}$ for some n. Let $\mathcal{C}_0 = \Sigma$, and let $x_0, y_0 \in X$ be any two alternatives such that, for all $S \in \mathcal{C}_0$, $x_0, y_0 \in S \Rightarrow c(S) = c(S \backslash \{y_0\})$ (alternatives such as x_0 and y_0 exist by Reducibility, and $S \backslash \{y_0\} \in \Sigma$ by assumption). For $0 < i$ define recursively $x_i, y_i \in X$ as any two alternatives such that $(x_i, y_i) \neq (x_j, y_j)$ for all $j < i$, and

$$\text{for all } S \in \bigcap\nolimits_{j < i} \mathcal{C}_j\text{: } x_i, y_i \in S \Rightarrow c(S) = c(S \backslash \{y_i\})$$

and

$$\mathcal{C}_i = \bigcap\nolimits_{j < i} \mathcal{C}_j \backslash \left\{ S \in \bigcap\nolimits_{j < i} \mathcal{C}_j : \{x_i, y_i\} \subseteq S \right\}$$

For all i, let $f_i(x_i) = 1$, $f_i(y_i) = -1$, $f_i(z) = 0$ for all $z \in X \backslash \{x_i, y_i\}$, and $\sigma = 1$. Because X is finite, for any i, unless $S \in \mathcal{C}_{i+1} \Rightarrow |S| = 1$ (i.e. unless $\mathcal{C}_i$ is a trivial collection), it is true by Reducibility that $\mathcal{C}_i \neq \mathcal{C}_{i+1}$. Therefore $S \in \bigcap\nolimits_{i \in I} \mathcal{C}_i \Rightarrow |S| = 1$.

This defines a basic lexicographic semiorder $f = (f_i)_{i \in I}$. As we show below, f induces c. Recall the definition of the survivor sets $M_i(S)$.

Fix $S \in \Sigma$. Suppose by induction that $c(S) \in M_i(S)$. It must be that $M_i(S) \in \mathcal{C}_i$. Otherwise, there would exist $k \leq i$ such that $f_k(x_k) = 1$, $f_k(y_k) = -1$ and $\{x_k, y_k\} \subseteq M_i(S) \in \mathcal{C}_k$, contradicting the definition of $M_i(S)$. If also $M_i(S) \in \mathcal{C}_{i+1}$, then $\{x_{i+1}, y_{i+1}\} \nsubseteq M_i(S)$ and so we have immediately $c(S) \in M_{i+1}(S)$. If $M_i(S) \notin \mathcal{C}_{i+1}$, then (since $M_i(S) \in \mathcal{C}_i$) it must be $\{x_{i+1}, y_{i+1}\} \subseteq S$. It cannot be $y_{i+1} = c(S)$ since,

[18]For choice correspondences one would change the qualifier that not all S in $\mathcal{C}$ are singletons with that that not all of them are such that $c(S) = S$.

by construction of the sequence $\{x_i, y_i\}_{i \in I}$, $c(S) = c(S \setminus \{y_1\}) = ... = c(S \setminus \{y_1, .., y_{i+1}\})$. Therefore $c(S) \in M_{i+1}(S)$.

We now show that for all $s \in S \setminus \{c(S)\}$ there exists a k such that $s \notin M_k(S)$. If not, let $\bigcap_{i \in I} M_i(S) = T$, and let $s \in T$. The definition of T implies that, for all $i \in I$, $\{x_i, y_i\} \nsubseteq T$ (otherwise $x_i, y_i \in M_i(S)$, which is impossible by construction since $f_i(x_i) = 1$ and $f_i(y_i) = -1$). Therefore $T \in \bigcap_{i \in I} \mathcal{C}_i$. But this is a contradiction with $c(S) \neq s \in T$ and $c(S) \in T$, since, as observed before, $T \in \bigcap_{i \in I} \mathcal{C}_i$ implies $|T| = 1$.

(iii) $\Rightarrow$ *(i).* Trivial. $\blacksquare$

Remark 1 *The proof of theorem 1 implies in fact an even more general result. The only feature of a lexicographic semiorder that we have used in the proof is transitivity. Therefore, the same characterisation would hold even for procedures which use generic partial orders at each stage of elimination.*

3.2 Sequentially rationalisable choice

Theorem 1 can be used together with proposition 3 to provide a characterization of acyclic sequential rationalisability:

Corollary 1 *Let X be finite and let Σ be the set of all nonempty subsets of X. Then a choice function on Σ is acyclic sequentially rationalisable if and only if it is reducible.*

In short, then, while acyclic sequential rationalizability and lexicographic semiorders coincide on finite sets, they are nested for choice functions defined over more general domains (see proposition 4). This observation prompts the following natural question: what types of behaviour can be explained by the sequential rationalisability model but not by the lexicographic semiorder model? To this aim we introduce a weakening of Reducibility:

Weak reducibility: For every non empty $\mathcal{C} \subseteq \Sigma$, there exists $S \in \mathcal{C}$ and a collection of pairs $\{x_i, y_i\}_{i=1,2,...}$, with $x_i, y_i \in S$ for all i, such that, for all $T \in \mathcal{C}$:

$$T \setminus \bigcup_{i: x_i \in T} \{y_i\} \in \mathcal{C} \Rightarrow c(T) = c\left(T \setminus \bigcup_{i: x_i \in T} \{y_i\}\right)$$

A choice function that satisfies Weak reducibility is called *weakly reducible*.

The only difference between Reducibility and Weak reducibility is that in the latter the single pair (x, y) has been replaced by a *collection* $\{x_i, y_i\}_{i=1,2,\ldots}$ of pairs. In other words, compared to a reducible choice function, a choice function which is only weakly reducible is such that some alternatives which are not individually $\mathcal{C}-$irrelevant (the removal of any one of those alternatives does affect choice) may nevertheless be 'collectively' $\mathcal{C}-$irrelevant (their collective removal from a choice set has no relevance for choice).

We show that the choice functions which are sequentially rationalisable but not cles are exactly those which are only weakly reducible but not reducible.

Theorem 2 *Let X be finite. Let c be a choice function defined on the domain Σ of all finite subsets of X. Then c is sequentially rationalisable if and only if it is weakly reducible.*

Proof. *Necessity.* Let c be sequentially rationalisable with rationales $\{P_i\}_{i \in I}$, and let $\mathcal{C} \subseteq \Sigma$. Let
$$j = \min\{i : M_i^*(S) \neq S \text{ for some } S \in \mathcal{C}\}$$
Let $A = \{(x, y) : x, y \in S \text{ for some } S \in \mathcal{C} \text{ and } (x, y) \in P_j\}$. A is nonempty by the definition of j. Enumerate the pairs in A to obtain $\{x_i, y_i\}_{i \in J}$ where J is the finite interval $\{1, 2\ldots n\}$ for some n. Let $\bar{j}(S) = \min\{j : M_j(S) = M_k(S) \text{ for all } k \geq j\}$. Note that $\bar{j}$ is well defined since c is sequentially rationalisable. It follows straightforwardly that $M_{\bar{j}(S)}^*(S) = M_{\bar{j}(S)}^*\left(S \backslash \bigcup_{i:x_i \in S}\{y_i\}\right)$ for all $S \in \mathcal{C}$. The sequential rationalisability of c thus implies that $c(S) = c\left(S \backslash \bigcup_{i:x_i \in S}\{y_i\}\right)$.

Sufficiency. Let c be weakly reducible. We construct the rationales explicitly. Let $\mathcal{C}_0 = \Sigma$, and define recursively

$$P_i = \{(x_{ji}, y_{ji})\}_{j=1,\ldots,n(i)}, \text{ where } \{x_{ji}, y_{ji}\}_{j=1,\ldots,n(i)} \text{ is any collection of pairs such that}$$

$$c(S) = c\left(S \backslash \bigcup_{j:x_{ji} \in S}\{y_{ji}\}\right) \quad \forall S \in \mathcal{C}_{i-1};$$

$$\mathcal{C}_i = \{S \in \mathcal{C}_{i-1} : S = M_i^*(T) \text{ for some } T \in \mathcal{C}_{i-1}\}$$

The P_i are well-defined by Weak reducibility. Similarly to the proof of theorem 1, unless $S \in \mathcal{C}_{i+1} \Rightarrow |S| = 1$ (i.e. unless $\mathcal{C}_i$ is a trivial collection), it is true by Weak

Reducibility that $\mathcal{C}_i \neq \mathcal{C}_{i+1}$. Therefore $S \in \bigcap_{i \in I} \mathcal{C}_i \Rightarrow |S| = 1$. We show that $\{P_i\}_{i \in I}$, where I is the interval $\{1, 2...n\}$ for some n, sequentially rationalise c.

Let $x = c(S)$. Whenever $S \in \mathcal{C}_{i-1}$ for some i, it cannot be $(y, x) = P_i$, since $c(S) \neq c(S \setminus (\{x\} \cup A))$ for any $A \subseteq X$, contradicting the definition of P_i. This implies that $x \in M_i^*(S)$ for all i.

We now show that for all $y \in S \setminus \{c(S)\}$ there exists a k such that $y \notin M_k(S)$. If not, let $\bigcap_{i \in I} M_i(S) = T$, and let $y \in T$. The definition of T implies that, for all $i \in I$, $\{x_{ji}, y_{ji}\}_{j=1,...,n(i)} \not\subseteq T$ (otherwise $x_{ji}, y_{ji} \in M_i(S)$, which is impossible by construction since $(x_{ji}, y_{ji}) \in P_i$). Therefore $T \in \bigcap_{i \in I} \mathcal{C}_i$. But this is a contradiction with $c(S) \neq y \in T$ and $c(S) \in T$, since, as observed before, $T \in \bigcap_{i \in I} \mathcal{C}_i$ implies $|T| = 1$. $\blacksquare$

Theorems 1 and 2 are interesting in themselves, as Manzini and Mariotti [21] left the characterization of sequential rationalisability as an open problem.

Apesteguia and Ballester [1] have pioneered a solution to that problem, in so doing offering key insights. Their characterisation of acyclic sequential rationalisability is in terms of a condition called Independence of One Irrelevant Alternative (IOIA). To quickly sketch that condition, we need to define some auxiliary terms. A *binary selector* is a function f which associates to every feasible set S including at least two alternatives a binary feasible set in S. A binary selector f that satisfies certain consistency properties[19] is called *consistent*. Then IOIA requires that $c(S) = c(S \setminus (f(S) \setminus \{c(f(S))\}))$ for some consistent binary selector. While this condition may appear involved, its broad logic is simple, as it essentially imposes a two-stage structure on the choice function c. This is convenient because it reduces the problem of detecting an arbitrarily long sequential structure on c to that of detecting a far simpler construction. Thus, IOIA and Reducibility, which by our results and [1]'s are equivalent conditions in the finite case, highlight different aspects of sequential rationalisability. In both cases, the key idea is to somehow identify the sequence by which eliminations are made. Roughly speaking, both axioms try to identify the next step in the elimination process. IOIA's consistent binary selector tells which pair gets compared next out of any given set. Reducibility tells which pair gets

---

[19]We refer the reader to Apesteguia and Ballester [1] for a precise statement of the definition, which requires substantially more notation extraneous to the purposes of this paper.

compared next out of a given set of sets.[20]

As we have already noted, in the case of finite domains what really matters is the move from asymmetric to acyclic rationales - further strengthening the discriminatory power of each criterion to transitivity implies no additional restrictions for observed behaviour. However, when moving to larger domains this conclusion no longer holds. More precisely, for the case of countable domains we already saw (proposition 4) that there are RSMs with acyclic rationales which are not cles; in addition, the proof of proposition 4 can be easily modified to show that there are choice functions that are RSMs with acyclic rationales but that cannot be induced by any sequence of transitive rationales.[21] Furthermore, it can be shown that there are choice functions which are sequentially rationalizable by means of transitive rationales, but not by a lexicographic semiorders.[22] In sum, for domains with countable X we have that sequential rationalizability by means of asymmetric rationales is more permissive than when acyclic rationales are used, which is in turn more permissive than is the case if transitive rationales are required; and in turn, cles are even more restrictive in terms of observable choice behaviour.

The challenge ahead is to provide characterisations in this vein for very general domains, including for example those of standard consumer theory. This remains an open question. Of course, in domains that are not finite, if one insists on finite termination

---

[20]We are grateful to Bart Lipman for suggesting this interpretation of the relation between the two conditions. An explicit proof of the equivalence between the two conditions in the finite case, as well as an example of their non-equivalence in the infinite case, is available from the authors upon request.

[21]Take the c function induced by the RSM in the proof of proposition 4, and suppose by contradiction, that $Q_1...Q_K$ is a list of transitive rationales that rationalises c. For any two alternatives i and j let $k(i,j) = \min\{k : (i,j) \in Q_k \text{ or } (j,i) \in Q_k\}$. Fix some i and j such that $(i,j) \in Q_1$, and observe that $j \neq 1$, for otherwise $c(\{1,2,..,l\}) = 1$ for all $l \in X$ could not be retrieved. Moreover, it must be that $i = j-1$: if either $i < j-1$ or $i = j+1$ then the choice $c(\{i,j\}) = j$ could not be retrieved, while if $i > j+1$ then $c(\{j, j+1, ..., i\}) = j$ could not be retrieved. So $(i, i+1) \in Q_1$, and since Q_1 is transitive, it must be $(i+2, i) \notin Q_1$, for otherwise $(i+2, i+1) \in Q_1$, and $c(\{i+1, i+2\}) = i+1$ could not be retrieved. But then if $(i+1, i+2) \in Q_1$ we cannot retrieve $c(\{i, i+2\}) = i+2$, since by transitivity $(i, i+2) \in Q_1$; while if $(i+1, i+2) \notin Q_1$ we cannot retrieve $c(\{i, i+1, i+2\}) = i$ (observe that $(i+2, i) \in Q_{k(i,i+2)}$ since $c(\{i, i+2\}) = i+2$).

[22]We are grateful to Gil Riella for providing us with the example. As the example is rather lengthy, we omit it from the paper; however details are available from the authors upon request.

on each set, the procedure underlying choice by lexicographic semiorder will only lead to a unique selection if (at least some of) the criteria are capable of eliminating an infinite number of alternatives. For those domains, one may have to accept in general that a choice correspondence, instead of a choice function, is the appropriate primitive. In particular, one could imagine that the procedure studied so far describes only a preselection of suitable alternatives (a 'shortlist'), while a final unique choice might delegated for example to a standard complete and transitive criterion that picks from the shortlist. Such a procedure would still have the merit of replacing the computation of explicit tradeoffs with basic comparisons until a stage where the choice set has been suitably reduced. In the light of remark 1, on the domain we have studied the behavioural restrictions imposed by the modified procedure would still include Reducibility (because of the transitivity of the final criterion).[23]

4   Concluding remarks

We have focussed especially on the most minimalist version of the model we are proposing, which attributes to the agent very weak powers of discrimination (basic lexicographic semiorders). On finite domains this version is coextensive with a natural restriction of the seemingly far more general sequentially rationalisable choice model of Manzini and Mariotti [21]. On broader domains the model restricts choice data more narrowly than even a stripped down version of sequential rationalisability (Rational Shortlist Methods).

---

[23]For a concrete illustration, consider the case of choice under uncertainty. A class of decision criteria that has gained attention because of their applicability to realistic situations is that of 'quantile maximisation' (see Rostek [27]). Such criteria generalise classical criteria such as maximin, by allowing the use of any other quantile of the induced distribution, instead of the worst outcome. For the reasons we have discussed, however, it is hard to imagine that a decision maker ignores any other aspect of a gamble except, say, its median utility. Rather, if the difference in the median outcomes of two gambles is not too large, he is likely to look at some other quantile, and so on. This procedure, which sequentially uses quantiles of interest to discard gambles, is likely to achieve a significant reduction of a choice set. After that reduction, the use of a cardinal, mean-based criterion to single out a gamble may be far more appealing, because its lack of robustness (which justifies the use of ordinal criteria first) becomes less of a drawback in small sets.

The Reducibility condition delimits exactly the restrictions on choice behaviour that our theory implies. The weakening of Reducibility we have studied illustrates the additional behaviours admitted by sequential rationalisability tout court.

While we would argue that Reducibility has more than a whiff of plausibility, we have eschewed defending it as an a priori compelling property of bounded rationality. The appeal of the theory stems mostly from its psychological basis, its tractability and its testability. Our main aim was to extend Tversky's idea into a model of choice and to tease out the observable implications of the model, in the spirit of the 'revealed preference approach' (see Caplin [3], Gul and Pesendorfer [10], Rubinstein and Salant [29] for methodological discussions of this issue). Reducibility is an easily interpretable and operationally workable concept (as demonstrated by our workouts), and as such we believe it fulfills this role. Our approach is thus in the same spirit as a recent body of work which seeks to characterise models of boundedly rational choice in terms of direct axioms on choice behaviour (e.g. Cherepanov, Feddersen and Sandroni [4]; Eliaz, Richter and Rubinstein [6]; Masatlioglu and Ok [22] and [23]; Masatlioglu and Nakajima [24]; Masatlioglu, Nakajima and Ozbay [25]; Salant and Rubinstein [30]; Tyson [35], beside those already discussed).

The present work is also related to the 'checklist' model of choice in Mandler, Manzini and Mariotti [19]. In that model, an agent goes through an ordered checklist of properties (unary relations), at each step eliminating the alternatives that do not have the specified property. For example, the agent who wishes to buy a house looks first for houses in a certain location, then for those in that location with a minimum square footage, and so on until a final selection is made. A choice by *basic* lexicographic semiorder could be interpreted as a weakening of a choice by checklist, in which the membership of a property is allowed to have three values instead of only two. On this interpretation, $f_i(x) = 1$ (resp., $f_i(x) = -1$) means that x definitely has (resp., does not have) property i, while $f_i(x) = 0$ means that x neither fully has nor fully does not have property i (it falls in a 'grey area' or 'is neutral' with respect to that property). For example, a house's location may neither be entirely convenient (e.g. close to both spouses' workplaces) nor entirely inconvenient (far from both spouses' workplaces).

Because (on certain domains) choosing by checklist is exactly equivalent to maximising a utility function (as shown in Mandler, Manzini and Mariotti [19]), a choice by lexicographic semiorder can also be seen as a versatile but minimal departure from the standard model of rational choice.

5   Appendix

It is instructive to see how the algorithm to construct the rationales of theorem 2 works. We use an example provided by Apesteguia and Ballester [1]. The grand set of alternatives is $X = \{\alpha, \beta, \gamma, \delta, \varepsilon, \varphi\}$. The inverse image of the choice function (i.e. the collection of sets from which each alternative is chosen) is given below:

$$
c^{-1}(\alpha) = \left\{
\begin{array}{l}
\{\alpha, \beta, \gamma, \delta, \varepsilon\}, \\
\{\alpha, \beta, \gamma, \varepsilon\}, \{\alpha, \beta, \gamma, \delta\}, \{\alpha, \beta, \delta, \varepsilon\}, \{\alpha, \gamma, \delta, \varepsilon\}, \\
\{\alpha, \beta, \delta\}, \{\alpha, \delta, \varepsilon\}, \{\alpha, \beta, \gamma\}, \{\alpha, \beta, \varepsilon\}, \{\alpha, \gamma, \varepsilon\}, \\
\{\alpha, \beta\}, \{\alpha, \varepsilon\}, \{\alpha, \delta\}
\end{array}
\right\}
$$

$$
c^{-1}(\beta) = \left\{
\begin{array}{l}
\{\beta, \gamma, \delta, \varepsilon, \varphi\}, \\
\{\beta, \gamma, \delta, \varepsilon\}, \{\beta, \delta, \varepsilon, \varphi\}, \{\beta, \gamma, \varepsilon, \varphi\}, \\
\{\beta, \gamma, \delta\}, \{\beta, \delta, \varepsilon\}, \{\beta, \gamma, \varepsilon\}, \{\beta, \varepsilon, \varphi\}, \\
\{\beta, \delta\}, \{\beta, \gamma\}, \{\beta, \varepsilon\}
\end{array}
\right\}
$$

$$
c^{-1}(\gamma) = \left\{
\begin{array}{l}
\{\gamma, \delta, \varepsilon, \varphi\}, \{\alpha, \gamma, \delta, \varphi\}, \\
\{\alpha, \gamma, \varphi\}, \{\alpha, \gamma, \delta\}, \{\gamma, \delta, \varepsilon\}, \{\gamma, \delta, \varphi\}, \\
\{\alpha, \gamma\}, \{\gamma, \delta\}, \{\gamma, \varphi\}
\end{array}
\right\}
$$

$$
c^{-1}(\delta) = \{\{\beta, \delta, \varphi\}, \{\delta, \varepsilon, \varphi\}, \{\delta, \varepsilon\}, \{\delta, \varphi\}\}
$$

$$
c^{-1}(\varepsilon) = \left\{
\begin{array}{l}
X, \{\alpha, \beta, \gamma, \varepsilon, \varphi\}, \{\alpha, \beta, \delta, \varepsilon, \varphi\}, \{\alpha, \delta, \gamma, \varepsilon, \varphi\}, \\
\{\alpha, \beta, \varepsilon, \varphi\}, \{\alpha, \gamma, \varepsilon, \varphi\}, \{\alpha, \delta, \varepsilon, \varphi\}, \\
\{\alpha, \varepsilon, \varphi\}, \{\gamma, \varepsilon, \varphi\}, \\
\{\gamma, \varepsilon\}, \{\varepsilon, \varphi\}
\end{array}
\right\}
$$

$$
c^{-1}(\varphi) = \left\{
\begin{array}{l}
\{\alpha, \beta, \delta, \gamma, \varphi\}, \\
\{\alpha, \beta, \gamma, \varphi\}, \{\beta, \gamma, \delta, \varphi\}, \{\alpha, \beta, \delta, \varphi\}, \\
\{\alpha, \beta, \varphi\}, \{\beta, \gamma, \varphi\}, \{\alpha, \delta, \varphi\}, \\
\{\alpha, \varphi\}, \{\beta, \varphi\}
\end{array}
\right\}
$$

The base relation $P_c = \{(a, b) \in X \times X : a = c(\{a, b\})\}$ is thus:

$$P_c = \left\{ \begin{array}{l} (\alpha, \beta), (\alpha, \varepsilon), (\alpha, \delta), (\delta, \varepsilon), (\delta, \varphi), (\beta, \delta), (\beta, \gamma), (\beta, \varepsilon), \\ (\gamma, \alpha), (\gamma, \delta), (\gamma, \varphi), (\varepsilon, \gamma), (\varepsilon, \varphi), (\varphi, \alpha), (\varphi, \beta) \end{array} \right\}$$

If the rationales P_i and the collections $\mathcal{C}_{i-1}$ are built according to the algorithm in the proof of theorem 2, obviously it can never be $(a, b) \in P_c \cap P_i$ for any a and b such that b is chosen from some $S \in \mathcal{C}_{i-1}$ that also contains a. Consequently we are going to construct the rationales by first ruling out as potential members of P_i all such pairs; then we will verify whether the residual subcollection of pairs in P_c which have not yet been 'allocated' to any previous rationale P_j, $j < i$, satisfy the requirement in the Weak reducibility axiom, removing more pairs if necessary until we have the largest collection that satisfies the axiom.

Beginning with $\mathcal{C}_0 = \Sigma$, inspection of the inverse images reveals that the only candidate pairs are (α, β), (α, δ), (β, γ), (γ, δ), (ε, φ) and (φ, α), since for all other pairs $(a, b) \in P_c$ it is always the case that b is chosen in some set where a is present.

However, δ is also the only alternative such that, when it is removed from sets that also contain α, leaves choice unchanged. To see this, observe that $(\alpha, \beta) \notin P_1$, since e.g. $\alpha = c(\{\alpha, \beta, \gamma\}) \neq c(\{\alpha, \gamma\}) = \gamma$; $(\beta, \gamma) \notin P_1$, since e.g. $\varphi = c(\{\beta, \gamma, \delta, \varphi\}) \neq c(\{\beta, \delta, \varphi\}) = \delta$; $(\gamma, \delta) \notin P_1$, since e.g. $\gamma = c(\{\gamma, \delta, \varepsilon\}) \neq c(\{\gamma, \varepsilon\}) = \varepsilon$; $(\varepsilon, \varphi) \notin P_1$, since e.g. $\varepsilon = c(\{\alpha, \varepsilon, \varphi\}) \neq c(\{\alpha, \varepsilon\}) = \alpha$; and $(\varphi, \alpha) \notin P_1$, since e.g. $\varepsilon = c(\{\alpha, \beta, \varepsilon, \phi\}) \neq c(\{\beta, \varepsilon, \phi\}) = \beta$.

Consequently,

$$P_1 = \{(\alpha, \delta)\}$$

The domain thus reduces from $\mathcal{C}_0$ to $\mathcal{C}_1$ as indicated in the display that follows (simply remove all sets containing α and δ), where observe that the first line is a subcollection of $c^{-1}(\alpha)$, the second line is a subcollection of $c^{-1}(\beta)$, and so on:

$$
\mathcal{C}_1 = \left\{
\begin{array}{l}
\{\alpha,\beta,\gamma,\varepsilon\}, \{\alpha,\beta,\gamma\}, \{\alpha,\beta,\varepsilon\}, \{\alpha,\gamma,\varepsilon\}, \{\boldsymbol{\alpha},\boldsymbol{\beta}\}, \{\alpha,\varepsilon\} \\[4pt]
\{\beta,\gamma,\delta,\varepsilon,\varphi\}, \{\beta,\gamma,\delta,\varepsilon\}, \{\beta,\delta,\varepsilon,\varphi\}, \{\beta,\gamma,\varepsilon,\varphi\}, \\[4pt]
\qquad \{\beta,\gamma,\delta\}, \{\beta,\delta,\varepsilon\}, \{\beta,\gamma,\varepsilon\}, \{\beta,\varepsilon,\varphi\}, \{\beta,\delta\}, \{\boldsymbol{\beta},\boldsymbol{\gamma}\}, \{\beta,\varepsilon\} \\[4pt]
\{\gamma,\delta,\varepsilon,\varphi\}, \{\alpha,\gamma,\varphi\}, \{\gamma,\delta,\varepsilon\}, \{\gamma,\delta,\varphi\}, \{\alpha,\gamma\}, \{\boldsymbol{\gamma},\boldsymbol{\delta}\}, \{\gamma,\varphi\} \\[4pt]
\{\beta,\delta,\varphi\}, \{\delta,\varepsilon,\varphi\}, \{\boldsymbol{\delta},\boldsymbol{\varepsilon}\}, \{\delta,\varphi\} \\[4pt]
\{\alpha,\beta,\gamma,\varepsilon,\varphi\}, \{\alpha,\beta,\varepsilon,\varphi\}, \{\alpha,\gamma,\varepsilon,\varphi\}, \{\alpha,\varepsilon,\varphi\}, \{\gamma,\varepsilon,\varphi\}, \{\gamma,\varepsilon\}, \{\boldsymbol{\varepsilon},\boldsymbol{\varphi}\} \\[4pt]
\{\alpha,\beta,\gamma,\varphi\}, \{\beta,\gamma,\delta,\varphi\}, \{\alpha,\beta,\varphi\}, \{\beta,\gamma,\varphi\}, \{\boldsymbol{\alpha},\boldsymbol{\varphi}\}, \{\beta,\varphi\}
\end{array}
\right\}
$$

Next, observe that α and φ are chosen in the presence of γ, so that our algorithm prescribes $(\gamma,\alpha) \notin P_2$ and $(\gamma,\varphi) \notin P_2$. Moreover, β is chosen in the presence of φ; γ is chosen in the presence of ε; δ and ε in the presence of β; ε is chosen in the presence of α; and φ is chosen in the presence of δ. This leaves only (α,β), (β,γ), (γ,δ), (δ,ε), (ε,φ) and (φ,α) as potential members of P_2 (appearing in boldface in the above display), and it is easy to verify that indeed the whole collection of 'candidate pairs'

$$
P_2 = \{(\alpha,\beta), (\beta,\gamma), (\gamma,\delta), (\delta,\varepsilon), (\varepsilon,\varphi), (\varphi,\alpha)\}
$$

is such that $c(S) = c\left(S \backslash \bigcup_{i:x_i \in S} y_i\right)$ for all $S \in \mathcal{C}_1$ as in the definition of Weak Reducibility. Note also that Reducibility fails on the collection $\mathcal{C}_1$: no set contains α and δ, and by the same considerations contained in the previous paragraphs, the only pairs of alternatives that might satisfy Reducibility are $\{\alpha,\beta\}$, $\{\beta,\gamma\}$, $\{\gamma,\delta\}$, $\{\delta,\varepsilon\}$, $\{\varepsilon,\varphi\}$ and $\{\varphi,\alpha\}$. However, none of them does: first of all, because all these binary sets are in $\mathcal{C}_1$, the 'losing' alternative must be the one that is not chosen in pairwise sets; in addition, $(x_2, y_2) \notin \{(\alpha,\beta), (\beta,\gamma), (\gamma,\delta), (\varepsilon,\varphi), (\varphi,\alpha)\}$ since the same sets for which $c(S) \neq c(S \backslash \{y_2\})$ when checking $\mathcal{C}_0$ are also in $\mathcal{C}_1$; while $x_2, y_2 \neq \delta, \varepsilon$ since e.g. $\beta = c(\{\beta,\gamma,\delta,\varepsilon,\varphi\}) \neq c(\{\beta,\gamma,\delta,\varphi\}) = \varphi$.

Going back to our algorithm, the construction of P_2 yields

$$\mathcal{C}_2 = \left\{ \begin{array}{c} \{\alpha, \gamma, \varepsilon\}, \{\boldsymbol{\alpha}, \boldsymbol{\varepsilon}\} \\ \{\beta, \delta\}, \{\boldsymbol{\beta}, \boldsymbol{\varepsilon}\} \\ \{\alpha, \gamma\}, \{\boldsymbol{\gamma}, \boldsymbol{\varphi}\} \\ \{\beta, \delta, \varphi\}, \{\boldsymbol{\delta}, \boldsymbol{\varphi}\} \\ \{\boldsymbol{\gamma}, \boldsymbol{\varepsilon}\} \\ \{\boldsymbol{\beta}, \boldsymbol{\varphi}\} \end{array} \right\}$$

For the next step, we note that δ is chosen in the presence of β; α is chosen in the presence of γ. So one can verify that all together the remaining candidate pairs provide a suitable P_3, that is:

$$P_3 = \{(\alpha, \varepsilon), (\varepsilon, \gamma), (\beta, \varepsilon), (\delta, \varphi), (\varphi, \beta), (\gamma, \varphi)\}$$

As a consequence, the subdomain reduces to:

$$\mathcal{C}_3 = \{\{\beta, \delta\}, \{\alpha, \gamma\}\}$$

so that we can build the final rationale

$$P_4 = \{(\beta, \delta), (\gamma, \alpha)\}$$

It is straightforward to double check that P_1, P_2, P_3, P_4 so defined sequentially rationalises c.

References

[1] Apesteguia, J. and M. A. Ballester (2010) "Choice by Sequential Procedures", mimeo, Universitat Autonoma de Barcelona and Universitat Pompeu Fabra.

[2] Binmore, K., A. Voorhoeve and B. Wallace (2008) "Similarity-Based Decision-Making in Moral Decisions: Experimental Results", mimeo, London School of Economics.

[3] Caplin, A. (2008) "Economic Theory and Psychological Data: Bridging the Divide", in: *The Foundations of Positive and Normative Economics: A Handbook*, A. Caplin and Andrew Schotter (eds.), Oxford University Press.

[4] Cherepanov, V., T. Feddersen and A. Sandroni (2008) "Rationalization", mimeo, Northwestern University.

[5] Ehlers, L. and Y. Sprumont (2008) "Weakened WARP and top-cycle choice rules", *Journal of Mathematical Economics*, 44: 87-94.

[6] Eliaz, K., M. Richter and A. Rubinstein (2011) "Choosing the Two Finalists", *Economic Theory*, 46: 211-219.

[7] Fishburn, P. C. (1970) "Intransitive Indifference with Unequal Indifference Intervals," *Journal of Mathematical Psychology*, 7: 144-149.

[8] Fishburn, P. C. (1985) *Interval Orders and Interval Graphs: A Study of Partially Ordered Sets.* New York: Wiley.

[9] Gigerenzer, G. and P. M. Todd. (1999) "Fast and Frugal Heuristics: The Adaptive Toolbox." In *Simple heuristics that make us smart*, ed. G. Gigerenzer, P. M. Todd, and the ABC Research Group, 3-34. New York: Oxford University Press.

[10] Gul, F. and W. Pesendorfer (2008) "The Case for Mindless Economics", in: *The Foundations of Positive and Normative Economics: A Handbook*, A. Caplin and Andrew Schotter (eds.), Oxford University Press.

[11] Houy, N. and K. Tadenuma (2009) "Lexicographic Compositions of Multiple Criteria for Decision Making", *Journal of Economic Theory*, 144: 1770-1782.

[12] Kalandrakis, T. (2010) "Rationalizable voting", *Theoretical Economics,* 5: 93–125

[13] Kohli, R., and Jedidi, K. (2007) "Representation and inference of lexicographic preference models and their variants" *Marketing Science*, 26: 380–399.

[14] Leland, J. W. (1994). "Generalized Similarity Judgments: An Alternative Explanation for Choice Anomalies", *Journal of Risk and Uncertainty*, 9(2): 151-72.

[15] Lombardi, M. (2008) "Uncovered set choice rules", *Social Choice and Welfare*, 31: 271 – 279.

[16] Luce, R. D. (1956) "Semiorders and a theory of utility discrimination", *Econometrica*, 24: 178–191.

[17] Luce, R. D. and H. Raiffa (1957) *Games and Decisions: Introduction and Critical Survey*, Wiley & Sons.

[18] Mandler, M. (2009) "Rational agents are the quickest", mimeo, Royal Holloway University of London.

[19] Mandler, M., P. Manzini and M. Mariotti (2008) "One Million Answers to Twenty Questions: Choosing by Checklist", mimeo, Queen Mary, University of London.

[20] Manzini, P. and M. Mariotti (2006) "A Vague Theory of Choice over Time", *Advances in Theoretical Economics*, Berkeley Electronic Press, 6: 1265-1265.

[21] Manzini, P. and M. Mariotti (2007) "Sequentially Rationalizable Choice", *American Economic Review*, 97: 1824-1839.

[22] Masatlioglu, Y. and E. Ok (2005) "Rational Choice with Status-quo Bias", *Journal of Economic Theory* 121: 1-29.

[23] Masatlioglu, Y. and E. Ok (2006) "Reference-Dependent Procedural Decision Making", working paper, NYU.

[24] Masatlioglu, Y. and D. Nakajima (2008) "Choice by Constrained Elimination" mimeo, University of Michigan.

[25] Masatlioglu, Y., Nakajima and E. Ozbay (2009) "Revealed Attention", forthcoming in the *American Economic Review*.

[26] Medina, B., S. Naeh and U. Segal (2010) "Ranking Ranking Rules", Working Paper, Boston College.

[27] Rostek, M. (2010), "Quantile Maximization in Decision Theory", *Review of Economic Studies*, 77: 339–371.

[28] Rubinstein, A. (1988) "Similarity and Decision Making Under Risk (Is there a utility theory resolution to the Allais Paradox?)", *Journal of Economic Theory*, 46: 145-153.

[29] Rubinstein, A. and Y. Salant (2008) "Some thoughts on the principle of revealed preference", in: *The Foundations of Positive and Normative Economics: A Handbook*, A. Caplin and Andrew Schotter (eds.), Oxford University Press.

[30] Salant, Y. and A. Rubinstein (2009) "(A,f): Choice with frames", *Review of Economic Studies*, 75: 1287-1296.

[31] Tadenuma, K. (2002) "Efficiency First or Equity First? Two Principles and Rationality of Social Choice", *Journal of Economic Theory*, 104: 462-472.

[32] Tadenuma, K. (2005), "Egalitarian-Equivalence and the Pareto Principle for Social Preferences," *Social Choice and Welfare* 24: 455-473.

[33] Tversky, A. (1969) "Intransitivity of Preferences", *Psychological Review*, 76: 31-48.

[34] Tversky, A. (1972) "Elimination by aspects: A theory of choice", *Psychological Review*, 79: 281–299.

[35] Tyson, C. J. "Cognitive constraints, contraction consistency, and the satisficing criterion", *Journal of Economic Theory*, 138: 51 – 70

[36] Yee, M., Dahan, E., Hauser, J. R., and Orlin, J. (2007) "Greedoid-based noncompensatory inference", *Marketing Science*: 26, 532–549.