PART IV

NEW RESULTS IN NONLINEAR PREFERENCE THEORY

The dozen papers in this part illustrate the diversity and applicability of nonlinear preference and utility theories. Chew, Segal, and Nau generalize earlier nonlinear theories in interesting ways. The papers by LaValle & Fishburn, Chew & Epstein, Nakamura, Gilboa, Hazen, and Karni & Safra analyze and consider application of various nonlinear theories. The other three papers are primarily expository: Blume et al. lead off with a tour through lexicographic choice models under uncertainty; Quiggin offers a survey of nonlinear models that is organized by alternative notions of dominance and independence; and LaValle concludes Part IV by identifying the class of nonlinear models for decision under uncertainty in which complex outcomes can be replaced by simpler proxies without violating basic preferences between acts.

Annals of Operations Research 19(1989)231–246

AN OVERVIEW OF LEXICOGRAPHIC CHOICE UNDER UNCERTAINTY*

Lawrence BLUME

Department of Economics, University of Michigan, Ann Arbor, Michigan 48109, USA

Adam BRANDENBURGER

Harvard Business School, Boston, Massachusetts 02163, USA

and

Eddie DEKEL

Department of Economics, University of California, Berkeley, California 94720, USA

Abstract

This overview focuses on lexicographic choice under conditions of uncertainty. First, lexicographic versions of traditional (von Neumann–Morgenstern) expected utility theory are described where the usual Archimedean axiom is weakened. The role of these lexicographic variants in explaining some well-known "paradoxes" of choice theory is reviewed. Next, the significance of lexicographic choice for game theory is discussed. Finally, some lexicographic extensions of the classical maximin decision rule are described.

1. Introduction

An underlying theme in most of the economic theory of choice behavior is the notion that it is always possible to induce a person to give up some part of a commodity which he possesses in exchange for some amount of a second commodity. This theme of substitutability, which goes back at least to Edgeworth, finds a counterpart in modern treatments of choice theory in a "continuity" or "Archimedean" axiom. When combined with the other axioms of "rational" choice, this axiom implies the existence of a real-valued utility function with which to represent choice behavior. Starting with Debreu's [15] pioneering work, many rigorous treatments of the question of the existence of a real-valued utility function have been provided.

*The authors wish to acknowledge financial support from NSF Grant IRI-8608964, Harvard Business School Division of Research, and the Miller Institute for Basic Research in Science. We are grateful to Irving LaValle and Peter Fishburn for encouragement and helpful comments.

One strand in the development of choice theory has centered on the implications of weakening, or even entirely abandoning, the continuity axiom. It has been argued that unlike the other axioms of rational choice (axioms of completeness, transitivity, and reflexivity), the continuity axiom is of a somewhat unpleasant existential nature and has perhaps been assumed for reasons more of mathematical convenience than behavioral validity (Richter [46]). Areas in which non-Archimedean choice theory has proved insightful include insurance, portfolio selection, and resolutions of the Allais "paradox" (section 2), game theory (section 3), and maximin-like choice criteria to capture the notion of "ignorance" (section 4).

Chipman [9] states a very general result on the representation of non-Archimedean choice behavior: a complete, transitive, and reflexive preference relation on any set X can be represented by a lexicographic order* on (possibly transfinite) sequences of real numbers, where each element of X is associated with one such sequence. An alternative representation is possible using nonstandard analysis (Richter [46], Skala [51], Narens [45, Chs. 4–6]). In this survey, the lexicographic approach is emphasized, and henceforth the terms "non-Archimedean" and "lexicographic" will be used interchangeably.

This survey focuses on lexicographic choice under conditions of uncertainty. For an overview of non-Archimedean utility theory, which is less specialized than the one in this paper, the reader is referred to the comprehensive survey by Fishburn [23]. Section 2 of this paper first reviews the traditional von Neumann–Morgenstern expected utility theory, and then shows how a lexicographic representation arises if the usual Archimedean axiom is modified or dropped. Section 3 explains how the notion of lexicographic choice arises naturally in game theory. Section 4 is a brief survey of some lexicographic choice criteria which have been proposed for situations of "complete uncertainty", i.e. when the decision maker is "unable" to assess a subjective probability distribution.

This paper is only an overview of work on lexicographic choice; it does not pretend to describe the technical results in any detail. We have also borrowed from several papers, which will be mentioned, and have depended particularly on the work of Chipman and Fishburn.

2. Lexicographic von Neumann–Morgenstern expected utility

This section begins with a review of the usual expected utility theory. The exposition follows that of Fishburn [24]. The individual has a weak preference relation $\dot{\gtrsim}$ over the set P of probability distributions on a finite☆ set $C = \{c_1, \ldots, c_n\}$

*An example of a lexicographic order is the order of words in a dictionary. As Chipman [10] notes, the concept of a lexicographic order was implicit in Cantor's work. In economics, the idea of a lexicographic utility function was mentioned in von Neumann and Morgenstern ([54, p. 631]).

☆For ease of exposition, finiteness assumptions will often be made in this paper.

of consequences. Let $p(c)$ denote the probability that gamble $p \in P$ assigns to consequence $c \in C$. The set P is made into a mixture space by defining, for $p, q \in P$ and $0 \leqslant \alpha \leqslant 1$, $\alpha p + (1-\alpha)q$ to be the gamble which assigns probability $\alpha p(c) + (1-\alpha)q(c)$ to each $c \in C$.* Letting $\dot{>}$ and $\dot{\sim}$ denote strict preference and indifference, the following axioms comprise the von Neumann–Morgenstern expected utility theory.

A1. (Order axiom) $\dot{\gtrsim}$ is a complete, transitive, reflexive binary relation on P.

A2. (Independence axiom) For all $p, q, r \in P$ and $0 < \alpha \leqslant 1$, if $p \mathrel{\dot{>}}$ (resp. $\dot{\sim}$)q, then $\alpha p + (1-\alpha) r \mathrel{\dot{>}}$ (resp. $\dot{\sim}$) $\alpha q + (1-\alpha)r$.

A3. (Archimedean axiom) If $p \mathrel{\dot{\gtrsim}} q \mathrel{\dot{\gtrsim}} r$, then there exists $0 \leqslant \gamma \leqslant 1$ such that $\gamma p + (1-\gamma)r \mathrel{\dot{\sim}} q$.

A real-valued function u on P is said to be linear if $u(\alpha p + (1-\alpha)q) = \alpha u(p) + (1-\alpha)u(q)$ for all $p, q \in P$ and $0 \leqslant \alpha \leqslant 1$. The representation theorem can now be stated.

PROPOSITION 2.1

Axioms A1–A3 hold if and only if there is a linear function $u : P \to \mathbb{R}$ such that, for all $p, q \in P$, $p \mathrel{\dot{\gtrsim}} q \iff u(p) \geqslant u(q)$. Furthermore, u is unique up to positive affine transformations.

The uniqueness statement in proposition 2.1 means that a linear function $u' : P \to \mathbb{R}$ represents the same preferences as does u if and only if $u' = au + b$ for real numbers $a > 0$, b. To express u in the more familiar expected utility form, define u on C as well by $u(c) = u(p)$, where $p(c) = 1$. Then the linearity of u means that $u(p)$ can be written as

$$u(p) = \sum_{i=1}^{n} p(c_i)u(c_i) \quad \text{for all } p \in P.$$

If the Archimedean axiom (A3) is dropped, a numerical representation of preferences is no longer possible. Instead, to each gamble is assigned a vector of expected utilities, and these vectors are ordered using the lexicographic ordering $>_L$.☆ The vector of expected utilities is calculated by taking expectations under the

*Strictly speaking, this definition only establishes P as a mixture *set* (Fishburn [24, p. 11]). This is sufficient for the Archimedean theory, but for the non-Archimedean case, a further requirement is needed, namely, that P be a mixture *space*. (This terminology is from Hausner [29].) For a precise statement of the requirement, see Fishburn [24, p. 39, axiom M5]. All the mixture sets considered in this paper are also mixture spaces.

☆$a >_L b$ if and only if there is an i such that $a_i > b_i$ and $a_j = b_j$ for all $j < i$.

gamble of a lexicographic hierarchy of utility functions. A similar representation -- for the case of partially ordered preferences – can be found in Fishburn [24, p. 60, theorem 5.5].

PROPOSITION 2.2

Axioms A1 and A2 hold if and only if there are linear functions $u_1, \ldots, u_K$ on P, for some integer K, such that, for all $p, q \in P$,

$$p \dot{\gtrsim} q \iff [u_1(p), \ldots, u_K(p)] \geqslant_{\mathrm{L}} [u_1(q), \ldots, u_K(q)].$$

Without loss of generality, $K < n$. Furthermore, each u_k is unique up to affine transformations of $u_1, \ldots, u_k$ which assign positive weight to u_k.

The uniqueness part of the proposition means that utility functions $u'_1, \ldots, u'_K$ represent the same preferences as do $u_1, \ldots, u_K$ if and only if for each k,

$$u'_k = \sum_{j=1}^{k} a_{kj} u_j + b_k,$$

where $a_{k1}, \ldots, a_{kk}, b_k \in \mathbb{R}$ with $a_{kk} > 0$.

The reason why the lexicographic hierarchy of utility functions can be restricted to $n-1$ (or fewer) levels rests on the fact that in an $(n-1)$-dimensional space, namely P, at most $n-1$ vectors can be linearly independent. In fact, a complete proof of proposition 2.2 using linear algeabra can be based on the construction of Hausner [29]. We will provide a sketch of an alternative method of proof, based on a successive separation argument, which is perhaps more illuminating as to how the lexicographic representation arises. The same argument has been used by Fishburn [24, sect. 5.6], whose exposition is followed here, and Myerson [44, lemma 2], among others. A gamble $p \in P$ will be represented by the vector $[p(c_1), \ldots, p(c_{n-1})] \in \mathbb{R}^{n-1}$. Let H be the convex cone in $\mathbb{R}^{n-1}$ generated by $\{p - q : p, q \in P$ with $p \dot{>} q\}$. Clearly, $0 \notin H$. The crucial step is to show that there are vectors $v^1, \ldots, v^K$ in $\mathbb{R}^{n-1}$, where $K \leqslant n-1$, such that $(v^1 \cdot a, \ldots, v^K \cdot a) >_{\mathrm{L}} 0$ for $a \in H$. The ith component of v^k is to be interpreted as the kth order utility of the lottery that gives consequence c_i for sure. This follows by successive separation. The separating hyperplane theorem implies that there is a $v^1 \in \mathbb{R}^{n-1}$ such that $v^1 \cdot a \geqslant 0$ for all $a \in H$, with strict inequality for some $a \in H$. Let $H^1 = H \cap \{a \in \mathbb{R}^{n-1} : v^1 \cdot a = 0\}$. The set H^1 is again a convex cone in $\mathbb{R}^{n-1}$ with $0 \notin H^1$, so there is a $v^2 \in \mathbb{R}^{n-1}$ such that $v^2 \cdot a \geqslant 0$ for all $a \in H^1$, with strict inequality for some $a \in H^1$. Let $H^2 = H^1 \cap \{a \in \mathbb{R}^{n-1} : v^2 \cdot a = 0\}$. Continuing in this fashion establishes that there are vectors $v^1, \ldots, v^K$, where $K \leqslant n-1$ since the dimension of the H^k's is strictly decreasing, such that $(v^1 \cdot a, \ldots, v^K \cdot a) >_{\mathrm{L}} 0$ for $a \in H$. It only remains to define the von Neumann–Morgenstern utility function u_k, for $k = 1, \ldots, K$, by

$$u_k(p) = \sum_{i=1}^{n-1} v_i^k p(c_i).$$

If the space of consequences C is infinite, then a lexicographic representation along the lines of that of proposition 2.2 is still possible. One route, which is taken in Fishburn [22], [24, p. 39, theorem 4.4], is to adopt an extra axiom which has the effect of restricting the lexicographic hierarchy to finitely many levels. Without this additional axiom, an infinite-dimensional representation may be necessary (and can be based on the arguments of Hausner and Wendel [30]).

Several authors have discussed applications of lexicographic expected utility. Chipman [11] considers the implications for portfolio choice of dropping the Archimedean axiom. He shows how "ruin" (e.g. bankruptcy) and "aspiration levels" can be modeled. Suppose, for example, that the consequences $c_1, \ldots, c_n$ represent possible wealth levels, where c_1 is negative (i.e. "ruin") and $c_2, \ldots, c_n$ are positive. A possible two-level lexicographic representation is:

$$u_1(c_1) = 0$$

$$u_1(c_i) = 1 \quad \text{for } i = 2, \ldots, n;$$

$$u_2(c_1) = -1$$

$$u_2(c_i) = i \quad \text{for } i = 2, \ldots, n.$$

Gamble p will be preferred to gamble q if and only if

$$\left[\sum_{i=2}^{n} p(c_i),\ -p(c_1) + \sum_{i=2}^{n} p(c_i)i\right] >_L \left[\sum_{i=2}^{n} q(c_i),\ -q(c_1) + \sum_{i=2}^{n} q(c_i)i\right].$$

In other words, u_1, u_2 represent a policy of minimizing the probability of ruin and, among gambles with equal probability of ruin, maximizing expected wealth. A similar two-level lexicographic representation can be used to model the policy of maximizing the probability of achieving some aspiration wealth level w and, among gambles with equal probability of achieving w, maximizing expected utility. Chipman [11] offers this second type of behavior as an alternative resolution to that provided by Friedman and Savage [26] of the "paradox" of simultaneous gambling and insurance purchases. More generally, any preference relation defined in terms of whether a constraint is or is not met can be represented lexicographically. Ençarnacion [18,19], Ferguson [20,21], and Thrall [53] discuss other instances of this phenomenon.

A slightly different form of lexicographic choice theory, based on the expected utility framework reviewed at the beginning of this section, is developed in Gilboa [27].

Rather than simply dropping the Archimedean axiom A3, Gilboa proposes a different set of axioms which includes a modification of A2 as well as of A3. These axioms characterize a lexicographic decision rule – the first component being maximin, the second a combination of maximin and expected utility. Gilboa offers his approach as an alternative resolution of the Allais "paradox" (Allais [1]) to other resolutions which weaken or drop the von Neumann–Morgenstern independence axiom (see Machina [36], Chew [8], Fishburn [25], Dekel [16], Yaari [55], among others).

3. Lexicographic choice in games

This section will trace the development of a lexicographic approach to refinements of Nash equilibrium. The program of refining the set of Nash equilibria in the context of extensive-form games was initiated by Selten [49]. He demonstrated that a Nash equilibrium may be supported by off-the-equilibrium-path behavior which is implausible, and proposed instead the concept of subgame-perfect equilibrium as a way to eliminate such implausible equilibria. A Nash equilibrium is subgame perfect if its component strategies induce a Nash equilibrium in every subtree of the original tree. Subgame perfection is in a sense a multi-person version of the backwards induction principle of single-person decision theory.

Kreps and Wilson [34] proposed the idea of a sequential equilibrium as a refinement of subgame perfection. The definition of a sequential equilibrium has two parts. The first part is "sequential rationality": for each information set I in the tree, the strategy of the player who moves at I must prescribe a choice which is optimal given some probability distribution over the nodes in I (and given the strategies of the other players). The key aspect of this definition is the specification of beliefs at *every* information set in the tree – even at those which have zero "prior" probability of being reached under the equilibrium strategies. The second part of the definition of a sequential equilibrium is a consistency requirement on how the players' strategies and their beliefs at the various information sets are related.

Kreps and Wilson give the term "lexicographic consistency" to one of the consistency requirements which they impose on beliefs.* Roughly speaking, lexicographic consistency demands that there be a sequence $\sigma_1, \ldots, \sigma_K$ of alternative strategy profiles (representing alternative beliefs as to how the game is played), and that this sequence be used in lexicographic fashion to assign beliefs to information sets in the tree. That is, the first ("base-case") strategy profile σ_1 is used to calculate beliefs at all information sets to which σ_1 assigns positive probability. Next, σ_2 is used to calculate beliefs at as many remaining information sets as possible. And so on. The notion of lexicographic consistency indicates an appealing way to analyze refinements

*See Kreps and Wilson [34, p.874]. Actually, there are some subtleties associated with their definition of lexicographic consistency – see Kreps and Ramey [35].

of Nash equilibrium: a player can be thought of as having a *lexicographic hierarchy* of beliefs, and uses this hierarchy in some fashion to compute an optimal strategy. The remainder of this section of the paper will be devoted to reviewing some recent work which makes precise this perspective on refinements, and gives it an axiomatic foundation.

Myerson [42,43] approached this question by providing a formal definition of a system of beliefs, which he called a *conditional probability system* (CPS), which includes beliefs conditional on events of prior probability 0.

DEFINITION 3.1

A conditional probability system (CPS) on a finite space Ω is a collection of functions $p(\cdot \mid S)$, one for each $\emptyset \neq S \subset \Omega$, such that:

(1) $p(\cdot \mid S)$ is a probability measure on Ω with $p(S \mid S) = 1$;

(2) if $V \subset T \subset S$ and $T \neq \emptyset$, then $p(V \mid S) = p(V \mid T) \times p(T \mid S)$.

Myerson [43] contains an axiomatic justification for CPSs. In order to describe this work, and to discuss subsequent developments, it will be helpful to first review the conventional theory which serves as the starting point.

An appropriate theory for analyzing the behavior of players in a game is *subjective* expected utility theory (SEU). The idea is that each player faces a space of possible states of the world, namely, the possible choices that the other players can make, and chooses a strategy to maximize his expected utility calculated according to his subjective probability distribution on the states. The foundational work on SEU is, of course, that of Savage [48]. However, for the purpose of analyzing finite games, the version due to Anscombe and Aumann [2] is more convenient; it is this theory which we now review.

As in sect. 2, let P be the set of probability distributions on a finite set C of consequences. The individual has a weak preference relation $\dot{\gtrsim}$ on P^{Ω}, where Ω is a finite set of states. If act $x \in P^{\Omega}$ is chosen, and state ω occurs, then the consequence is the gamble $x_\omega \in P$, where x_ω is the ωth component of x. The set P^{Ω} is made into a mixture space by defining, for $x, y \in P^{\Omega}$ and $0 \leqslant \alpha \leqslant 1$, $\alpha x + (1-\alpha)y$ to be the act which, if state ω occurs, assigns probability $\alpha x_\omega(c) + (1-\alpha)y_\omega(c)$ to each $c \in C$. Given an event $S \subset \Omega$, let x_S be the tuple $(x_\omega)_{\omega \in S}$ and $x_{-S} \equiv x_{\Omega - S}$. Letting $\dot{>}$ and $\dot{\sim}$ denote strict preference and indifference, the following axioms comprise the Anscombe and Aumann theory.

B1. (Order axiom) $\dot{\gtrsim}$ is a complete, transitive, reflexive binary relation on P^{Ω}.

B2. (Independence axiom) For all $x, y, z \in P^{\Omega}$ and $0 < \alpha \leqslant 1$, if $x \dot{>}$ (resp. $\dot{\sim}$) y, then $\alpha x + (1-\alpha)z \dot{>}$ (resp. $\dot{\sim}$) $\alpha y + (1-\alpha)z$.

B3. (Non-triviality axiom) There are $x, y \in P^{\Omega}$ such that $x \dot{>} y$.

B4. (Archimedean axiom) If $x \dot{\gtrsim} y \dot{\gtrsim} z$, then there exists $0 \leqslant \gamma \leqslant 1$ such that $\gamma x + (1 - \gamma) z \dot{\sim} y$.

We next define conditional preferences (Savage [48]) and null events.

DEFINITION 3.2

$x \dot{\gtrsim}_S y$ if for some $z \in P^{\Omega}$, $(x_S, z_{-S}) \dot{\gtrsim} (y_S, z_{-S})$.*

DEFINITION 3.3

The event $S \subset \Omega$ is null if $x \dot{\sim}_S y$ for all $x, y \in P^{\Omega}$.

B5. (State-independence axiom) For all states $\omega, \omega' \in \Omega$ which are non-null, and for any two acts $x, y \in P^{\Omega}$ with $x_\omega = x_{\omega'}$, $y_\omega = y_{\omega'}$, $x \dot{\gtrsim}_{\{\omega\}} y$ if and only if $x \dot{\gtrsim}_{\{\omega'\}} y$.

PROPOSITION 3.1 (Anscombe and Aumann [2] and Fishburn [24, p. 111, theorem 9.2])

Axioms B1–B5 hold if and only if there is a linear function $u : P \to I\!R$ and a probability measure p on Ω such that, for all $x, y \in P^{\Omega}$,

$$x \dot{\gtrsim} y \iff \sum_{\omega \in \Omega} p(\omega) u(x_\omega) \geqslant \sum_{\omega \in \Omega} p(\omega) u(y_\omega).$$

Furthermore, u is unique up to positive affine transformations, p is unique, and $p(S) = 0$ if and only if the event S is null.

Returning to the theme of this section, Myerson [43] augments the conventional theory by supposing that the decision maker has, in addition to the preference relation $\dot{\gtrsim}$ on P^{Ω}, a preference relation $\dot{\gtrsim}_S$ on P^S for each non-empty subset S of Ω. It is important to note that these additional preferences $\dot{\gtrsim}_S$ with which Myerson endows the decision maker are *not* the conditional preferences of definition 3.2, but are, a priori, quite distinct preference relations. Myerson supposes that each preference $\dot{\gtrsim}_S$ satisfies the analogs to axioms B1–B5 on the space P^S, and hence each $\dot{\gtrsim}_S$ can be represented by a utility function $u_S : P \to I\!R$ and a probability measure $p(\cdot \mid S)$ on S. Myerson then shows that, by specifying further axioms on how the various $\dot{\gtrsim}_S$ are related, one can conclude that there is a single utility function ($u_S = u$ for all S) and that the collection of probability measures $\{p(\cdot \mid S) : \emptyset \neq S \subset \Omega\}$ is a CPS.

*The relation $\dot{\gtrsim}_S$ is well defined since by axioms B1 and B2 the definition is independent of z.

Notice that at no point has the Archimedean axiom B4 been weakened, so that the representation which emerges cannot be lexicographic in nature. Nevertheless, a lexicographic character can be introduced by specifying *above and beyond* the preference relations $\dot{\gtrsim}_S$ a *rule* for how the representation should be applied. Consider the following example. Suppose $p(\cdot \mid \Omega)$ has support S, where S is a strict subset of Ω, and $p(\cdot \mid \Omega - S)$ has support $\Omega - S$. Let X be a non-empty, finite subset of acts from P^{Ω} that is closed under subjective mixtures with respect to the partition $\{S, \Omega - S\}$.* According to the preference relation $\dot{\gtrsim}$, an act $x \in X$ is optimal if and only if it yields at least as much expected utility, calculated under the probability measure $p(\cdot \mid \Omega)$ (which is the same as the measure $p(\cdot \mid S)$), as does any other act $y \in X$. As an alternative, one could bring the preference relation $\dot{\gtrsim}_{\Omega - S}$ into the picture as well, by specifying the following rule for optimality: $x \in X$ is optimal if and only if:

$$\sum_{\omega \in S} p(\omega \mid S)\,[u(x_\omega) - u(y_\omega)] \geqslant 0$$

and

$$\sum_{\omega \in \Omega - S} p(\omega \mid \Omega - S)\,[u(x_\omega) - u(y_\omega)] \geqslant 0$$

for all $y \in X$. This rule for applying CPSs faithfully captures the two key features of sequential equilibrium discussed earlier. Sequential rationality is captured by requiring that x be optimal not only conditional on S, but also conditional on $\Omega - S$, where the latter has "prior" probability 0. Lexicographic consistency is captured in the way that first $p(\cdot \mid S)$ and then $p(\cdot \mid \Omega - S)$ are used to assign beliefs. The paper by McLennan [38] makes precise how CPSs can be used to characterize sequential equilibrium, and also provides an existence proof using CPSs.

A thoroughgoing axiomatic approach would *derive* the lexicographic ranking of acts implicit in the optimality rule just described from the decision maker's basic preference relation $\dot{\gtrsim}$. Such an approach would also be more parsimonious: the preferences $\dot{\gtrsim}_S$, for $\emptyset \neq S \subset \Omega$, could be deduced from – rather than postulated in addition to – the basic preference relation $\dot{\gtrsim}$.☆ An explicitly lexicographic approach will now be described.

In the von Neumann–Morgenstern framework of sect. 2, the lexicographic expected utility representation consisted of a hierarchy of utility functions (proposition 2.2). The representation which is appropriate for the objectives discussed in this section consists of a single utility function $u: P \to \mathbb{R}$ but a lexicographic hierarchy of probability measures.

*That is, if $x, y \in X$, then $(x_S, y_{\Omega - S}) \in X$.

☆In fact, $\dot{\gtrsim}_S$ will turn out to be precisely the conditional preference relation given S.

DEFINITION 3.4

A lexicographic probability system (LPS) on a finite space Ω is a K-tuple $\lambda = (p_1, \ldots, p_K)$, for some integer K, of probability distributions on Ω such that for every $\omega \in \Omega$, $p_k(\omega) > 0$ for some k.[☆]

With this representation, an act x will be preferred to an act y, according to the basic preference relation $\dot{\gtrsim}$, if and only if

$$\left[\sum_{\omega \in \Omega} p_k(\omega) u(x_\omega)\right]_{k=1}^{K} \geqslant_L \left[\sum_{\omega \in \Omega} p_k(\omega) u(y_\omega)\right]_{k=1}^{K} \qquad (\star).$$

To see how this representation embodies the optimality rule described above, consider a two-level LPS $\hat{\lambda} = (p_1, p_2)$, where $p_1(\cdot) = p(\cdot \mid S)$ and $p_2(\cdot) = p(\cdot \mid \Omega - S)$. Then, $x \dot{\gtrsim} y$ if and only if

$$\left[\sum_{\omega \in S} p_1(\omega)\,[u(x_\omega) - u(y_\omega)],\ \sum_{\omega \in \Omega - S} p_2(\omega)\,[u(x_\omega) - u(y_\omega)]\right]$$

lexicographically exceeds $(0, 0)$.

The representation $(\star)$ is a natural extension of the lexicographic expected utility representation of proposition 2.2 to the richer environment of SEU. We now turn to the relationship between CPSs and LPSs.

DEFINITION 3.5

A lexicographic conditional probability system (LCPS) is an LPS $(p_1, \ldots, p_K)$ in which the supports of the p_k's are disjoint and for every $\omega \in \Omega$, $p_k(\omega) > 0$ for some k.

An example of an LCPS is the LPS $\hat{\lambda}$ constructed above. Although we have seen that the preference structures underlying CPSs and LCPSs are different, the second being genuinely lexicographic while the first is not, there is nevertheless a natural equivalence between the probabilities in the two structures. More formally, the spaces of CPSs and LCPSs are isomorphic, as we now demonstrate. Given a CPS $\{p(\cdot \mid S) : \emptyset \neq S \subset \Omega\}$, an LCPS $\lambda = (p_1, \ldots, p_K)$ can be defined as follows (here, Supp denotes the support of a measure):

☆This definition of an LPS rules out null events (definition 3.3). The more general case in which null events are permitted is treated in Blume et al. [7].

$$p_1 = p(\cdot \mid \Omega)$$

$$p_2 = p(\cdot \mid N_1), \quad \text{where } N_1 = \Omega - \operatorname{Supp} p(\cdot \mid \Omega)$$

$$p_3 = p(\cdot \mid N_2), \quad \text{where } N_2 = N_1 - \operatorname{Supp} p(\cdot \mid N_1)$$

The sequence of p_k's will terminate at some p_K such that Supp $p_K = N_{K-1}$. The number of levels K in the LCPS λ is finite since the N_k's are strictly decreasing and Ω is finite. Conversely, given an LCPS $\lambda = (p_1, \ldots, p_K)$, a CPS $\{p(\cdot \mid S) : \emptyset \neq S \subset \Omega\}$ can be defined as follows. For $\emptyset \neq S \subset \Omega$, let $k(S) = \min\{k : p_k(S) > 0\}$ and for each $\omega \in \Omega$ define $p(\omega \mid S) = p_{k(S)}(\omega)/p_{k(S)}(S)$. The probability distribution $p(\cdot \mid S)$ defined in this way is readily seen to satisfy conditions (1) and (2) of definition 3.1.

The more general space of LPSs is equivalent to the set of all convex combinations, in an appropriate sense, of LCPSs. Given two LCPSs $\lambda = (p_1, \ldots, p_K)$ and $\mu = (q_1, \ldots, q_K)$, both of length K, and $0 \leqslant \alpha \leqslant 1$, define $\alpha\lambda + (1-\alpha)\mu = [\alpha p_1 + (1-\alpha)q_1, \ldots, \alpha p_K + (1-\alpha)q_K]$.* Clearly, according to this definition the convex combination of two LCPSs need not be an LCPS, but it *will* be an LPS.

It is the larger space of LPSs that turns out to be the one suitable for analyzing refinements of Nash equilibrium, such as (normal-form) perfect equilibrium (Selten [50]) and proper equilibrium (Myerson [41]) which are defined on the *normal*, as opposed to the extensive, form. In fact, the following three issues appear related: (1) the property that convex combinations of LCPSs may be LPSs; (2) the appropriateness of LCPSs for extensive-form refinements and of LPSs for normal-form refinements; and (3) questions of normal-form equivalence (see Dalkey [14], Elmes and Remy [17], Kohlberg and Mertens [32], Mertens [39], and Thompson [52]). We believe that the precise relationship between (1), (2), and (3) merits further research.

An axiomatic characterization of subjective expected utility with lexicographic hierarchies of beliefs, i.e. the representation (★) above, has been developed by Blume et al. [7]. These authors go on to show how, in the context of single-person decision theory, this lexicographic variant of SEU provides a synthesis of both admissibility (Luce and Raiffa [35, Ch. 13]) and backwards induction with expected utility. They also show how SEU with lexicographic beliefs can be used to characterize the normal-form refinements: (normal-form) perfect equilibrium (Selten [50]) and proper equilibrium (Myerson [41]). The topological properties of lexicographic representations have been investigated by McLennan [38] and Hammond [28]. For the case of CPSs, McLennan has demonstrated a useful topology in terms of extended log-likelihood

*This definition may require further refinement if the p_k's and q_k's are not of comparable "order". This issue will not be pursued further here.

ratios. Hammond [28] looks at the topological properties of – and the relationships between – the various lexicographic belief systems discussed in this section.

We conclude this section by mentioning the use of a lexicographic decision criterion in the context of repeated games. Rubinstein [47] proves Folk theorems for the case in which the players evaluate streams of payoffs using the so-called "overtaking" criterion.* As Rubinstein points out, the overtaking criterion does not have a real-valued representation; it is in fact lexicographic in nature.

4. Lexicographic decision rules under "complete uncertainty"

This section is devoted to a brief description of some lexicographic choice criteria which arise in the context of decision making under conditions of so-called "complete uncertainty". The starting point for analysis is similar to that for SEU described in sect. 3. The decision maker is faced with a number of alternative possible actions to take. The consequence of each action depends on the state of the world – which is unknown to the decision maker at the time at which an action must be chosen. The distinguishing feature of "complete uncertainty" is the assumption that the decision maker has no way of assigning subjective probabilities to states. In this respect, decision making under "complete uncertainty" has sometimes been claimed to be more faithful to the conceptions of uncertainty held by Keynes, Knight, and others, than is the subjective probability approach.☆

In the literature on choice under "complete uncertainty", it is customary to suppose that the decision maker has a preference relation on possible consequences and from this to infer, by postulating various axioms, preferences over actions. There are two strands in the literature, distinguished according to the nature of assumptions made about preferences over consequences. The first strand, which includes papers by Milnor [40], Arrow and Hurwicz [3], Maskin [37, sect. I and II], Cohen and Jaffray [12,13], and Barbera and Jackson [4], assumes a cardinal ranking of consequences. Since cardinal utility is usually derived from a set of probabilistic axioms, it is perhaps a strange component of a theory that eschews probabilities. The second strand of the literature supposes only an ordinal ranking of consequences; this is the route taken by Maskin [37, sect. III], Barrett and Pattanaik [5], and Kelsey [31]. Although the ordinal approach is the more attractive, for expositional ease we shall nevertheless think of an action x as a vector of real numbers, the ith component x_i representing the von Neumann–Morgenstern utility if state i occurs.

* A sequence $\{x_t\}_{t=1}^{\infty}$ is (strictly) preferred to a sequence $\{y_t\}_{t=1}^{\infty}$ according to the overtaking criterion if $\lim_T \inf \Sigma_{t=1}^{T}(x_t - y_t) > 0$.

☆ Recently, Bewley [6], following Aumann [4], has developed a variant of SEU, which he terms "Knightian", in which the decision maker has a (possibly large) convex set of subjective probability distributions.

The classical criterion for ranking acts under conditions of "complete uncertainty" is, of course, the *maximin* criterion (see e.g. Luce and Raiffa [35, Ch. 13]). Given two acts $x, y \in \mathbb{R}^n$, x is (strictly) preferred to y according to the maximin criterion if and only if

$$\min\{x_i : i = 1, \ldots, n\} > \min\{y_i : i = 1, \ldots, n\}.$$

The maximin criterion plays a central role in the theory of two-person zero-sum games. Nevertheless, maximin is a rather "weak" decision rule in that it makes no use of the characteristics of an act beyond the minimum utility that the act guarantees. One way to take additional information about acts into account, while remaining faithful to the spirit of maximin, is to use a *lexicographic* criterion: Acts are first compared on the basis of their minimum utility levels. If deemed indifferent on this basis, then information about their next levels of utility is taken into account. And so on. Two such criteria which have been proposed in the literature are the *protective* criterion and the *leximin* criterion.

To define the protective and leximin criteria, some extra notation will be useful. Given two acts $x, y \in \mathbb{R}^n$, let $I(x, y) = \{i : x_i \neq y_i\}$. The protective criterion states that act x is (strictly) preferred to act y if and only if

$$\min\{x_i : i \in I(x, y)\} > \min\{y_i : i \in I(x, y)\}.$$

This definition is best understood by means of a simple example. Suppose $x = (30, 60, 10, 50)$ and $y = (30, 40, 10, 70)$. Then x and y are deemed indifferent under maximin. Under the protective criterion, neither the lowest utility level of 10 nor the next lowest of 30 permits a distinction between x and y, but once attention is focused on the utilities in the remaining states, the guaranteed level under x exceeds that under y. Consider now a second example: $x = (30, 60, 10, 50)$ and $y = (30, 10, 70, 10)$. Here, x and y are ranked indifferent under the protective criterion. The leximin criterion works by counting the number of times the lowest utility level is realized, then if necessary, the number of times the next lowest is attained. And so on. Given an act $x \in \mathbb{R}^n$ and a number $a \in \mathbb{R}$, let $C(x, a)$ denote the cardinality of the set $\{i : x_i = a\}$. Then, formally speaking, the leximin criterion states that act $x \in \mathbb{R}^n$ is (strictly) preferred to act $y \in \mathbb{R}^n$ if and only if there is a number a such that

$$C(x, b) = C(y, b) \quad \text{for all } b < a$$

$$C(x, a) < C(y, a).$$

In the second example, x is preferred to y under the leximin criterion.

An ordinal characterization of the protective criterion has been given by Barrett and Pattanaik [5]. Barbera and Jackson [4] provide a cardinal axiomatization.

Maskin [37] gives both ordinal and cardinal characterizations of the leximin criterion, while again Barrett and Pattanaik, and Barbera and Jackson, provide ordinal and cardinal characterizations, respectively. The reader is referred to the paper by Barrett and Pattanaik for a comparison of the various different axiom systems.

Kelsey [31] considers the case of "partial uncertainty", in which the decision maker still does not possess a subjective probability distribution on states, but *is* able to rank the states in order of likelihood. A particular criterion which Kelsey characterizes is the *lex-likelihood* rule: Given two acts x, y, the decision maker compares their utilities in the most likely state. If the utility of x is greater than that of y, then x is preferred. If the utilities are the same, then the respective utilities in the second most likely state are compared. And so on. There is an obvious formal similarity between this criterion and a special case of the lexicographic behavior discussed in sect. 3, in which choices are made on the basis of expected utility with an LCPS $(p_1, \ldots, p_K)$ in which each p_k is concentrated on a single state.

References

[1] M. Allais, Le comportement de l'homme rationnel devant le risque: Critique des postulats et axioms de l'ecole Americaine, Econometrica 21(1953)503.

[2] F. Anscombe and R. Aumann, A definition of subjective probability, Ann. Math. Stat. 34(1963)199.

[3] K. Arrow and L. Hurwicz, An optimality criterion for decision-making under ignorance, in: *Uncertainty and Expectations in Economics*, ed. C. Carter and J. Ford (Augustus M. Kelley, New Jersey, 1972).

R. Aumann, Utility theory without the completeness axiom, Econometrica 30(1962)445.

[4] S. Barbera and M. Jackson, Maximin, leximin, and the protective criterion: Characterizations and comparisons, J. Economic Theory, forthcoming.

[5] R. Barrett and P. Pattanaik, Decision-making under complete uncertainty, Research Memorandum 1/86, Department of Economics, University of Birmingham (1986).

[6] T. Bewley, Knightian decision theory: parts I and II, Cowles Foundation Discussion Paper Nos. 807 and 835, Cowles Foundation, Yale University (1986 – 87).

[7] L. Blume, A. Brandenburger and E. Dekel, Lexicographic probabilities and equilibrium refinements, I and II (1988), in preparation.

[8] S. Chew, A generalization of the quasi-linear mean with applications to the measurement of income inequality and decision theory resolving the Allais paradox, Econometrica 57 (1983)1065.

[9] J. Chipman, The foundations of utility, Econometrica 28(1960)193.

[10] J. Chipman, On the lexicographic representation of preference orderings, in: *Preference, Utility, and Demand*, ed. J. Chipman, L. Hurwicz, M. Richter and H. Sonnenschein (Harcourt Brace Jovanovich, New York, 1971).

[11] J. Chipman, Non-Archimedean behavior under risk: An elementary analysis – with application to the theory of assets, in: *Preference, Utility, and Demand*, ed. J. Chipman, L. Hurwicz, M. Richter and H. Sonnenschein (Harcourt Brace Jovanovich, New York, 1971).

[12] M. Cohen and J.-Y. Jaffray, Rational behavior under ignorance, Econometric 48(1980)1281.

[13] M. Cohen and J.-Y. Jaffray, Approximations of rational criteria under complete ignorance and the independence axiom, Theory and Decision 15(1983)121.

[14] N. Dalkey, Equivalence of information patterns and essentially determinate games, in: *Contributions to the Theory of Games*, II, ed. H. Kuhn and A. Tucker, Annals of Mathematical Studies 28 (Princeton University Press, Princeton, 1953).
[15] G. Debreu, Representation of a preference ordering by a numerical function, in: *Decision Processes*, ed. R. Thrall, C. Coombs and R. Davis (Wiley, New York, 1954).
[16] E. Dekel, An axiomatic characterization of preferences under uncertainty: Weakening the independence axiom, J. Economic Theory 40(1986)304.
[17] S. Elmes and P. Reny, The equivalence of games with perfect recall, Department of Economics, Princeton University (1987), unpublished.
[18] J. Encarnación, Constraints and the firm's utility function, Review of Economic Studies 31(1964)113.
[19] J. Encarnacion, On decisions under uncertainty, Economic Journal 75(1965)442.
[20] C. Ferguson, An essay on cardinal utility, Southern Economic Journal 25(1958)11.
[21] C. Ferguson, The theory of multidimensional utility analysis in relation to multiple-goal business behavior: A synthesis, Southern Economic Journal 32(1965)169.
[22] P. Fishburn, A study of lexicographic expected utility, Manag. Sci. 17(1971)672.
[23] P. Fishburn, Lexicographic orders, utilities, and decision rules: A survey, Manag. Sci. 20 (1974)1442.
[24] P. Fishburn, *The Foundations of Expected Utility* (Reidel, Dordrecht, 1982).
[25] P. Fishburn, Transitive measurable utility, J. Economic Theory 31(1983)293.
[26] M. Friedman and L. Savage, The utility analysis of choices involving risk, J. Political Economy 56(1948)279.
[27] I. Gilboa, A combination of expected utility and maximin decision criteria, Working Paper No. 12-86, Foerder Institute for Economic Research, Tel-Aviv University (1986).
[28] P. Hammond, Extended probabilities for decision theory and games, Department of Economics, Stanford University (1987), unpublished.
[29] M. Hausner, Multidimensional utilities, in: *Decision Processes*, ed. R. Thrall, C. Coombs and R. Davis (Wiley, New York, 1954).
[30] M. Hausner and J. Wendel, Ordered vector spaces, Proc. Amer. Math. Soc. 3(1952)977.
[31] D. Kelsey, Decision making under partial uncertainty, Churchill College, Cambridge (1988), unpublished.
[32] E. Kohlberg and J.-F. Mertens, On the strategic stability of equilibria, Econometrica 54 (1986)1003.
[33] D. Kreps and G. Ramey, Structural consistency, consistency, and sequential rationality, Econometrica 55(1987)1331.
[34] D. Kreps and R. Wilson, Sequential equilibria, Econometrica 50(1982)863.
[35] R. Luce and H. Raiffa, *Games and Decisions* (Wiley, New York, 1957).
[36] M. Machina, Expected utility analysis without the independence axiom, Econometrica 50(1982)277.
[37] E. Maskin, Decision-making under ignorance with implications for social choice, Theory and Decision 11(1979)319.
[38] A. McLennan, Consistent conditional systems in noncooperative game theory, Discussion Paper 5217-86, Mathematical Sciences Research Institute, University of California at Berkeley (1986).
[39] J.-F. Mertens, Ordinality in non cooperative games, CORE Discussion Paper 8728, Université Catholique de Louvain (1987).
[40] J. Milnor, Games against nature, in: *Decision Processes*, ed. R. Thrall, C. Coombs and R. Davis (Wiley, New York, 1954).
[41] R. Myerson, Refinements of the Nash equilibrium concept, Int. J. Game Theory 7(1978)73.
[42] R. Myerson, Multistage games with communication, Econometrica 54(1986)323.

[43] R. Myerson, Axiomatic foundations of Bayesian decison theory, Discussion Paper 671, J.L. Kellogg Graduate School of Management, Northwestern University (1986).
[44] R. Myerson, Acceptable and predominant correlated equilibria, Int. J. Game Theory 15 (1986).
[45] L. Narens, *Abstract Measurement Theory* (MIT Press, Cambridge, MA, 1985).
[46] M. Richter, Rational choice, in: *Preference, Utility, and Demand,* ed. J. Chipman, L. Hurwicz, M. Richter and H. Sonnenschein (Harcourt Brace Jovanovich, New York, 1971).
[47] A. Rubinstein, Equilibrium in supergames with the overtaking criterion, J. Economic Theory 21(1979)21.
[48] L. Savage, *The Foundations of Statistics* (Wiley, New York, 1954).
[49] R. Selten, Spieltheoretische behandlung eines oligopolmodels mit nachfragetragheit, Zeitschrift für die Gesamte Staatswissenschaft 121(1965)401; 667.
[50] R. Selten, Reexamination of the perfectness concept for equilibrium points in extensive games, Int. J. Game Theory 4(1975)25.
[51] H. Skala, *Non-Archimedean Utility Theory* (Reidel, Dordrecht, 1975).
[52] F. Thompson, Equivalence of games in extensive form, Research Memorandum RM-759, The RAND Corporation, Santa Monica (1952).
[53] R. Thrall, Applications of multidimensional utility theory, in: *Decision Processes,* ed. R. Thrall, C. Coombs and R. Davis (Wiley, New York, 1954).
[54] J. von Neumann and O. Morgenstern, *Theory of Games and Economic Behavior,* 2nd ed. (Princeton University Press, Princeton, 1947).
[55] M. Yaari, The dual theory of choice under risk, Econometrica 55(1987)95.