SSE Working Paper Series in Economics
No. 2015:2

Characterizing lexicographic preferences

Henrik Petri [a]
Mark Voorneveld [b]

[a] Stockholm School of Economics
Department of Finance

[b] Stockholm School of Economics
Department of Economics

Characterizing lexicographic preferences

Henrik Petri[1] and Mark Voorneveld*[2]

[1]Department of Finance, Stockholm School of Economics, Box 6501, 113 83 Stockholm, Sweden, henrik.petri@hhs.se

[2]Department of Economics, Stockholm School of Economics, Box 6501, 113 83 Stockholm, Sweden, mark.voorneveld@hhs.se

April 15, 2015

SSE Working Paper Series in Economics, No. 2015:2

Abstract

We characterize lexicographic preferences on product sets of finitely many coordinates. The main new axiom is a robustness property. It roughly requires this: Suppose x is preferred to y; many of its coordinates indicate that the former is better and only a few indicate the opposite. Then the decision maker is allowed a change of mind turning one coordinate in favor of x to an indifference: even if one less argument supports the preference, the fact that we started with many arguments in favor of x suggests that such a small change is not enough to give rise to the opposite preference.

Keywords: lexicographic preferences; lexicographic order; characterization; preferences

JEL codes: D01; D11; D70

*Corresponding author. Financial support by the Wallander-Hedelius Foundation under grant P2010-0094:1 is gratefully acknowledged.

1 Introduction

In advanced courses on microeconomics, lexicographic preferences over real vectors often serve as a cogent reminder that utility functions are useful tools in economic analysis, but cannot represent all plausible preference relations. An early decision-theoretic paper that makes this point is Debreu (1954). But theoretical overview articles (Fishburn, 1974; Martínez-Legaz, 1998), empirical studies (Yee et al., 2007), and references therein illustrate the use of lexicographic preferences/orders in many other branches of economics as well.

The purpose of this paper is to characterize lexicographic preferences on a product set $X = \times_{i \in I} X_i$ of finitely many coordinates. The challenge, of course, is to do so without explicit reference to the order of importance of the different coordinates: the same set X allows for different lexicographic preferences by looking at different priorities of the coordinates. Informally, preferences $\succ$ ('is preferred to') are lexicographic if we can find (a) an order on the coordinates $i \in I$ and (b) for each coordinate $i \in I$, a preference $\succ_i$ on the coordinate set X_i such that $x = (x_i)_{i \in I}$ is preferred to $y = (y_i)_{i \in I}$ if and only if some coordinate makes a difference ($x_i \succ_i y_i$ or $y_i \succ_i x_i$) and the first such coordinate makes x look better than y; precise definitions are in Sec. 2.

Our motivation is that parts of the existing characterizations of lexicographic preferences in Fishburn (1975, 1976) are problematic. We will discuss the reasons in detail in Section 6, once we can appeal to earlier notation and axioms. Informally, the axioms in Fishburn (1975) are nice, but his characterization has an unnatural domain restriction, since (a) one of its axioms points out that only *two* equivalence classes per coordinate matter to determine the entire order, but (b) to accommodate a special method of proof, it is assumed that each coordinate has *three or more* equivalence classes. This implies, for instance, that his characterization cannot address natural lexicographic models (Mandler et al., 2012) where alternatives have binary characteristics, like possessing or not possessing a desirable property. Fishburn (1976) considers, just like we do, the general domain. But he adds — without supporting intuitive motivation — a technical axiom, superadditivity, that is very strong and rather explicitly lexicographic. So our aim is a characterization that applies to the general domain (without the 3-equivalence-class restriction) and that uses an axiom that is both less demanding than the technical superadditivity assumption and has a straightforward intuition.

This new axiom is *robustness*. Its precise formulation — axiom A4 — is in Section 3; but what is the intuition behind it? Well, suppose x is preferred to y. Indeed, while there are attributes (coordinates) supporting either alternative, many are in favor of x, only few in favor of y. Then the decision maker can have a change of mind when it comes to some attribute in favor of x: instead of thinking that it argues for x, the decision maker judges the alternatives to be indistinguishable in that coordinate. So there is one less attribute in support of x. This may change the strict preference to a weak one. But since we started with a fairly large support anyway, the change is not enough to all of a sudden reverse the preference. In other words, if many attributes are in favor of x and only a few in favor of y, then a preference of x over y is robust to (not reversed by) a small change of mind, turning some attribute in favor of x to an equality.

Our contributions include the following:

1. Theorem 4.1 characterizes lexicographic preferences using four axioms: robustness and three other axioms that are (weaker forms of those) familiar from earlier papers on the topic.

2. The axioms are shown to be logically independent in Section 5.

3. In Section 6 we prove in detail how the assumptions in the two earlier characterizations of lexicographic preferences in Fishburn (1975, 1976) imply our axioms. That is, these two characterizations are special cases/corollaries of ours: see Cor. 6.1 and 6.2. But in contrast, our purpose was to avoid the unnatural domain restriction in the former and the technical and demanding superadditivity axiom in the latter.

The paper is organized as follows. After preliminaries in Section 2, we introduce and discuss our axioms in Section 3. The main characterization is in Section 4. Its axioms are shown to be logically independent in Section 5. The relation with Fishburn (1975, 1976) is treated in Section 6. Finally, Section 7 addresses variants of our robustness axiom and products of infinitely many sets. One proof is in the first appendix; the second provides additional examples of robust preferences.

2 Notation and definition of lexicographic preferences

Our notation follows Fishburn (1975, 1976) as much as possible. For completeness, subsection 2.1 contains standard definitions/notation for binary relations and product sets. Subsection 2.2 defines our object of study, lexicographic preferences on a product set over finitely many indices, and contains a lemma characterizing the order within each coordinate set. We briefly discuss infinite products in our concluding remarks, Sec. 7.

2.1 Standard definitions

Binary relations. A binary relation $\succ$ on a set X is a subset of $X \times X$. If $(x, y) \in \succ$, we write $x \succ y$. For $x, y \in X$, define

$$x \sim y \Leftrightarrow (\text{not } x \succ y \text{ and not } y \succ x) \qquad \text{and} \qquad x \succsim y \Leftrightarrow (x \sim y \text{ or } x \succ y).$$

A binary relation $\succ$ on a set X is:

- ⊠ *irreflexive* if, for all $x \in X$: not $x \succ x$.
- ⊠ *asymmetric* if, for all $x, y \in X$: $x \succ y$ implies that not $y \succ x$.
- ⊠ *transitive* if for all $x, y, z \in X$: $x \succ y$ and $y \succ z$ imply $x \succ z$.
- ⊠ *negatively transitive* if for all $x, y, z \in X$: $x \succ y$ implies $x \succ z$ or $z \succ y$.
- ⊠ a *linear order* if it is irreflexive, transitive, and for all $x, y \in X$ with $x \neq y$, either $x \succ y$ or $y \succ x$.

Asymmetry and negative transitivity imply transitivity: let $x, y, z \in X$ have $x \succ y$ and $y \succ z$. By negative transitivity, $x \succ z$ or $z \succ y$. Asymmetry rules out $z \succ y$, so $x \succ z$.

Sets. $\mathbb{R}$ is the set of real numbers. For subsets A and B of a set I ($A \subseteq I, B \subseteq I$), we write $A \setminus B = \{a \in A : a \notin B\}$ and denote the complement of A w.r.t. I as $I \setminus A$ or A^c if I is evident from the context. $|A|$ is the cardinality/number of elements of a finite set A.

For each i in a nonempty index set I, let X_i be a set. Denote their product set by $X = \times_{i \in I} X_i = \{(x_i)_{i \in I} : x_i \in X_i \text{ for each } i \in I\}$. As usual, we refer to elements $i \in I$ as indices, coordinates, or attributes. Conventional notational shortcuts are used. For instance, for $i \in I$, $X_{-i} = \times_{j \in I \setminus \{i\}} X_j$. Let $i, j \in I$ and $A \subseteq I$ with $i \in A$. Elements $x = (x_k)_{k \in I} \in X$ may be denoted by (x_i, x_{-i}) or (x_i, x_j, x_{-ij}) or $(x_A, x_{-A}) = (x_{A \setminus \{i\}}, x_i, x_{-A})$ if we want to stress coordinates i, j, or those in A.

2.2 Lexicographic preferences

From here on, $\succ$ ('is preferred to') is an asymmetric binary relation on a nonempty product set $X = \times_{i \in I} X_i$ with finitely many (but at least two) attributes $i \in I$. Informally, $\succ$ is lexicographic if we can order the coordinates $i \in I$ and find, for each $i \in I$, a binary relation $\succ_i$ on X_i, such that $x \succ y$ if and only if some coordinate i makes a difference ($x_i \succ_i y_i$ or $y_i \succ_i x_i$) and the first such coordinate makes x look better than y.

Definition 2.1. An asymmetric binary relation $\succ$ on $X = \times_{i \in I} X_i$ is *lexicographic* if there is a linear order $<_0$ on coordinates I and, for each $i \in I$, a binary relation $\succ_i$ on X_i such that for all $x = (x_i)_{i \in I}$ and $y = (y_i)_{i \in I}$ in X, $x \succ y$ if and only if $x_i \succ_i y_i$ for some $i \in I$ and for each $k \in I$ with $y_k \succ_k x_k$ there is a $j \in I$ with $j <_0 k$ and $x_j \succ_j y_j$. ◁

Equivalently, stressing the first coordinate according to $<_0$ that makes a difference, $\succ$ is lexicographic if, for all $x, y \in X$:

$$x \succ y \Leftrightarrow \text{there is an } i \in I \text{ such that } \begin{cases} x_i \succ_i y_i \text{ is true, but} \\ y_j \succ_j x_j \text{ is false for all } j <_0 i. \end{cases} \tag{1}$$

If preferences are lexicographic, relations $\succ_i$ on X_i are easily characterized (Fishburn (1975, p. 416) alludes to this result, but without proof and under stronger assumptions):

Lemma 2.1. *If preferences are lexicographic, relations $\succ_i$ in Definition 2.1 satisfy, for all $i \in I$ and $x_i, y_i \in X_i$:*

$$x_i \succ_i y_i \Leftrightarrow \textit{for all } z_{-i} \in X_{-i} : (x_i, z_{-i}) \succ (y_i, z_{-i}). \tag{2}$$

Proof. If $\succ$ is lexicographic, then for all $x, y \in X$:

$$x \sim y \Leftrightarrow \{i \in I : x_i \succ_i y_i \text{ or } y_i \succ_i x_i\} = \emptyset \Leftrightarrow x_i \sim_i y_i \text{ for all } i \in I. \tag{3}$$

Asymmetry of $\succ$ implies $x \sim x$ for all $x \in X$. With (3), $x_i \sim_i x_i$ for all $i \in I$ and $x_i \in X_i$. We use this to prove (2).

$\Rightarrow$: Let $i \in I, x_i, y_i \in X_i$ have $x_i \succ_i y_i$. Let $z_{-i} \in X_{-i}$. Since $x_i \succ_i y_i$ and $z_j \sim_j z_j$ for all $j \neq i$, lexicographic preference gives $(x_i, z_{-i}) \succ (y_i, z_{-i})$.

$\Leftarrow$: Let $i \in I, x_i, y_i \in X_i$, and $z_{-i} \in X_{-i}$ have $(x_i, z_{-i}) \succ (y_i, z_{-i})$. Since $z_j \sim_j z_j$ for all $j \neq i$, lexicographic preference must come from a difference in coordinate i: $x_i \succ_i y_i$. $\square$

So, given asymmetric relation $\succ$ on X, let us define, for each $i \in I$, relation $\succ_i$ on X_i by (2). Since $\succ$ is asymmetric, each $\succ_i$ is asymmetric. In contrast, in some examples we start with relations $(\succ_i)_{i \in I}$ on $(X_i)_{i \in I}$ to define $\succ$ on $X = \times_{i \in I} X_i$. To avoid ambiguity, we will then check that these relations $\succ_i$ satisfy (2).

Following Fishburn (1976, p. 395), we call coordinate i *essential* if there are $a_i, b_i \in X_i$ with $a_i \succ_i b_i$ and we assume throughout that each coordinate is essential. This is without loss of generality (w.l.o.g.): lexicographic preferences depend only on essential coordinates. (A bit more explicitly, for lexicographic preferences or — for that matter — any preference $\succ$ satisfying the noncompensation axiom A1 in Section 3, this is without loss of generality: if coordinate i is not essential, then $x \succ y$ if and only if $(z_i, x_{-i}) \succ (z_i, y_{-i})$ for all $z_i \in X_i$. Hence we may consider the restriction of $\succ$ to $\{z_i\} \times X_{-i}$ for some $z_i \in X_i$. Roughly speaking, we may 'forget' about inessential coordinates. In fact, all our axioms impose restrictions only in terms of essential coordinates.)

3 Axioms and discussion

In this section we introduce and discuss our axioms. For an ordered pair $(x, y) \in X \times X$, let $P(x, y) = \{i \in I : x_i \succ_i y_i\}$ denote the indices according to which x_i is preferred to y_i. Informally, when comparing x and y, set $P(x, y)$ gives the attributes or coordinates in favor of x. By asymmetry of $\succ_i$, $P(x, y)$ and $P(y, x)$ are disjoint.

We characterize lexicographic preferences using the following axioms on preferences $\succ$:

A1 *Noncompensation:* for all $x, y, w, z \in X$,

$$[P(x, y) = P(w, z), P(y, x) = P(z, w)] \quad \Rightarrow \quad [x \succ y \Leftrightarrow w \succ z].$$

A2 $\succ$ is transitive on a product set $X' = \times_{i \in I}\{a_i, b_i\} \subseteq X$, where $a_i \succ_i b_i$ for each $i \in I$.

A3 *Weak decisiveness:* for all $x, y \in X$, if $x_i \succ_i y_i$ and $y_j \succ_j x_j$ for some distinct $i, j \in I$, then $x \succ y$ or $y \succ x$.

A4 *Robustness:* For all $x, y \in X$,

if $x \succ y$, $P(y, x) \neq \emptyset$, and $|P(x, y)|/|P(y, x)| \geq 2$,

then there is an $i \in P(x, y)$ with $(z_i, x_{-i}) \succsim (z_i, y_{-i})$ for all $z_i \in X_i$ (equivalently: there is no $z_i \in X_i$ with $(z_i, y_{-i}) \succ (z_i, x_{-i})$).

The notation of the robustness axiom A4 is cumbersome, but the intuition behind it is simple. Robustness basically says that if many attributes are in favor of x and only a few in favor of y, then a small change of mind — turning some attribute in favor of x to an indifference/equality — is not enough to reverse a preference of x over y. In a bit more detail, let x be preferred to y. Indeed, suppose the attributes make a strong case for this: although there are attributes in favor of each of the two alternatives, many are in favor of x, only few (here, at most half as many) in favor of y. Then the decision maker can have a change of mind when it comes to some attribute in favor of x and instead judge the alternatives to be indistinguishable in that coordinate. That will slightly decrease the set of attributes in favor of x (by one element). But given that we started with a fairly large support anyway, that is not enough to reverse the preference. Variants of this axiom will be discussed in the concluding remarks, Sec. 7.

Lexicographic preferences and many other plausible ones satisfy this property. Two prominent examples are the weak Pareto order (Example 5.3) and simple majority preferences defined, for asymmetric relations $(\succ_i)_{i \in I}$ on finitely many sets $(X_i)_{i \in I}$, as follows:

$$\text{for all } x, y \in X = \times_{i \in I} X_i : \qquad x \succ y \quad \Leftrightarrow \quad |P(x, y)| > |P(y, x)|, \tag{4}$$

i.e., x is preferred to y if more attributes are in favor of x than in favor of y. Indeed, if $x, y \in X$ satisfy the assumptions of robustness, let $i \in P(x, y), z_i \in X_i$, and $x' = (z_i, x_{-i}), y' = (z_i, y_{-i})$. Then

$$\begin{aligned} x \succ y \Rightarrow |P(x, y)| &= |P(x', y')| + 1 > |P(y, x)| = |P(y', x')| \\ &\Rightarrow |P(y', x')| < |P(x', y')| + 1 \\ &\Rightarrow |P(y', x')| \leq |P(x', y')| \\ &\Rightarrow (\text{not } y' \succ x'), \text{ i.e., } x' = (z_i, x_{-i}) \succsim y' = (z_i, y_{-i}). \end{aligned}$$

Since z_i was arbitrary, majority preferences (4) satisfy robustness. (Of course, this asymmetric relation also satisfies noncompensation A1.) For the interested reader, the second appendix provides two additional classes of robust preferences.

All other axioms are either straight from Fishburn or are weaker forms of his assumptions. We mention those only briefly. The precise relation is discussed in Section 6, where we prove in detail how the axioms in his characterizations imply ours.

The noncompensation axiom A1 is taken from Fishburn (1975, Axiom 3) and Fishburn (1976, Definition 1). A similar property in the characterization of Pareto dominance is called 'ordinality' (Voorneveld, 2003, p. 8): the noncompensation axiom stresses the ordinal character of preferences. That is, to order alternatives x and y, what matters are the set $P(x, y)$ of coordinates in favor of x and the set $P(y, x)$ of coordinates in favor of y. But not the actual 'size' of the coordinates x_i and y_i: preferences are not affected by order-preserving transformations on each coordinate.

Axiom A2 relaxes the negative transitivity property in Fishburn (1975, Axiom 1) by requiring transitivity only on a product set obtained by restricting attention to two elements per coordinate: one 'good' element a_i and one 'bad' element b_i, in the sense that $a_i \succ_i b_i$. In its definition, we use the innocent assumption that all coordinates are essential. If one wants to allow for some inessential coordinates, it can be modified easily by requiring transitivity on a product set of the form $X' = \times_{i \in I} X'_i \subseteq X$ with

$$X'_i = \begin{cases} \{a_i, b_i\} & \text{for some } a_i, b_i \in X_i \text{ with } a_i \succ_i b_i \text{ if coordinate } i \text{ is essential;} \\ \{c_i\} & \text{for some } c_i \in X_i \text{ if coordinate } i \text{ is not essential.} \end{cases}$$

One might surmise that this axiom plus noncompensation A1 give transitivity on all of X, but that is not the case:

Example 3.1 (No implied transitivity)**.** Consider lexicographic preferences $\succ$ on $X_1 \times X_2 = \{a, b, c\} \times \{0, 1\}$ with order $1 <_0 2$ on the coordinates, nontransitive relation $\succ_1$ on $X_1 = \{a, b, c\}$ defined by $a \succ_1 b$ and $b \succ_1 c$ (so $a \sim_1 c$) and $\succ_2$ on $X_2 = \{0, 1\}$ defined by $1 \succ_2 0$. By direct verification or invoking characterization theorem 4.1, lexicographic $\succ$ satisfies all axioms. In particular, it is transitive on $\{a, b\} \times \{0, 1\}$. But $\succ$ is not transitive: $(a, 0) \succ (b, 0)$ and $(b, 0) \succ (c, 0)$, but $a \sim_1 c$ and $0 \sim_2 0$ imply that $(a, 0) \succ (c, 0)$ is false. $\triangleleft$

So our setting is general enough to allow nontransitive preferences on coordinate sets X_i. Fishburn (1991) reviews the literature on such preferences and their behavioral motivations. But the weak transitivity requirement A2, which is (Thm. 4.1) a necessary condition for lexicographic preferences, does require a certain transitivity *across* attributes: e.g., if it holds on $X' = \times_{i=1}^{3} \{a_i, b_i\}$, and $x = (a_1, b_2, b_3), y = (b_1, a_2, b_3), z = (b_1, b_2, a_3)$ all have a unique, distinct attribute in their favor, then $x \succ y$ and $y \succ z$ imply $x \succ z$.

Weak decisiveness A3 assures that if there are attributes in favor of x and others in favor of y, then preferences have enough cutting power to order x and y. It relaxes the decisiveness property C6 (cf. axiom A7 below) in Fishburn (1976). It is satisfied, for instance, by all preferences over X that can be represented by an injective (one-to-one) utility function. But it does not require that there are few indifferences ('small equivalence classes'):

Example 3.2 (Lexicographic preferences over integer parts)**.** $\mathbb{Z} = \{\dots, -2, -1, 0, 1, 2, \dots\}$ denotes the set of integers. For each $r \in \mathbb{R}$, let $\lfloor r \rfloor = \max\{z \in \mathbb{Z} : z \leq r\}$ be the integer part or 'floor' of r. Relation $\succ$ on $\mathbb{R}^2$ with

$$\text{for all } x, y \in \mathbb{R}^2: \qquad x \succ y \quad \Leftrightarrow \quad \lfloor x_1 \rfloor > \lfloor y_1 \rfloor \text{ or } (\lfloor x_1 \rfloor = \lfloor y_1 \rfloor \text{ and } \lfloor x_2 \rfloor > \lfloor y_2 \rfloor)$$

is lexicographic. It orders coordinates $1 <_0 2$. For each $i \in \{1, 2\}$, $\succ_i$ on coordinate set $\mathbb{R}$ satisfies $x_i \succ_i y_i$ if and only if $\lfloor x_i \rfloor > \lfloor y_i \rfloor$. Hence all x and y in the uncountable set $\{z \in \mathbb{R}^2 : 0 \leq z_1 < 1, 0 \leq z_2 < 1\}$ satisfy $x \sim y$. $\triangleleft$

To prepare for our characterization of lexicographic preferences in the next section, Lemma 3.1 provides relations between our axioms and the usual Pareto dominance or monotonicity properties of preferences over product sets.

Lemma 3.1 (Pareto)**.** *Let $\succ$ be an asymmetric relation on $X = \times_{i \in I} X_i$.*

(a) Let $\succ$ satisfy noncompensation A1. For all $x, y \in X$, if $x_i \sim_i y_i$ for all $i \in I$: $x \sim y$.

(b) Let $\succ$ satisfy transitivity requirement A2 on $X' = \times_{i \in I} \{a_i, b_i\} \subseteq X$. For all $x, y \in X'$ with $x_i \succsim_i y_i$ for all $i \in I$ and $x_i \succ_i y_i$ for some $i \in I$: $x \succ y$.

Proof. (a) Let $x, y \in X$ have $x_i \sim_i y_i$ for all $i \in I$. Since $P(x, y) = P(y, x) = \emptyset$, noncompensation implies that $x \succ y$ if and only if $y \succ x$. By asymmetry, neither $x \succ y$, nor $y \succ x$. So $x \sim y$.

(b) With $x, y \in X'$ as in the statement of (b), $A = P(x, y)$ is nonempty. If $i \in I \setminus A$, then $x_i \sim_i y_i$ and, by definition of X', $x_i = y_i$:

$$x_i = y_i \text{ for all } i \in I \setminus A. \tag{5}$$

We prove (b) by induction on the cardinality $|A|$ of A.

If $|A| = 1$, then $x_i \succ_i y_i$ for some $i \in I$ and $x_j \sim_j y_j$ for all $j \neq i$. By (5), $x_{-i} = y_{-i}$. By definition of $x_i \succ_i y_i$: $x = (x_i, x_{-i}) \succ (y_i, x_{-i}) = (y_i, y_{-i}) = y$.

Next, let $k \in \mathbb{N}$ satisfy $1 \leq k < |I|$ and suppose (b) is true if $|A| \leq k$. Let $x, y \in X'$ be such that $|A| = k + 1$. Pick $i \in A$. By (5), $x_{-A} = (x_j)_{j \in I \setminus A} = (y_j)_{j \in I \setminus A} = y_{-A}$, so using the induction hypothesis twice to obtain the strict preferences, we have:

$$x = (x_{A \setminus \{i\}}, x_i, x_{-A}) \succ (y_{A \setminus \{i\}}, x_i, x_{-A}) \succ (y_{A \setminus \{i\}}, y_i, x_{-A}) = (y_{A \setminus \{i\}}, y_i, y_{-A}) = y.$$

So $x \succ y$ by transitivity on X'. $\square$

4 Characterization of lexicographic preferences

Theorem 4.1 provides our characterization of lexicographic preferences. Following the argument of Debreu (1954) that such preferences over $\mathbb{R}^2$ cannot be represented by a utility function, the standard textbook case often looks at product sets of two coordinates. We consider that case separately in Corollary 4.2 and show that the robustness axiom A4 becomes superfluous.

Theorem 4.1. *Let $\succ$ be an asymmetric binary relation on $X = \times_{i \in I} X_i$. The preference relation $\succ$ is lexicographic if and only if it satisfies noncompensation A1, the weak transitivity requirement A2, weak decisiveness A3, and robustness A4.*

Proof. "Only if": Let $\succ$ be lexicographic on X with linear order $<_0$ on the coordinates I.
A1 holds: $P(x, y)$ and $P(y, x)$ completely determine the order between $x, y \in X$.
A2 holds: We assumed w.l.o.g. that for each $i \in I$ there are $a_i, b_i \in X_i$ with $a_i \succ_i b_i$. Let $X' = \times_{i \in I}\{a_i, b_i\}$. To show that $\succ$ is transitive on X', let $x, y, z \in X'$ have $x \succ y$ and $y \succ z$: there is an $i \in I$ with $x_i \succ_i y_i$ and $x_k \sim_k y_k$ for all $k <_0 i$. By definition of $X'_k = \{a_k, b_k\}$, the latter implies that $x_k = y_k$ for all $k <_0 i$. Likewise, there is a $j \in I$ with $y_j \succ_j z_j$ and $y_k = z_k$ for all $k <_0 j$. Now $i \neq j$: otherwise, $x_i \succ_i y_i, y_i \succ_i z_i$, and asymmetry of $\succ_i$ contradict $|X'_i| = |\{a_i, b_i\}| = 2$. If $i <_0 j$, then $x_k = z_k$ for all $k <_0 i$ and $x_i \succ_i y_i = z_i$, so $x \succ z$. If $j <_0 i$, then $x_k = z_k$ for all $k <_0 j$ and $x_j = y_j \succ_j z_j$, so $x \succ z$.
A3 holds: $x \sim y$ can only hold if $P(x, y) = P(y, x) = \emptyset$.
A4 holds: Let $x, y \in X$ satisfy the assumptions of A4. In particular, $x \succ y$ and there are distinct $i, j \in P(x, y)$. W.l.o.g. $j <_0 i$. Since coordinate i is checked after coordinate j in the lexicographic order, changing coordinate i does not affect the preference: $(z_i, x_{-i}) \succ (z_i, y_{-i})$ and hence $(z_i, x_{-i}) \succsim (z_i, y_{-i})$ for all $z_i \in X_i$.

"If": Let asymmetric relation $\succ$ satisfy the axioms. By A2, $\succ$ is transitive on a subset $X' = \times_{i \in I}\{a_i, b_i\} \subseteq X$ with $a_i \succ_i b_i$ for each $i \in I$. If A is a subset of I, define $e_A \in X'$ by

$$(e_A)_i = a_i \text{ if } i \in A \text{ and } (e_A)_i = b_i \text{ if } i \in I \setminus A. \tag{6}$$

If $A = \{j\}$ is a singleton, we often write e_j instead of $e_{\{j\}}$.

Define binary relation $<_0$ on I by $i <_0 j$ if and only if $e_i \succ e_j$. Then $<_0$ is a linear order: by A3, $e_i \succ e_j$ or $e_j \succ e_i$ for all distinct $i, j \in I$. And $<_0$ inherits transitivity and irreflexivity from $\succ$ on X'.

We show that relations $<_0$ on I and $\succ_i$ on X_i defined in (2) make $\succ$ lexicographic. By definition (1), we need to prove, for all $x, y \in X$:

$$x \succ y \Leftrightarrow \text{there is an } i \in I \text{ such that } \begin{cases} x_i \succ_i y_i \text{ is true, but} \\ y_j \succ_j x_j \text{ is false for all } j <_0 i. \end{cases}$$

($\Leftarrow$): Let $x, y \in X$. Let $i \in I$ be such that $x_i \succ_i y_i$ is true, but $y_j \succ_j x_j$ is false for all $j <_0 i$. To show that $x \succ y$, we use noncompensation A1 to simplify the argument:
STEP 1: Let $k \in I$. We prove by induction on the cardinality of B:

$$\text{for each } B \subset I: \quad \text{if } k <_0 j \text{ for all } j \in B, \text{ then } e_k \succ e_B. \tag{7}$$

This follows from Pareto dominance, Lemma 3.1(b), if $|B| = 0$ and from the definition of $<_0$ if $|B| = 1$. Now let $m \in \mathbb{N}, m \geq 2$, and suppose (7) is true if B has fewer than $m \in \mathbb{N}$ elements. We show that it is true also if B has m elements. Suppose, to the contrary, that $e_k \succ e_B$ is false. By axiom A3, $e_B \succ e_k$. Since $|P(e_B, e_k)|/|P(e_k, e_B)| = |B|/1 \geq 2$, robustness A4 implies that $e_{B\setminus\{\ell\}} \succsim e_k$ for some $\ell \in B$. But $B\setminus\{\ell\}$ has only m elements and satisfies the conditions in (7), so $e_k \succ e_{B\setminus\{\ell\}}$ by the induction hypothesis: a contradiction. Conclude that (7) is true by induction.
STEP 2: If $A, B \subseteq I$ satisfy $A \cap B = \emptyset$ and there is a $k \in A$ with $k <_0 j$ for all $j \in B$, then

$$e_A \succ e_B. \tag{8}$$

This follows from (7) if A is a singleton. If A has more than one element, (7) gives $e_k \succ e_B$. The Pareto property, Lemma 3.1(b), gives $e_A \succ e_k$. Transitivity of $\succ$ on X' gives $e_A \succ e_B$.
STEP 3: Let $A = P(x, y), B = P(y, x)$. Then $A \cap B = \emptyset$. By assumption, $i \in A$ and $i <_0 j$ for all $j \in B$. By (8), $e_A \succ e_B$. By noncompensation A1 with $w = e_A$ and $z = e_B$: $x \succ y$.

($\Rightarrow$): if $x \succ y$, Lemma 3.1 implies that $\{j \in I : x_j \succ_j y_j \text{ or } y_j \succ_j x_j\}$ is nonempty. It is finite, since I is. Using linear order $<_0$, it has a smallest element i. By construction, $x_j \sim_j y_j$ for all $j <_0 i$. Since $x \succ y$, direction "$\Leftarrow$" gives $x_i \succ_i y_i$, finishing the proof. $\square$

Robustness A4 imposes restrictions only on pairs $x, y \in X$ where $P(y, x)$ has at least one and, consequently, $P(x, y)$ has at least two elements. Since $P(x, y) \cap P(y, x) = \emptyset$ by asymmetry, such x and y must have at least three coordinates. So robustness holds vacuously on product sets of two coordinates:

Corollary 4.2 (Lexicographic preferences over products of two sets)**.** *Let $\succ$ be an asymmetric binary relation on $X = X_1 \times X_2$. The preference relation $\succ$ is lexicographic if and only if it satisfies axioms A1, A2, and A3.*

5 Logical independence

None of the axioms used in Theorem 4.1 is implied by the others:

Proposition 5.1. *Axioms A1, A2, A3, and A4 are logically independent.*

The result is proven by four examples of asymmetric binary relations on product sets, each of which violates exactly one of the axioms.

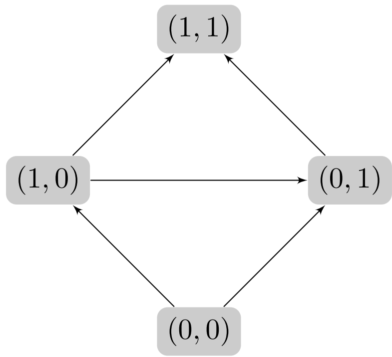

Figure 1: Violation of A2.

Example 5.1 (A1 violated)**.** The linear order $\succ$ on $X = \{0, 2\} \times \{0, 1, 4\}$ represented by utility function $u : X \to \mathbb{R}$ with $u(x) = x_1 + x_2$ has

$$(2, 4) \succ (0, 4) \succ (2, 1) \succ (2, 0) \succ (0, 1) \succ (0, 0).$$

So $\succ$ is asymmetric. The utility function is increasing in each coordinate, so $\succ_1$ and $\succ_2$ coincide with the usual order $>$ on integers.

Since $\succ$ is transitive, A2 holds; it has singleton equivalence classes, so A3 holds. A4 holds: it imposes restrictions only on products of three or more sets, here we have two.

A1 is violated: let $x = w = (2, 0), y = (0, 1), z = (0, 4)$. Then $P(x, y) = P(w, z) = \{1\}, P(y, x) = P(z, w) = \{2\}$, but $x \succ y$ and $z \succ w$. ◁

Example 5.2 (A2 violated)**.** Define a relation $\succ$ on $X = \times_{i=1}^{2}\{0, 1\}$ by $x \succ y$ if and only if there is a directed edge from y to x in Figure 1. So $\succ$ is asymmetric. Moreover, $1 \succ_1 0$, since there is an edge from $(0, 0)$ to $(1, 0)$ and from $(0, 1)$ to $(1, 1)$. Similarly, $1 \succ_2 0$, since there is an edge from $(0, 0)$ to $(0, 1)$ and from $(1, 0)$ to $(1, 1)$.

A1 holds: X has only two factors $X_i = \{0, 1\}$ and $1 \succ_i 0$ all $i = 1, 2$.

A3 holds: it demands only that $(0, 1)$ and $(1, 0)$ are ordered. Here, $(0, 1) \succ (1, 0)$.

A4 holds: it imposes restrictions only if $|I| \geq 3$, but here $|I| = 2$.

A2 is violated: $(1, 1) \succ (1, 0)$ and $(1, 0) \succ (0, 0)$, but not $(1, 1) \succ (0, 0)$. ◁

Example 5.3 (A3 violated)**.** For $n \in \mathbb{N}, n \geq 2$, let $\succ$ be the weak Pareto order on $\mathbb{R}^n$:

$$\text{for all } x, y \in \mathbb{R}^n : \qquad x \succ y \Leftrightarrow \begin{cases} x_i \geq y_i & \text{for all } i \in \{1, \ldots, n\}, \\ x_i > y_i & \text{for some } i \in \{1, \ldots, n\}. \end{cases}$$

So $\succ$ is asymmetric. For each $i \in \{1, \ldots, n\}$, $\succ_i$ is the usual order $>$ on $\mathbb{R}$, so

$$\text{for all } x, y \in \mathbb{R}^n : \qquad x \succ y \Leftrightarrow P(x, y) \neq \emptyset \text{ and } P(y, x) = \emptyset.$$

Hence, A1 holds. A2 holds: $\succ$ is transitive on $\mathbb{R}^n$ and in particular on $X' = \times_{i=1}^n \{0,1\}$. A4 holds vacuously: there are no x, y satisfying its assumptions.

A3 is violated: neither $(1,0,\ldots,0) \succ (0,\ldots,0,1)$, nor $(0,\ldots,0,1) \succ (1,0,\ldots,0)$. $\lhd$

Example 5.4 (A4 violated)**.** The linear order $\succ$ on $X = \times_{i=1}^3 \{0,1\}$ represented by utility function $u : X \to \mathbb{R}$ with $u(x) = 2x_1 + 3x_2 + 4x_3$ has

$$(1,1,1) \succ (0,1,1) \succ (1,0,1) \succ (1,1,0) \succ (0,0,1) \succ (0,1,0) \succ (1,0,0) \succ (0,0,0).$$

So $\succ$ is asymmetric. The utility function is increasing in each coordinate, so $1 \succ_i 0$ for all $i \in \{1,2,3\}$. Let $w = (2,3,4) \in \mathbb{R}^3$. Since $x_i, y_i \in \{0,1\}$ for all $i = 1,2,3$, we can rewrite

$$x \succ y \quad \Leftrightarrow \quad \sum_{i=1}^{3} w_i(x_i - y_i) > 0 \quad \Leftrightarrow \quad \sum_{i \in P(x,y)} w_i > \sum_{i \in P(y,x)} w_i.$$

So A1 holds. A2 holds: $\succ$ is transitive. As $\succ$ has singleton equivalence classes, A3 holds.

A4 is violated: $(1,1,0) \succ (0,0,1)$, so A4 would imply that $(0,1,0) \succsim (0,0,1)$ (if coordinate $i = 1$ were changed) or $(1,0,0) \succsim (0,0,1)$ (if coordinate $i = 2$ were changed). But both are false. $\lhd$

6 Deriving Fishburn's characterizations as corollaries

The purpose of this section is, firstly, to provide a critical assessment of the earlier characterizations of lexicographic preferences (Fishburn, 1975, 1976) and, secondly, to prove that these characterizations are corollaries of ours.

Informally, our critique of Fishburn (1975) is that his axioms are nice, but the domain is not. Why not? Well, we (and he) assumed w.l.o.g. that each attribute $i \in I$ was essential: $a_i \succ_i b_i$ for some $a_i, b_i \in X_i$. By noncompensation (A1), the order on $X' = \times_{i \in I} \{a_i, b_i\}$ with only *two* elements per coordinate tells all there is to know about the order between *any* pair of alternatives. For $A \subseteq I$, define $e_A \in X'$ as in (6): $(e_A)_i = a_i$ if $i \in A$, $(e_A)_i = b_i$ if $i \in I \setminus A$. By A1, for all $x, y \in X$: $x \succ y$ if and only if $e_{P(x,y)} \succ e_{P(y,x)}$.

If we only need information about two elements per coordinate $i \in I$, it seems redundant and unnatural to assume like Fishburn (1975) does — condition (9) below — that there are *three* elements $a_i, b_i, c_i \in X_i$ with $a_i \succ_i b_i$ and $b_i \succ_i c_i$.

So why does he have that assumption? It is needed only to make his proof technique work. Without going into detail, he identifies — like Mitra and Sen (2014) do in a recent alternative proof — some similarities with proofs of Arrow's classical impossibility theorem (Arrow, 1963) and the latter hinges critically on there being three alternatives. Therefore, he stresses that it plays a 'key role' (Fishburn, 1975, p. 416) in his proof.

Yet one easily imagines scenarios where each coordinate has only two feasible values. Mandler et al. (2012) is a case in point: they consider $n \in \mathbb{N}$ linearly ordered desirable properties. Each alternative gives rise to a list of n yes/no answers, depending on whether the corresponding property is satisfied; alternatives are ordered lexicographically. Such natural cases are not covered by Fishburn (1975). We did not want to impose a domain restriction solely in the service of our proof technique, so we proved Theorem 4.1 on the general domain. We now prove that our Theorem 4.1 implies the characterization in Fishburn (1975). The weak Pareto order on $\mathbb{R}^n$ in Example 5.3 shows that in this corollary, negative transitivity cannot be weakened to our mild transitivity property A2.

Corollary 6.1 (Fishburn (1975), Thm. 1)**.** *Let $\succ$ be an asymmetric and negatively transitive relation on product set $X = \times_{i \in I} X_i$. Let X be such that*

$$\textit{for all } i \in I, \textit{ there are } a_i, b_i, c_i \in X_i \textit{ with } a_i \succ_i b_i \textit{ and } b_i \succ_i c_i. \tag{9}$$

The relation $\succ$ is lexicographic if and only if it satisfies noncompensation A1.

Proof. Lexicographic preferences satisfy A1, so we only need to show that the proposition implies all other axioms of Theorem 4.1.

A2 HOLDS: Recall: since $\succ$ is asymmetric and negatively transitive, $\succ$ is transitive on X. By the domain restriction (9), for each $i \in I$ we can choose $a_i, b_i \in X_i$ with $a_i \succ_i b_i$. Since $\succ$ is transitive on X, it is transitive on $\times_{i \in I}\{a_i, b_i\} \subseteq X$.

A3 HOLDS: Let $x, y \in X$ be such that $x_i \succ_i y_i$ and $y_j \succ_j x_j$ for distinct $i, j \in I$. To show that $x \succ y$ or $y \succ x$, suppose to the contrary that $x \sim y$. We derive a contradiction.

The set X_i contains elements a_i, b_i, c_i with $a_i \succ_i b_i, b_i \succ_i c_i$. Since $\succ$ is transitive, $\succ_i$ is transitive, so $a_i \succ_i c_i$. A1 and $x \sim y$ imply that $(a_i, x_{-i}) \sim (b_i, y_{-i})$ and $(a_i, x_{-i}) \sim (c_i, y_{-i})$. Since $\succ$ is negatively transitive, $\sim$ is transitive, so $(b_i, y_{-i}) \sim (c_i, y_{-i})$. This contradicts the assumption that $b_i \succ_i c_i$.

A4 HOLDS: Let $x, y \in X$ be as in A4. Then $x \succ y$ and $x_i \succ_i y_i, x_j \succ_j y_j$ for distinct $i, j \in I$. By A3, $(x_i, y_j, x_{-ij}) \succ (y_i, x_j, x_{-ij})$ or $(y_i, x_j, x_{-ij}) \succ (x_i, y_j, x_{-ij})$. Consider the former (the latter proceeds analogously): $(x_i, y_j, x_{-ij}) \succ (y_i, x_j, x_{-ij})$. By (9), there are $a_i, b_i, c_i \in X_i$ with $a_i \succ_i b_i, b_i \succ_i c_i$ and $a_j, b_j, c_j \in X_j$ with $a_j \succ_j b_j, b_j \succ_j c_j$. So,

$$\begin{aligned}
(a_i, c_j, x_{-ij}) &\succ (b_i, b_j, x_{-ij}) && \text{by A1,}\\
(b_i, b_j, x_{-ij}) &\succ (c_i, c_j, y_{-ij}) && \text{by A1 and } x \succ y,\\
(a_i, c_j, x_{-ij}) &\succ (c_i, c_j, y_{-ij}) && \text{by transitivity of } \succ,\\
(c_j, x_{-j}) = (x_i, c_j, x_{-ij}) &\succ (y_i, c_j, y_{-ij}) = (c_j, y_{-j}) && \text{by A1.}
\end{aligned}$$

A1 and $(c_j, x_{-j}) \succ (c_j, y_{-j})$ give $(z_j, x_{-j}) \succ (z_j, y_{-j})$ and in particular $(z_j, x_{-j}) \succsim (z_j, y_{-j})$ for *all* $z_j \in X_j$: A4 holds. □

Fishburn (1976) characterizes, like we do, lexicographic preferences on the general domain. But one of his axioms, superadditivity, is a technical assumption that is *(i)* offered without supporting intuitive motivation and *(ii)* very strong and rather explicitly lexicographic. For instance, in $\mathbb{R}^n$, if the standard basis vectors are ordered

$$e_1 = (1, 0, \ldots, 0) \succ e_2 = (0, 1, 0, \ldots, 0) \succ \cdots \succ e_n = (0, \ldots, 0, 1),$$

it explicitly requires that $e_1 = (1, 0, \ldots, 0) \succ (0, 1, 1, \ldots, 1)$, stressing the lexicographic dominance of the first coordinate. Of course, that is the conclusion we *want* for a lexicographic preference, but it would have been nice if the axioms left some work still to be done and didn't need to assume the desired conclusion.

To be precise and to show that his result is a corollary of ours, we need some of his notation. Define a relation $\rhd$ on subsets of the attributes I as follows: for all $A, B \subseteq I$:

$$A \rhd B \Leftrightarrow A \cap B = \emptyset \text{ and } x \succ y \text{ for all } x, y \in X \text{ with } (P(x, y), P(y, x)) = (A, B). \qquad (10)$$

With the noncompensation axiom A1, this is equivalent with

$$A \rhd B \Leftrightarrow A \cap B = \emptyset \text{ and } x \succ y \text{ for } \textit{some } x, y \in X \text{ with } (P(x, y), P(y, x)) = (A, B). \quad (11)$$

By definition of $\succ_i$: $\{i\} \rhd \emptyset$ for all $i \in I$. Since $\succ$ is asymmetric, so is $\rhd$. Fishburn (1976) calls relation $\succ$:

A5 *attribute acyclic* if $\rhd$ is acyclic: there are no $n \in \mathbb{N}$ and subsets $A_1, \ldots, A_n$ of I with $A_i \rhd A_{i+1}$ for all $i \in \mathbb{N}, i < n$, and $A_n \rhd A_1$.

A6 *superadditive* if, for all $A, B, C, D \subseteq I$,

$$\left.\begin{array}{rcl} A & \rhd & B, \\ C & \rhd & D, \\ (A \cup C) \cap (B \cup D) & = & \emptyset \end{array}\right\} \quad \Rightarrow \quad A \cup C \rhd B \cup D.$$

A7 *decisive* if for all $A, B \subseteq I$ where $A \cap B = \emptyset$ and at least one of them is nonempty: $A \rhd B$ or $B \rhd A$.

By A1, A5, and A7, $\rhd$ linearly orders the attributes: using $\{i\} \rhd \emptyset$ for all i and relabeling if necessary, we may assume w.l.o.g. that $I = \{1, \ldots, n\}$ for $n \in \mathbb{N}, n \geq 2$, and that $\{1\} \rhd \{2\} \rhd \cdots \rhd \{n\} \rhd \emptyset$. Superadditivity A6 yields $A \rhd \emptyset$ for all nonempty $A \subseteq I$. It follows easily that $\succ$ is lexicographic with coordinates ordered $1 <_0 2 <_0 \cdots <_0 n$. By A1, it suffices to show that for all nonempty, disjoint $A, B \subseteq I$ with $\min A < \min B$: $A \rhd B$. So let $i = \min A$. If $|A| = 1, A = \{i\}$. If $|A| > 1$, A6 and $\{i\} \rhd \{b\}$ for all $b \in B$ give $\{i\} \rhd B$. With $A \rhd \emptyset$, A6 gives $A = A \cup \{i\} \rhd \emptyset \cup B = B$.

In fact, Fishburn's proof that axioms A1, A5, A6, and A7 make the order lexicographic is less detailed than ours in the preceding paragraph and has only five lines; see Fishburn (1976), pp. 401–402, "Suppose next…satisfies C12."

In contrast, the aim of our robustness axiom A4 is to provide a requirement that is both less demanding than the technical superadditive assumption and has a straightforward intuition. Recall the discussion in section 3: informally, it requires that if many attributes are in favor of x and only a few in favor of y, then a preference of x over y is robust to (not reversed by) a small change of mind, turning an attribute in favor of x to an indifference. The characterization in Fishburn (1976) is a corollary of Theorem 4.1:

Corollary 6.2 (Fishburn (1976), Thm. 1(f))**.** *Let $\succ$ be an asymmetric relation on product set $X = \times_{i \in I} X_i$. Relation $\succ$ is lexicographic if and only if it satisfies noncompensation A1, attribute acyclicity A5, superadditivity A6, and decisiveness A7.*

Proof. That lexicographic preferences satisfy the conditions is straightforward. For the opposite direction: we summarized Fishburn's proof above. But to clarify that it is a special case of our Theorem 4.1, we derive our axioms from his.

The proof that A2 holds is not difficult, but has many steps; it is in the appendix.

As its name suggests, weak decisiveness A3 is implied by decisiveness: let $x, y \in X$ satisfy $x_i \succ_i y_i$ and $y_j \succ_j x_j$ for distinct $i, j \in I$. Then $P(x,y) \neq \emptyset, P(y,x) \neq \emptyset$, and $P(x,y) \cap P(y,x) = \emptyset$ by asymmetry of relations $\succ_k$. By decisiveness, $P(x,y) \rhd P(y,x)$ or $P(y,x) \rhd P(x,y)$. In particular, $x \succ y$ or $y \succ x$.

Robustness A4 is implied by noncompensation A1 and superadditivity A6: Let $x, y \in X$ be as in A4. Then $x \succ y$, $A = P(x,y)$ has at least two elements, $B = P(y,x)$ is nonempty. By asymmetry of relations $\succ_k, k \in I$, $A \cap B = \emptyset$. With A1, (11), and $x \succ y$, it follows that $A \rhd B$. Choose distinct $i, j \in A$. Since $(A \setminus \{i\}) \cup (A \setminus \{j\}) = A$, superadditivity implies that $B \rhd A \setminus \{i\}$ is false or $B \rhd A \setminus \{j\}$ is false. Assume w.l.o.g. that $B \rhd A \setminus \{i\}$ is false. To prove A4, we show that $(z_i, x_{-i}) \succsim (z_i, y_{-i})$ for all $z_i \in X_i$. Let $z_i \in X_i, x' = (z_i, x_{-i}), y' = (z_i, y_{-i})$. Then $P(x', y') = A \setminus \{i\} \neq \emptyset, P(y', x') = B \neq \emptyset$. Since $B \rhd A \setminus \{i\}$ is false, (11) implies that $y' \succ x'$ is false. So $x' \succsim y'$, proving A4. □

We showed above that noncompensation and superadditivity imply robustness. But superadditivity is strictly more demanding; see Example 6.1 below. That superadditivity is more demanding should not come as a surprise. For instance, it imposes restrictions already in the case of a product $X = X_1 \times X_2$ of two sets, while robustness A4 does not.

Most effort in this section has been devoted to proving that the assumptions in Fishburn's earlier characterizations are stronger than (imply) ours. There is one exception. Attribute acyclicity is a weaker assumption than our transitivity requirement A2:

Lemma 6.3. *If asymmetric relation $\succ$ satisfies transitivity requirement A2 on $X' = \times_{i \in I}\{a_i, b_i\} \subseteq X$ with $a_i \succ_i b_i$ for all $i \in I$, then $\succ$ is attribute acyclic A5.*

Proof. Let $n \in \mathbb{N}$ and let subsets $A_1, \ldots, A_n$ of I satisfy $A_i \rhd A_{i+1}$ for all $i \in \mathbb{N}, i < n$. If $n = 1$, acyclicity of $\succ$ rules out that $A_n \rhd A_1$. So let $n > 1$. As in the proof of Theorem

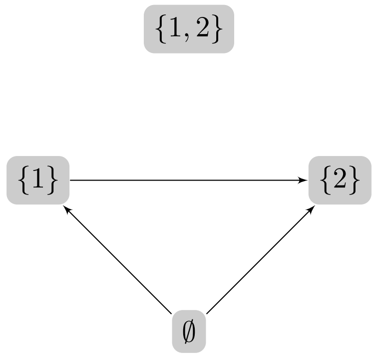

Figure 2: Relation $\rhd$ in Example 5.2. An arrow from, e.g., $\{1\}$ to $\{2\}$ means $\{2\} \rhd \{1\}$.

4.1 define, for each subset A of I, alternative $e_A \in X'$ by $(e_A)_i = a_i$ if $i \in A$ and $(e_A)_i = b_i$ if $i \in I \setminus A$. By definition of $\rhd$, $e_{A_i} \succ e_{A_{i+1}}$ for all $i \in \mathbb{N}, i < n$. By transitivity of $\succ$ on X', $e_{A_1} \succ e_{A_n}$. If $A_1 \cap A_n \neq \emptyset$, (10) rules out $A_n \rhd A_1$ and we are done. If $A_1 \cap A_n = \emptyset$, $e_{A_1} \succ e_{A_n}$ and asymmetry of $\succ$ rule out that $e_{A_n} \succ e_{A_1}$ and hence that $A_n \rhd A_1$. $\square$

So it is natural to ask if Theorem 4.1 remains valid if transitivity requirement A2 is weakened to attribute acyclicity A5. Example 5.2 shows that this is not the case. Nor can one replace superadditivity with our weaker robustness assumption A4 in Corollary 6.2:

Example 6.1. Recall that relation $\succ$ in Example 5.2 is not lexicographic: $1 \succ_1 0$ and $1 \succ_2 0$, but not $(1,1) \succ (0,0)$.

The corresponding relation $\rhd$ depicted in Figure 2 is clearly acyclic: $\succ$ is attribute acyclic. It also satisfies all axioms in Theorem 4.1 except A2. Conclude: A2 cannot be replaced by attribute acyclicity in the characterization of lexicographic preferences in Theorem 4.1.

Relation $\succ$ is not superadditive: $\{1\} \rhd \emptyset$ and $\{2\} \rhd \emptyset$, so superadditivity would imply $\{1,2\} = \{1\} \cup \{2\} \rhd \emptyset$. Yet neither $\{1,2\} \rhd \emptyset$, nor $\emptyset \rhd \{1,2\}$. But $\succ$ does satisfy all other axioms in Corollary 6.2 plus robustness A4. Hence, superadditivity cannot be weakened to robustness in the characterization of lexicographic preferences in Corollary 6.2. $\lhd$

It is, however, possible in Theorem 4.1 to weaken transitivity property A2 on $X' = \times_{i \in I}\{a_i, b_i\} \subseteq X$ to acyclicity, provided that one strengthens weak decisiveness:

Lemma 6.4. *Let $\succ$ be an asymmetric relation on $X = \times_{i \in I} X_i$. Transitivity property A2 and weak decisiveness A3 hold if and only if $\succ$ satisfies the following two properties:*

(a) $\succ$ is acyclic on a product set $X' = \times_{i \in I}\{a_i, b_i\} \subseteq X$, where $a_i \succ_i b_i$ for each $i \in I$;

(b) For all $x, y \in X$, if $x_i \succ_i y_i$ or $y_i \succ_i x_i$ for some $i \in I$, then $x \succ y$ or $y \succ x$.

Proof. "Only if": Asymmetry of $\succ$ and transitivity on X' imply acyclicity on X': (a) holds. For (b), let $x, y \in X$ have $x_i \succ_i y_i$ or $y_i \succ_i x_i$ for some $i \in I$: $P(x, y) \cup P(y, x) \neq \emptyset$. If both $P(x, y)$ and $P(y, x)$ are nonempty, A3 gives $x \succ y$ or $y \succ x$. If only one, say $P(x, y)$, is nonempty, Pareto Lemma 3.1 gives $x \succ y$.

"If": (b) implies A3. For A2, let $x, y, z \in X'$ satisfy $x \succ y$ and $y \succ z$. If $P(x, z) \cup P(z, x) = \emptyset$, then for all $i \in I$, $X'_i = \{a_i, b_i\}$ with $a_i \succ_i b_i$ implies that $x_i = z_i$. But $x \succ y$ and $y \succ z = x$ contradict acyclicity. So $P(x, z) \cup P(z, x) \neq \emptyset$ and $x \succ z$ or $z \succ x$ by (b). Acyclicity rules out $z \succ x$, so $x \succ z$, proving that $\succ$ is transitive on X'. $\square$

Example 5.2 shows that it is not possible to relax transitivity in A2 to acyclicity (a) without strengthening another axiom in Theorem 4.1: Figure 1 shows that $\succ$ is acyclic and it satisfies all axioms in Theorem 4.1 except A2, but $\succ$ is not lexicographic.

7 Concluding remarks

The main new axiom in our characterization of lexicographic preferences is:

> *Robustness:* For all $x, y \in X$,
>
> if $x \succ y$, $P(y, x) \neq \emptyset$, and $|P(x, y)|/|P(y, x)| \geq 2$,
>
> then there is an $i \in P(x, y)$ with $(z_i, x_{-i}) \succsim (z_i, y_{-i})$ for all $z_i \in X_i$.

Informally, if many attributes are in favor of x and only a few in favor of y, then a small change of mind — turning some attribute in favor of x to an indifference — is not enough to reverse a preference of x over y.

Our argument (see the proof of Thm. 4.1) that lexicographic preferences satisfy robustness was as follows: let $x, y \in X$ satisfy the assumptions of robustness. In particular, $x \succ y$ and there are distinct $i, j \in P(x, y)$. W.l.o.g. $j <_0 i$. Since coordinate i is checked after coordinate j in the lexicographic order, changing coordinate i does not affect the preference. But this means that the preference is not affected by *anything* you do to the i-th coordinate: the requirement

> there is an $i \in P(x, y)$ with $(z_i, x_{-i}) \succsim (z_i, y_{-i})$ for all $z_i \in X_i$

in the last sentence of the definition of robustness can be strengthened to a strict preference:

> there is an $i \in P(x, y)$ with $(z_i, x_{-i}) \succ (z_i, y_{-i})$ for all $z_i \in X_i$,

or to an arbitrary change of mind concerning the i-th coordinate:

> there is an $i \in P(x, y)$ with $(z_i, x_{-i}) \succsim (w_i, y_{-i})$ for all $z_i, w_i \in X_i$,

or even both:

there is an $i \in P(x,y)$ with $(z_i, x_{-i}) \succ (w_i, y_{-i})$ for all $z_i, w_i \in X_i$.

All these variants, however, are heavier requirements than our robustness axiom: fewer preferences satisfy them. In characterization theorems it is therefore common practice to impose the least demanding assumption (here, our formulation of the robustness axiom).

Talking about robustness informally, we used phrases like "there are many attributes in favor of x and only a few in favor of y". The axiom models this by requiring that the attributes in favor of x are at least twice as many as those in favor of y: $|P(x,y)|/|P(y,x)| \geq 2$. Comparing Theorem 4.1 with Corollary 4.2 sheds light on the question 'Why (at least) twice?'. Although robustness is not needed to characterize lexicographic preferences over products of two sets (Cor. 4.2), Example 5.4 illustrates the need for an additional axiom that kicks in when there are three or more coordinates. Robustness does precisely that: it imposes restrictions only on pairs $x, y \in X$ where $P(y,x)$ has at least one and, consequently, $P(x,y)$ has at least two elements. By asymmetry, $P(x,y)$ and $P(y,x)$ are disjoint, so such x and y have at least three attributes!

We considered lexicographic preferences over product sets of finitely many coordinates. It might be of mathematical interest to extend the analysis to infinite index sets. But we refrained from doing so because it will require arbitrary modeling choices that lead to counterintuitive results or mathematical complications that go far beyond what is likely to concern actual decision makers. Let's elaborate briefly.

The key element in lexicographic preferences is the first element (given an order $<_0$ on the coordinates) in the set $S = \{i \in I : x_i \succ_i y_i \text{ or } y_i \succ_i x_i\}$ of attributes that make a difference. But if we stick to a linear order $<_0$ on I and the set S is infinite, it need not have a first element and we loose the desirable result (Lemma 3.1) that lexicographic preferences extend the Pareto order:

Example 7.1. Let $I = \{-1, -2, -3, \ldots\}$ have its usual order $\cdots <_0 -3 <_0 -2 <_0 -1$. For each $i \in I$, $X_i = \{0, 1\}$ with $1 \succ_i 0$. Define lexicographic relation $\succ$ on $X = \times_{i \in I}\{0,1\}$ as follows:

$$\text{for all } x, y \in X: \qquad x \succ y \quad \Leftrightarrow \quad \begin{cases} S = \{j \in I : x_j \succ_j y_j \text{ or } y_j \succ_j x_j\} \text{ is nonempty,} \\ S \text{ has a first element } i, \text{ and} \\ x_i \succ_i y_i. \end{cases}$$

If $x \in X$ has some, but finitely many coordinates equal to 1, then $x \succ \mathbf{0} = (0,0,0,\ldots)$. But if x has infinitely many coordinates equal to 1, the set

$$S = \{j \in I : x_j \succ_j 0 \text{ or } 0 \succ_j x_j\} = \{j \in I : x_j = 1\}$$

has no first element: neither $x \succ \mathbf{0}$, nor $\mathbf{0} \succ x$. Pareto dominance would require $x \succ \mathbf{0}$. ◁

Alternatively, one can change the linear order $<_0$ to a well-ordering of I. Then, by definition, every nonempty subset of I *will* have a first element. But the well-ordering theorem, the result that every I can be well-ordered, is equivalent with the Axiom of Choice; see for instance Munkres (2000, Ch. 1). This is a nonconstructive result; roughly speaking, well-orders typically cannot be explicitly defined, placing lexicographic reasoning with well-orders outside the scope of your common-or-garden variety decision maker. Fishburn (1974, p. 1450) discusses further mathematical complications.

A Appendix: an omitted proof

We prove that noncompensation A1, attribute acyclicity A5, superadditivity A6, and decisiveness A7 imply the transitivity property A2. So let $X' = \times_{i \in I}\{a_i, b_i\} \subseteq X$ with $a_i \succ_i b_i$ for each $i \in I$. With e_A as usual, see (6), we must prove for all $A, B, C \subseteq I$:

$$\text{if } e_A \succ e_B \text{ and } e_B \succ e_C, \text{ then } e_A \succ e_C. \tag{12}$$

We start by showing that for all $A, B, C, D \subseteq I$, the following implications hold:

$$A \neq \emptyset \implies A \rhd \emptyset \tag{13}$$

$$A \rhd B \implies A \neq \emptyset \tag{14}$$

$$A \rhd B \cup C \implies A \rhd B \tag{15}$$

$$\begin{cases} (A \cup B) \cap (C \cup D) = \emptyset \\ A \cup B \rhd C \cup D \end{cases} \implies A \rhd C \cup D \text{ or } B \rhd C \cup D \tag{16}$$

$$\begin{cases} A \rhd B \\ B \rhd C \\ A \cap C = \emptyset \end{cases} \implies A \rhd C \tag{17}$$

Proof. (13): For each $i \in I$, $e_i \succ e_\emptyset$, so $\{i\} \rhd \emptyset$ by (11). Superadditivity gives (13).
(14): If $B = \emptyset$, attribute acyclicity rules out $A = \emptyset$. If $B \neq \emptyset$, then $B \rhd \emptyset$ by (13). Since $A \rhd B$, attribute acyclicity once again rules out $A = \emptyset$.
(15): By (14), $A \neq \emptyset$. If $C = \emptyset$, $A \rhd B \cup C = B$. If $C \neq \emptyset$, $C \rhd \emptyset$ by (13). If $A \rhd B$ were false, decisiveness gives $B \rhd A$. Superadditivity gives $B \cup C \rhd A \cup \emptyset = A$. But $A \rhd B \cup C$, contradicting attribute acyclicity.
(16): If $A = \emptyset$ or $B = \emptyset$, this is trivial, so let both be nonempty. If — to the contrary — neither $A \rhd C \cup D$ nor $B \rhd C \cup D$ were true, (13) makes $C \cup D$ nonempty. By decisiveness, $C \cup D \rhd A$ and $C \cup D \rhd B$. By superadditivity, $C \cup D \rhd A \cup B$. But $A \cup B \rhd C \cup D$, contradicting attribute acyclicity.
(17): By (14), $A \neq \emptyset$. By decisiveness, $A \rhd C$ or $C \rhd A$. Since $A \rhd B$ and $B \rhd C$, attribute acyclicity rules out $C \rhd A$. So $A \rhd C$. □

This leaves us properly equipped for the proof of (12): let $A, B, C \subseteq I$ satisfy $e_A \succ e_B$ and $e_B \succ e_C$. To show: $e_A \succ e_C$. Consider six disjoint sets (see Fig. 3):

$$S_1 = A \cap B^c \cap C^c \qquad S_2 = A^c \cap B \cap C^c \qquad S_3 = A^c \cap B^c \cap C$$
$$S_4 = A \cap B \cap C^c \qquad S_5 = A \cap B^c \cap C \qquad S_6 = A^c \cap B \cap C$$

Note that

$$\text{for all } X, Y \subseteq I \text{ with } X \cap Y = \emptyset: \qquad e_X \succ e_Y \overset{(11)}{\iff} X \rhd Y. \tag{18}$$

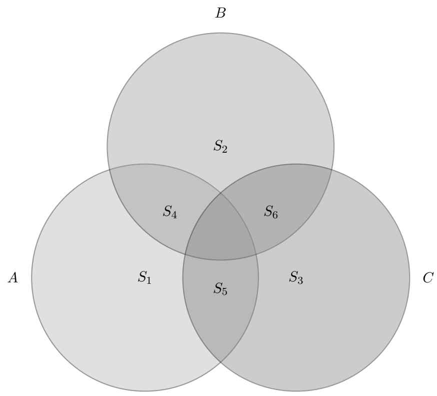

Figure 3: Cutting up A, B, and C.

Therefore,

$$\begin{aligned} e_A \succ e_B &\overset{A1}{\iff} e_{A\setminus B} = e_{S_1\cup S_5} \succ e_{S_2\cup S_6} = e_{B\setminus A} \overset{(18)}{\iff} S_1 \cup S_5 \rhd S_2 \cup S_6 \\ &\overset{(16)}{\implies} S_1 \rhd S_2 \cup S_6 \text{ or } S_5 \rhd S_2 \cup S_6 \\ &\overset{(15)}{\implies} (S_1 \rhd S_2 \text{ and } S_1 \rhd S_6) \quad \text{or} \quad (S_5 \rhd S_2 \text{ and } S_5 \rhd S_6) \end{aligned} \tag{19}$$

Likewise, writing $B \setminus C = S_2 \cup S_4$ and $C \setminus B = S_3 \cup S_5$:

$$e_B \succ e_C \implies (S_2 \rhd S_3 \text{ and } S_2 \rhd S_5) \quad \text{or} \quad (S_4 \rhd S_3 \text{ and } S_4 \rhd S_5). \tag{20}$$

Finally, with $A \setminus C = S_1 \cup S_4$ and $C \setminus A = S_3 \cup S_6$:

$$e_A \succ e_C \iff S_1 \cup S_4 \rhd S_3 \cup S_6. \tag{21}$$

We prove the right-hand side of (21). By (19) and (20), we need to distinguish four cases:

1. $(S_1 \rhd S_2$ and $S_1 \rhd S_6)$ and $(S_2 \rhd S_3$ and $S_2 \rhd S_5)$:

   $S_1 \rhd S_6$ and $S_1 \rhd S_2 \rhd S_3$, so $S_1 \rhd S_3$ by (17). By superadditivity, $S_1 \rhd S_3 \cup S_6$. If $S_4 = \emptyset$, $S_1 \cup S_4 = S_1 \rhd S_3 \cup S_6$. If $S_4 \neq \emptyset$, $S_4 \rhd \emptyset$ by (13); superadditivity then gives $S_1 \cup S_4 \rhd (S_3 \cup S_6) \cup \emptyset = S_3 \cup S_6$.

2. $(S_1 \rhd S_2$ and $S_1 \rhd S_6)$ and $(S_4 \rhd S_3$ and $S_4 \rhd S_5)$:

   $S_1 \rhd S_6$, $S_4 \rhd S_3$, and superadditivity give $S_1 \cup S_4 \rhd S_3 \cup S_6$.

3. $(S_5 \rhd S_2$ and $S_5 \rhd S_6)$ and $(S_2 \rhd S_3$ and $S_2 \rhd S_5)$:

   This is impossible: $S_5 \rhd S_2$ and $S_2 \rhd S_5$ contradict attribute acyclicity.

4. $(S_5 \rhd S_2$ and $S_5 \rhd S_6)$ and $(S_4 \rhd S_3$ and $S_4 \rhd S_5)$:

   $S_4 \rhd S_3$ and $S_4 \rhd S_5 \rhd S_6$, so $S_4 \rhd S_6$ by (17). Arguing as in the first case, $S_1 \cup S_4 \rhd S_3 \cup S_6$.

B Appendix: further examples of robust preferences

Robustness A4 is a new axiom, so let us provide two more plausible examples of asymmetric preferences that satisfy it (and noncompensation A1). The first is a parametric class of outnumbering relations that generalizes our earlier example of simple majority preferences. The second is a stylized example of dialectic reasoning where an attribute in favor of one alternative can be trumped by a 'counterargument' supporting the other one.

Example B.1 (Outnumbering)**.** Let $\alpha \in \{0, 1, 2, \dots\}$. Given asymmetric relations $(\succ_i)_{i\in I}$ on finitely many coordinate sets $(X_i)_{i\in I}$, define, for all $x, y \in X = \times_{i\in I} X_i$:

$$x \succ y \iff \begin{cases} (P(x,y) \neq \emptyset \text{ and } P(y,x) = \emptyset) \text{ or} \\ |P(x,y)| > |P(y,x)| + \alpha. \end{cases} \tag{22}$$

So x is preferred to y if x Pareto dominates y or if the attributes in favor of x outnumber those in favor of y by more than α. Simple majority preferences (4) are the special case with $\alpha = 0$. Relation $\succ$ is asymmetric and noncompensation A1 clearly holds. The first line of (22) makes $\succ_i$ unambiguously defined: it satisfies (2). Also robustness A4 holds: Let $x, y \in X$ be such that $x \succ y, P(y,x) \neq \emptyset$, and $|P(x,y)|/|P(y,x)| \geq 2$. So we are in the second line of (22). Let $i \in P(x,y), z_i \in X_i$, and $x' = (z_i, x_{-i}), y' = (z_i, y_{-i})$. Then

$$\begin{aligned} x \succ y &\Rightarrow |P(x,y)| = |P(x',y')| + 1 > |P(y,x)| + \alpha = |P(y',x')| + \alpha \\ &\Rightarrow |P(y',x')| < |P(x',y')| + 1 - \alpha \quad (\in \mathbb{Z}) \\ &\Rightarrow |P(y',x')| \leq |P(x',y')| - \alpha \leq |P(x',y')| + \alpha \\ &\Rightarrow (\text{not } y' \succ x'), \text{ i.e., } x' = (z_i, x_{-i}) \succsim y' = (z_i, y_{-i}). \end{aligned}$$

Since z_i was arbitrary, robustness holds. ◁

Example B.2 (Reasoning by valid counterargument)**.** Let $I = \{1, \dots, n\}$ for some $n \in \mathbb{N}, n \geq 2$, and for each $i \in I$, $X_i = \{0, 1\}$ with $1 \succ_i 0$. Given an attribute $i \in I$, let $v(i) \subseteq I \setminus \{i\}$ be a possibly empty set of 'valid counterarguments'; roughly speaking, if i is an argument raised to support a preference of one alternative over another, the 'valid counterarguments' might be used to argue in favor of the opposite alternative. More precisely, define, for all $x, y \in X$:

$$x \succ y \iff \begin{cases} v(j) \cap P(x,y) \neq \emptyset & \text{for each } j \in P(y,x), \\ v(k) \cap P(y,x) = \emptyset & \text{for some } k \in P(x,y). \end{cases}$$

So x is preferred to y if *(i)* each attribute j in favor of y has a valid counterargument in favor of x and *(ii)* some attribute k in favor of x has no valid counterargument in favor of y. Roughly speaking, all arguments in favor of y are trumped by counterarguments in favor of x, but the converse is not true.

The second requirement makes $\succ$ asymmetric. Since $\succ$ respects Pareto dominance, relation $\succ_i$ is unambiguously defined: it satisfies (2). Noncompensation A1 is obvious. To verify robustness A4, let $x, y \in X$ have $x \succ y, P(y,x) \neq \emptyset$, and $|P(x,y)|/|P(y,x)| \geq 2$.

Firstly, let $|P(x,y)| = 2$. Then $|P(y,x)| = 1$, so $P(y,x) = \{j\}$ for some $j \in I$.
Case 1: There is a $k \in P(x,y)$ with $k \in v(j)$ and $j \in v(k)$. Since $|P(x,y)| = 2$, there is an $i \in P(x,y) \setminus \{k\}$. For each $z_i \in X_i$, $x' = (z_i, x_{-i})$ and $y' = (z_i, y_{-i})$ have $P(x', y') = \{k\}$ and $P(y', x') = \{j\}$. Since j and k are valid counterarguments to each other, $x' \sim y'$.
Case 2: There is no $k \in P(x,y)$ with $k \in v(j)$ and $j \in v(k)$. Since $x \succ y$, there is a $k \in v(j) \cap P(x,y)$. By assumption, $j \notin v(k) \cap P(y,x)$. As above, there is an $i \in P(x,y) \setminus \{k\}$. For each $z_i \in X_i$, $x' = (z_i, x_{-i})$ and $y' = (z_i, y_{-i})$ have $P(x', y') = \{k\}$ and $P(y', x') = \{j\}$. Since k is valid against j, but j is not valid against k, $x' \succ y'$.

So for all $z_i \in X_i$, $x' = (z_i, x_{-i}) \succsim (z_i, y_{-i}) = y'$, i.e., robustness holds if $|P(x,y)| = 2$.

Secondly, let $|P(x,y)| > 2$. Since $x \succ y$, there is a $k \in P(x,y)$ with $v(k) \cap P(y,x) = \emptyset$. We show that some attribute $i \in P(x,y) \setminus \{k\}$ is 'redundant' in the sense that every $j \in P(y,x)$ has a valid counterargument in $P(x,y)$ *other* than i:

$$\text{There is an } i \in P(x,y) \setminus \{k\} \text{ with } v(j) \cap (P(x,y) \setminus \{i\}) \neq \emptyset \text{ for all } j \in P(y,x). \tag{23}$$

Suppose not. Then for each $i \in P(x,y) \setminus \{k\}$, we can fix a $j_i \in P(y,x)$ with $v(j_i) \cap (P(x,y) \setminus \{i\}) = \emptyset$. Since j_i must have a valid counterargument, it follows that $v(j_i) \cap P(x,y) = \{i\}$. This equality makes the function $f : i \mapsto j_i$ from $P(x,y) \setminus \{k\}$ to $P(y,x)$ injective. Hence, $|P(x,y) \setminus \{k\}| = |P(x,y)| - 1 \leq |P(y,x)|$. Since $|P(y,x)| \leq |P(x,y)|/2$, this implies that $|P(x,y)| \leq 2$, contradicting our assumption that $|P(x,y)| > 2$.

Pick a 'redundant' $i \in P(x,y) \setminus \{k\}$ as in (23). For each $z_i \in X_i$, $x' = (z_i, x_{-i})$ and $y' = (z_i, x_{-i})$ have $P(x', y') = P(x,y) \setminus \{i\}$ and $P(y', x') = P(y,x)$. By (23), $v(j) \cap P(x', y') \neq \emptyset$ for each $j \in P(x', y')$. By definition of k, $v(k) \cap P(x', y') = \emptyset$. So $x' \succ y'$ and, in particular, $x' = (z_i, x_{-i}) \succsim (z_i, y_{-i}) = y'$. As $z_i \in X_i$ was arbitrary, robustness holds if $|P(x,y)| > 2$. $\triangleleft$

The relations in our final example, on reasoning by valid counterargument, are the topic of our ongoing study.

References

K. J. Arrow. *Social Choice and Individual Values.* Wiley, 2nd edition, 1963.

G. Debreu. Representation of a preference ordering by a numerical function. In R. Thrall, C. Coombs, and R. Davis, editors, *Decision processes*, chapter XI, pages 159–165. Chapman & Hall, 1954.

P. C. Fishburn. Lexicographic orders, utilities and decision rules: A survey. *Management Science*, 20:1442 – 1471, 1974.

P. C. Fishburn. Axioms for lexicographic preferences. *Review of Economic Studies*, 42: 415–419, 1975.

P. C. Fishburn. Noncompensatory preferences. *Synthese*, 33:393–403, 1976.

P. C. Fishburn. Nontransitive preferences in decision theory. *Journal of Risk and Uncertainty*, 4:113–134, 1991.

M. Mandler, P. Manzini, and M. Mariotti. A million answers to twenty questions: Choosing by checklist. *Journal of Economic Theory*, 147:71–92, 2012.

J. E. Martínez-Legaz. Lexicographic utility and orderings. In S. Barberà, P. J. Hammond, and C. Seidl, editors, *Handbook of Utility Theory. Volume 1, Principles*, chapter 9, pages 345–369. Kluwer Academic Publishers, 1998.

M. Mitra and D. Sen. An alternative proof of Fishburn's axiomatization of lexicographic preferences. *Economics Letters*, 124:168–170, 2014.

J. R. Munkres. *Topology.* Prentice Hall, Inc., 2nd edition, 2000.

M. Voorneveld. Characterization of Pareto dominance. *Operations Research Letters*, 31: 7–11, 2003.

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