PETER C. FISHBURN

SUBJECTIVE EXPECTED UTILITY: A REVIEW OF NORMATIVE THEORIES

Abstract: This paper reviews theories of subjective expected utility for decision making under uncertainty. It focuses on normative interpretations and discusses the primitives, axioms and representation-uniqueness theorems for a number of theories. Similarities and differences among the various theories are highlighted. The interplay between realistic decision structures and structural axioms that facilitate mathematical derivations is also emphasized.

The review attempts to be complete up to 1980. Among others, it includes theories developed by Ramsey; Savage; Suppes; Davidson and Suppes; Anscombe and Aumann; Pratt, Raiffa and Schlaifer; Fishburn; Bolker; Jeffrey; Pfanzagl; Luce and Krantz.

1. INTRODUCTION

This paper reviews theories for decision making under uncertainty that represent numerically the beliefs and preferences of a presumably rational individual in a personalistic, or subjective, expected utility model. Although there is abundant evidence[1] that people's casual or carefully considered decisions often violate the assumptions of subjective expected utility theories, these theories are felt by many writers to provide the most satisfactory normative approach to decision making under uncertainty. The present survey will focus on the normative approach and describe in modest detail the primitives, axioms, and numerical representations of a variety of theories that have been proposed since 1926, when Frank P. Ramsey set down his ideas for measuring beliefs on the basis on the extents to which we are prepared to act on them.

Three preliminary sections will preface our more detailed review of particular theories. The first of these identifies the primitives and constructed notions that are used in most of the theories. We then outline some of the numerical representations and types of axioms that will be encountered later. The final preliminary section summarizes the linear utility theory of John von Neumann and Oskar Morgenstern, which first appeared in 1944. Although the von Neumann-Morgenstern theory bypasses the issue of subjective (personal, psychological) probability, it is used in some subjective expected utility

Theory and Decision 13 (1981) 139–199. 0040–5833/81/0132–0139$06.10.
Copyright © 1981 *by D. Reidel Publishing Co., Dordrecht, Holland, and Boston, U.S.A.*

theories. It appears in the representation derivation in Leonard J. Savage's theory (1954) which, to the best of my knowledge, was the first – and, by some accounts, still the finest – complete theory of subjective expected utility. Savage also draws on the pioneering ideas of Bruno de Finetti in subjective probability. Direct uses of von Neumann-Morgenstern lotteries in the axioms of other theories will be discussed in section 8.

Particular theories of subjective expected utility are examined in sections 5 through 10. A partial sectional outline follows.

(5) Ramsey's (1931) ethically neutral proposition and his proposal for measuring utilities and subjective probabilities. Theories for equally-likely events. Pfanzagl's (1967, 1968) completion of Ramsey's outline.

(6) Savage's (1954) act-oriented theory, constant acts, and continuously divisible events. Other Savage-type theories.

(7) Suppes' (1956) merger of Ramsey and Savage with acts and an ethically neutral proposition.

(8) Extraneous scaling probabilities and the use of probability lotteries in subjective expected utility. Theories by Anscombe and Aumann (1963), Pratt, Raiffa and Schlaifer (1964, 1965), Fishburn (1967, 1969) and others.

(9) Luce and Krantz's (1971) conditional acts on nonnull events, and a theory of conditional subjective expected utility.

(10) Jeffrey's (1965a, 1978) and Bolker's (1967) mono-set theories, with utilities and subjective probabilities defined on the same entities. Domotor's (1978) finite version.

The final section of the paper presents an evaluative summary of the theories reviewed.

2. PRIMITIVES AND DEFINITIONS

The primitive notions of the theories that we shall review consist of one or more nonempty sets (or algebras) and one or more preference relations on these sets or on sets that are constructed from the basic sets. In this section we shall assign symbols to and provide interpretations of the basic and constructed sets that will be used later. Special sets not mentioned here will be introduced as needed.

2.1. *Consequences*

Most theories employ a set $\mathscr{C}$ of *consequences*, which are the atomic holistic entities that have value to the individual. Ideally, a consequence $c \in \mathscr{C}$ provides a complete description of everything the individual may be concerned about, and the occurrence of one consequence precludes the occurrence of any other consequence. Again, as an ideal, one might wish to exclude all traces of uncertainty from consequences although this is not possible in practice. Consequences are discussed further by Ramsey (1931), Savage (1954) and Fishburn (1964). Special structures for $\mathscr{C}$ arise in some of the theories considered later.

The utility $u(c)$ of consequence c is a number that provides a measure of the consequence's subjective value in relation to other consequences. Probabilities may be assigned to propositions such as "if I do such-and-so, then consequence c will occur" (Fishburn, 1964; Gibbard and Harper, 1978) but, except for mono-set theories (section 10), probabilities are not assigned to consequences alone.

2.2. *Events*

The carriers of uncertainty in most theories are propositions (Ramsey) or events (Savage), the latter of which are subsets of a set of "states of the world". Savage's description (1954, p. 9) of a state sounds very much like the foregoing description of a consequence. Indeed, the two notions appear to merge in mono-set theories. However, in most cases, states are treated as entities that are the basis of the individual's uncertainty and that have value only to the extent that they lead to specific consequences that depend on the course of action adopted by the individual. Moreover, it is usually presumed that the 'true' state, or state that obtains (e.g., 'rain' or 'no rain', 'heads' or 'tails'), which is initially unknown by the individual, cannot be changed by the individual's actions.

Throughout the paper, S denotes a nonempty set of *states*, and $\mathscr{S}$ denotes a set of *events* (subsets of S) that contains the empty event $\emptyset$ and is closed under complementation: if $A \in \mathscr{S}$ then $S \backslash A$ (also denoted $\bar{A}$) is in $\mathscr{S}$. A *measurable partition* of S is a collection of nonempty mutually disjoint events in $\mathscr{S}$ whose union equals S.

The event set $\mathscr{S}$ is a *Boolean algebra* if it is closed under finite unions:

$A, B \in \mathscr{S} \Rightarrow A \cup B \in \mathscr{S}$. It is a *σ-algebra* if it is closed under countable unions. Several theories (Pfanzagl, 1967, 1968; Bolker, 1967) use abstract algebras without an underlying S, but no real harm is done by imagining these to be algebras of subsets of some set S. Bolker's $\mathscr{S}$ is *atom free*: for every nonempty $A \in \mathscr{S}$ there are nonempty disjoint B and C in $\mathscr{S}$ for which $B \cup C \subseteq A$.

Following Hausdorff (1957), I shall use $A + B$ to denote $A \cup B$ when $A \cap B = \emptyset$. *All instances of $A + B$ carry the unwritten implication that A and B are disjoint.*

An individual's subjective probability for event $A \in \mathscr{S}$ will be denoted by $P(A)$. Without exception, P will denote a real valued function on $\mathscr{S}$ for which $P(\emptyset) = 0, P(A) \geqslant 0$ for all $A \in \mathscr{S}$, and $P(A) + P(\bar{A}) = 1$ for all $A \in \mathscr{S}$. When $\mathscr{S}$ is an algebra, P is a (finitely additive) *probability measure* if $P(A + B) = P(A) + P(B)$ for all $A, B \in \mathscr{S}$, and a *countably additive probability measure* if $P(\cup A_i) = \Sigma P(A_i)$ for every countable collection of mutually disjoint events whose union is in $\mathscr{S}$.

Several theories distinguish a set $\mathscr{N}$ of *null events*, which the individual believes cannot obtain and whose subjective probabilities are zero. When P is a probability measure on an algebra $\mathscr{S}$, and $P(A) = 0$ iff $A \in \mathscr{N}$, then $\emptyset \in \mathscr{N}$, $A, B \in \mathscr{N} \Rightarrow A \cup B \in \mathscr{N}$, and $[A \in \mathscr{S}, A \subseteq B, B \in \mathscr{N}] \Rightarrow A \in \mathscr{N}$ so that $\mathscr{N}$ is an *ideal* and, in fact, a *proper* ideal since $P(S) = 1$ with $S \notin \mathscr{N}$. The null ideal $\mathscr{N}$ is *principal* if the union of all events in $\mathscr{N}$ is in $\mathscr{N}$. Null events are usually identified with the aid of the individual's indifference relation.

In later material, 2^S is the set of all subsets of S, $\mathscr{S}_0 = \mathscr{S} \backslash \{\emptyset\}$ – the set of all nonempty events, and $\mathscr{S} \backslash \mathscr{N}$ is the set of nonnull events.

2.3. *Mixture sets and lotteries*

Following Herstein and Milnor (1953), a set $\mathscr{M}$ is said to be a *mixture set* if for any $x, y \in \mathscr{M}$ and any $\lambda \in [0, 1]$ we can associate another element in $\mathscr{M}$, written as (x, λ, y) or as $\lambda x + (1 - \lambda)y$, such that

$$(x, 1, y) = x$$

$$(x, \lambda, y) = (y, 1 - \lambda, x)$$

$$((x, \lambda, y), \mu, y) = (x, \lambda\mu, y)$$

for all $x, y \in \mathscr{M}$ and all $\lambda, \mu \in [0, 1]$.

A *lottery* on a nonempty set X is a function $p: X \to [0,1]$ such that $p(x)=0$ for all but a finite number of $x \in X$, and $\Sigma p(x) = 1$. The set of all lotteries on X is denoted as $\mathscr{M}(X)$, which is a mixture set under the interpretation that (p, λ, q) is the convex combination $\lambda p + (1-\lambda)q$ of lotteries p and q, with $(\lambda p + (1-\lambda)q)(x) = \lambda p(x) + (1-\lambda)q(x)$ for all $x \in X$.

We usually think of $\mathscr{M}(X)$ as the set of all simple probability distributions on X, with $p(x)$ the probability that x will occur if lottery p is 'played'. Lotteries will have this interpretation in the theories discussed in section 8.

2.4. *Acts*

Broadly speaking, an act is a decision or course of action that an individual might choose to pursue. In a few theories, acts are basic primitives; but in most theories, acts are constructed primitives based on consequences and states or events.

In the former approach, we might start with a set $\mathscr{F}$ of acts, identify consequences that might follow from the acts, and consider a model based on utilities for consequences and probabilities for propositions of the form "if I do $f \in \mathscr{F}$, then $c \in \mathscr{C}$ will occur". There are no Savage-type states in this model, but we can construct entities that function like states (Fishburn, 1964, 1970). This is done by viewing a 'state' s as a function from $\mathscr{F}$ into $\mathscr{C}$, i.e. as an element in the function set $\mathscr{C}^{\mathscr{F}}$. Each such $s \in \mathscr{C}^{\mathscr{F}}$ specifies the consequence that ensues from each act and therefore removes (as an ideal) all uncertainty in the situation. Moreover, the act chosen by the individual does not affect the 'true' state.

The preceding formulation of $S \subseteq \mathscr{C}^{\mathscr{F}}$ is often very complex from the states viewpoint. The alternative approach, with S as a basic primitive, is usually simpler to visualize on a conceptual level. When $\mathscr{C}$ and S are basic primitives, we shall let $\mathscr{F}$ denote a nonempty subset of $\mathscr{C}^S$ and let $\mathscr{G}$ be a nonempty subset of $\cup\{\mathscr{C}^A: A \in \mathscr{S}_0\}$. An act in $\mathscr{F}$ is a *Savage act* that assigns a consequence to each state of the world: if f is chosen and s obtains, then $f(s)$ occurs. An 'act' in $\mathscr{G}$ is a function that maps each state in a nonempty event A into a consequence. Such a function may be denoted by f_A to make its domain explicit. Various interpretations can be given to f_A, including:

(i) in Savage's theory, f_A is the restriction to A of a Savage act $f \in \mathscr{C}^S$. It is not an object of choice unless $A = S$;

(ii) Narens (1976) views f_A as a Savage act that yields consequence $f_A(s)$ if $s \in A$ obtains, and yields 'nothing' (or the status quo) if $s \in \bar{A}$ obtains;

(iii) Roberts (1974) views f_A as a Savage act that yields $f_A(s)$ if $s \in A$ obtains, and yields a 'null consequence' in an extensive measurement structure for $\mathscr{C}$ if $s \in \bar{A}$ obtains;

(iv) Luce and Krantz (1971), who require $A \in \mathscr{S} \setminus \mathscr{N}$, view f_A as a potential object for choice, with A the relevant universe of states if f_A is adopted. They refer to f_A as a conditional decision. Luce-Krantz states differ from Savage states, as discussed in section 9.

Act $f: A \to \mathscr{C}$ is *measurable* if the set of all nonempty $f^{-1}(c)$ is a measurable partition of A, i.e. if $\{s \in A: f(s) = c\}$ is in $\mathscr{S}$ for all $c \in \mathscr{C}$. In addition, f is *simple* if $f^{-1}(c) = \emptyset$ for all but a finite number of consequences. Simple measurable acts are sometimes referred to as gambles.

Other types of 'acts' use lotteries in their constructions. Two examples from Anscombe and Aumann (1963) are the set $\mathscr{H} = [\mathscr{M}(\mathscr{C})]^S$ of *horse lotteries* and the set $\mathscr{M}(\mathscr{H})$ of lotteries on horse lotteries. A horse lottery h assigns a consequence lottery to each state of the world. Implementation of $p \in \mathscr{M}(\mathscr{H})$ might unfold in three steps: the $p(h)$ probabilities first determine a horse lottery h; the chosen h then yields the consequence lottery $h(s) \in \mathscr{M}(\mathscr{C})$ when s obtains; finally, the $h(s)$ thus obtained determines a consequence according to the probabilities $h(s)(c)$, where $h(s)(c)$ is the probability that $h(s)$ assigns to c.

2.5. *Preferences*

Throughout, $\succ$ will denote an individual's *strict preference relation* on a set X which, depending on context, might be $\mathscr{F}$, $\mathscr{G}$, $\mathscr{H}$, $\mathscr{C} \times \mathscr{S} \times \mathscr{C}$, or another designated set. A triple (c, A, d) in $\mathscr{C} \times \mathscr{S} \times \mathscr{C}$ will be viewed as an 'act' that yields consequence c if A obtains (i.e. some $s \in A$ obtains) and consequence d if $\bar{A}$ obtains. In many cases, $\succ$ on X induces a strict preference relation $\succ'$ on a set – such as $\mathscr{C}$ – used in the formation of X. For example, we define $c \succ' d$ iff $(c, S, c) \succ (d, S, d)$ when $\succ$ is a primitive on $\mathscr{C} \times \mathscr{S} \times \mathscr{C}$. In Savage's theory, $c \succ' d$ iff $f \succ g$ when $f(s) = c$ and $g(s) = d$ for all $s \in S$. For notational convenience, the prime of $\succ'$ will usually be omitted.

The preference relation $\succ$ is often used to define a *qualitative probability* *relation* $\succ^*$ on the set $\mathscr{S}$ of events. (Read $x \succ y$ as "x is preferred to y" and $A \succ^* B$ as "A is more probable than B".) Given $\succ$ on $\mathscr{C} \times \mathscr{S} \times \mathscr{C}$, we define $A \succ^* B$ if and only if $c \succ d$ and $(c, A, d) \succ (c, B, d)$ for some $c, d \in \mathscr{C}$. Intuitively, if you prefer c to d, then you will prefer (c, A, d) to (c, B, d) if you believe that A is more likely than B to obtain. Note, however, that this makes good sense only if all components of value reside in the consequences. If valued aspects of events are not included in the consequences then we would be reluctant to assert $A \succ^* B$ when $c \succ d$ and $(c, A, d) \succ (c, B, d)$.

We shall always assume that $\succ$ *is asymmetric*: $x \succ y \Rightarrow$ not $(y \succ x)$. The *indifference* relation $\sim$ and *nonstrict preference* relation $\succsim$ associated with $\succ$ are defined by

$$x \sim y \quad \text{iff} \quad \text{neither} \quad x \succ y \quad \text{nor} \quad y \succ x;$$

$$x \succsim y \quad \text{iff} \quad x \succ y \quad \text{or} \quad x \sim y.$$

Many writers use a nonstrict rather than strict preference relation as a basic primitive, but for uniformity I shall use $\succ$ except in discussing the theory of Pratt, Raiffa and Schlaifer (1964, 1965) in section 8.

The relation $\succ$ will be called a *strict partial order* if it is transitive ($x \succ y$ and $y \succ z \Rightarrow x \succ z$), an *asymmetric weak order* if it is negatively transitive ($x \succ z \Rightarrow x \succ y$ or $y \succ z$), and a *linear order* if it is a complete ($x \neq y \Rightarrow x \succ y$ or $y \succ x$) asymmetric weak order. It is easily seen that $\succ$ is an asymmetric weak order if and only if both $\succ$ and $\sim$ are transitive, and in this case $\succsim$ is transitive also and is commonly referred to as a weak order. If a real valued function u on X satisfies $u(x) > u(y)$ iff $x \succ y$, for all $x, y \in X$, then $\succ$ must be an asymmetric weak order.

3. AXIOMS AND REPRESENTATIONS

Each theory reviewed later consists of a set of primitives, axioms based on the primitives, and a numerical representation implied by the axioms that reflects preferences in a subjective expected utility model. A uniqueness theorem for utilities and probabilities usually accompanies the representation. Three examples of representations, taken respectively from Pfanzagl (1967, 1968), Savage (1954), and Luce and Krantz (1971) are:

(i) $\forall (c, A, d), (c', B, d') \in \mathscr{C} \times \mathscr{S} \times \mathscr{C}$:

$$(c,A,d) \succ (c',B,d') \quad \text{iff} \quad P(A)u(c) + P(\bar{A})u(d) > P(B)u(c') + P(\bar{B})u(d');$$

(ii) $\forall f,g \in \mathscr{F} = \mathscr{C}^S$ with $\mathscr{S} = 2^S$:

$$f \succ g \quad \text{iff} \quad \int_S u(f(s))\, \mathrm{d}P(s) > \int_S u(g(s))\, \mathrm{d}P(s);$$

(iii) for all simple measurable $f_A, g_B \in \mathscr{G}$:

$$f_A \succ g_B \quad \text{iff} \quad \frac{1}{P(A)} \int_A u(f_A(s))\, \mathrm{d}P(s) > \frac{1}{P(B)} \int_B u(g_B(s))\, \mathrm{d}P(s).$$

Other features, such as boundedness of u (Savage) and connections between P and a qualitative probability relation $\succ^*$ on $\mathscr{S}$, will be noted later.

3.1. *Uniqueness*

Let $\mathscr{R}$ denote the set of all pairs (v, Q) of utility and probability functions that satisfy a specified representation, and let (u,P) be one pair in $\mathscr{R}$. Then P is said to be *unique* if $Q = P$ for all $(v, Q) \in \mathscr{R}$. When P is unique, u is *unique up to a positive affine* (linear) *transformation* if

$$\mathscr{R} = \{(au + b, P)\colon a \text{ and } b \text{ are real numbers and } a > 0\},$$

and u is *unique up to a similarity transformation* if

$$\mathscr{R} = \{(au, P)\colon a > 0\}.$$

The theories cited in the preceding paragraph have P unique and u unique up to a positive affine transformation. Special structures for $\mathscr{C}$ that are discussed later yield u unique up to a similarity transformation. Somewhat different transformations apply to the mono-set theories in section 10.

3.2. *Axiom development*

The process of theory development can be roughly described by two steps. First, the primitives and numerical representation that one is interested in are formulated. Then axioms that imply the desired representation are developed. For example, the representations (i) to (iii) require $\succ$ to be an asymmetric weak order. Moreover, if $x \succ y$ whenever $x \in \{(c, S, c), (d, S, d)\}$ and $y \in \{(c', S, c'), (d', S, d')\}$ in representation (i), then we must have $(c, A, d) \succ (c', B, d')$.

It may even be possible to deduce from the representation a set of necessary axioms which implies the representation. Granting the nonnecessary but empirically reasonable assumptions that $\mathscr{C}$ and S are finite, Richter (1975) does this for the one-way Savage-type model[2]

$$f \succ g \Rightarrow \int_S u(f(s))\, \mathrm{d}P(s) > \int_S u(g(s))\, \mathrm{d}P(s).$$

However, Richter's necessary and sufficient condition for this representation involves abstract polynomial rings, and as such it is very complex and nonintuitive even though it does illuminate the mathematical structure of the representation.

In fact, because any set of necessary and sufficient axioms for a subjective expected utility model is bound to include complicated preference axioms that have little intuitive appeal, virtually all the theories reviewed later contain fairly strong structural assumptions that are not necessary for the representation. The use of structural axioms highlights a dilemma that is pervasive in developments of subjective expected utility theories. On the one hand, we would like our axioms to be simple, interpretable, intuitively clear, and capable of convincing others that they are appealing criteria of coherency and consistency in decision making under uncertainty, but to do this it seems essential to invoke strong structural conditions. On the other hand, we would like our theory to adhere to the loose structures that often arise in realistic decision situations, but if this is done then we will be faced with fairly complicated axioms that accommodate these loose structures.

Many of the developments in theories of subjective expected utility have arisen from the conflict between axiomatic simplicity-interpretability and structural flexibility. For example, Savage's theory is axiomatically quite elegant but suffers from structural constrictions that others have attempted to relax.

We can distinguish three essentially different types of axioms on the basis of structure. The first type is a purely structural axiom that does not involve the preference relation, such as "$\mathscr{C}$ is finite" or "$\mathscr{S}$ is a nonatomic and complete Boolean algebra". The second type is a necessary condition on preferences, such as "$\succ$ is an asymmetric weak order on the set of acts", or "if A, B and $A+B$ are in $\mathscr{S}_0$ and $A \succ B$, then $A \succsim A+B$ and $A+B \succsim B$". The third type mixes $\succ$ with structural assumptions, such as "$x \succ y$ for some x and y" and "if $A \in \mathscr{S} \setminus \mathscr{N}$ and $g_B \in \mathscr{E}$, then there exists $f_A \in \mathscr{G}$ for which $f_A \sim g_B$". A further classification of preference axioms is given at the end of the next section.

4. LINEAR UTILITY

We consider two axiomatizations of von Neumann-Morgenstern (1947) linear utilities.[3] In each case, the axioms apply $\succ$ to a mixture set $\mathscr{M}$ and are meant to hold for all $x, y, z \in \mathscr{M}$. The following are from Jensen (1967) or Fishburn (1970).

(A1) $\succ$ on $\mathscr{M}$ is an asymmetric weak order.

(A2) If $x \succ y$ and $0 < \lambda < 1$ then $\lambda x + (1-\lambda)z \succ \lambda y + (1-\lambda)z$.

(A3) If $x \succ y$ and $y \succ z$ then $\alpha x + (1-\alpha)z \succ y$ and $y \succ \beta x + (1-\beta)z$ for some $\alpha, \beta \in (0, 1)$.

Our second set of axioms is from Herstein and Milnor (1953).

(B1) $\succ$ on $\mathscr{M}$ is an asymmetric weak order.

(B2) If $x \sim y$ then $\frac{1}{2}x + \frac{1}{2}z \sim \frac{1}{2}y + \frac{1}{2}z$.

(B3) $\{\alpha\colon\ \alpha x + (1-\alpha)z \succ y\}$ and $\{\beta\colon\ y \succ \beta x + (1-\beta)z\}$ are open subsets of $[0, 1]$.

The two sets of axioms are equivalent even though this is far from obvious on comparing (A2) and (A3) with (B2) and (B3). Each set holds if, and only if, there is a (real valued function) u on $\mathscr{M}$ that is *order-preserving*:

$$\forall x, y \in \mathscr{M}\colon\ x \succ y \quad \text{iff} \quad u(x) > u(y),$$

and *linear*:

$$\forall x, y \in \mathscr{M}, \forall \lambda \in [0, 1]\colon\ u(\lambda x + (1-\lambda)y) = $$
$$= \lambda u(x) + (1-\lambda)u(y).$$

Moreover, such a u is unique up to a positive affine transformation.

Let $\mathscr{M}$ be the lottery set $\mathscr{M}(X)$, let x^* denote the lottery that assigns probability 1 to $x \in X$, and define u on X by $u(x) = u(x^*)$. Then linearity implies the expected utility form

$$u(p) = \sum_X p(x)u(x)$$

for every lottery $p \in \mathscr{M}(X)$. Fishburn (1977) discusses the extension $u(p) = \int u(x)\mathrm{d}p(x)$ to more general probability measures p on an algebra of subsets of X. The extension involves the following dominance axiom, which applies to all Y in the algebra and all $x \in X$.

(A4) If $p(Y) = 1$ and $x^* \succ y^*$ for all $y \in Y$, then $x^* \succsim p$;

if $p(Y) = 1$ and $y^* \succ x^*$ for all $y \in Y$, then $p \succsim x^*$.

This axiom, which is not implied by (A1) through (A3) when $\mathscr{M}$ is a mixture set of probability measures that properly includes $\mathscr{M}(X)$, seems very appealing. For example, its first part says that if you prefer x to every y in a subset of X that is certain to contain the outcome of p, then x will be at least as preferable as p.

The second axiom in each set cited above is an independence condition between outcomes $(x, y, \ldots)$ and mixing numbers (λ) that is essential for linearity. Axiom (A2) says that if you prefer x to y then you will prefer a nontrivial mixture of x and z to a similar mixture of y and z; (B2) says that indifference between x and y entails indifference between equal mixtures of x with z and y with z.

Although axiom (B3) contains traces of independence, (A3) and (B3) are primarily Archimedean or continuity axioms which ensure that utilities will be real numbers rather than non-standard numbers (Robinson, 1966) or real vectors (Hausner, 1954).

4.1. *Subjective linear utility*

The passage from linear utility, with

$$u(x, \lambda, y) = \lambda u(x) + (1 - \lambda)u(y),$$

to *subjective* linear utility, with

$$u(x, A, y) = P(A)u(x) + [1 - P(A)]u(y),$$

can be viewed as a move to replace the given probability λ with the uncertain event $A \in \mathscr{S}$ along with its subjective probability $P(A)$. Not surprisingly, the axioms used for subjective linear utility and for more general versions of subjective expected utility reflect many of the ideas contained in the axioms presented earlier in this section.

In particular, all axiomatizations of subjective expected utility involve ordering and independence axioms, and most – exceptions being Davidson and Suppes (1956) and Richter (1975) – involve Archimedean or continuity axioms. In addition, there are often nontriviality conditions (e.g. $x \succ y$ for some x and y), special conditions to ensure that P is additive, and dominance conditions that are needed when we want the expected utility form to hold for nonsimple measurable acts. For example, Savage uses an ordering axiom, three independence axioms, a nontriviality condition, an Archimedean condition, and a dominance axiom.

Independence axioms, which along with order axioms constitute the cores of subjective expected utility theories, often serve to impute order to subjective values of consequences and to comparative qualitative probabilities of events, and to separate value from uncertainty to yield the decomposition into utility × probability that is found in most representations. Examples of independence axioms for $\succ$ on $\mathscr{C} \times \mathscr{S} \times \mathscr{C}$ are

$$(c, S, c) \succ (d, S, d) \quad \text{iff} \quad (c, A, c) \succ (d, A, d);$$

$$\{c \succ d, c' \succ d', (c, A, d) \succ (c, B, d)\} \Rightarrow (c', A, d') \succ (c', B, d');$$

The first of these is a consistency axiom for $\succ$ on $\mathscr{C}$, and the second is a consistency axiom for $\succ^*$ on $\mathscr{S}$.

Another type of independence axiom involves 'averaging' over disjoint events. Let $f_A \cup g_B$ denote the Luce-Krantz act on $A \cup B$ that equals f_A on A and g_B on B, given $A \cap B = \emptyset$ and $A, B \in \mathscr{S} \backslash \mathscr{N}$. Their third axiom is

$$f_A \sim g_B \Rightarrow f_A \sim f_A \cup g_B,$$

which might be interpreted as "if f_A is indifferent to g_B, then it is indifferent to the 'average' of f_A and g_B". With $A \cap B = \emptyset$ and $A, B \in \mathscr{S} \backslash \mathscr{N}$, an averaging condition for mono-set theories is

$$A \succ B \Rightarrow A \succ A + B \succ B.$$

According to one interpretation (Jeffrey, 1965), $A \succ B \Rightarrow A \succ A + B$ says that if you would be happier to hear that A rather than B obtains, then you would be happier to hear that A rather than A or B obtains.

5. RAMSEY

This section summarizes Ramsey's pioneering proposal for quantifying values and beliefs. It then discusses representations with equally-likely states, followed by the equally-spaced utilities theory of Davidson and Suppes (1956). The section concludes with Pfanzagl's (1967, 1968) theory, which in many respects can be viewed as a natural completion of Ramsey's ideas.

5.1. *Ramsey's proposal*

Ramsey's ideas for quantifying values and beliefs on the basis of preferences and an underlying model for subjective expected utility are extremely rich, insightful, and very carefully reasoned. The frequency with which they arise in the works of later theorists testifies to their broad appeal. At the same time, his mathematical treatment is deliberately concise and sometimes cryptic. I am aware of no account that attempts to reconstruct the complete details of his theory, except that Pfanzagl (1967, 1968) comes close, without however making use of Ramsey's "ethically neutral proposition". In the next few paragraphs I shall outline an approximate version of Ramsey's theory.

Ramsey's (1931) basic primitives are $\mathscr{C}$, a finite state set S (Ramsey uses 'propositions' rather than states and events), and $\succ$. His event set might be presumed to be $\mathscr{S} = 2^S$, and he applies $\succ$ to some subset of

$$\mathscr{C} \cup (\mathscr{C} \times \mathscr{S} \times \mathscr{C}) \cup (\mathscr{C} \times \mathscr{S} \times \mathscr{C} \times \mathscr{S} \times \mathscr{C}).$$

Act $cAd \in \mathscr{C} \times \mathscr{S} \times \mathscr{C}$ is interpreted as before: get c if A obtains, and d if $\bar{A}$ obtains. Act $(cAd)Be \in \mathscr{C} \times \mathscr{S} \times \mathscr{C} \times \mathscr{S} \times \mathscr{C}$ is interpreted as: get c if $A \cap B$ obtains, get d if $\bar{A} \cap B$ obtains, and e if $\bar{B}$ obtains. The latter acts are used to define conditional subjective probabilities like "the degree of belief in A given B".

Ramsey's quantification procedure proceeds through four steps. First, determine an "ethically neutral" event E, which satisfies $cEd \sim dEc$ for all $c, d \in \mathscr{C}$ (with $c \succ d$ for some c and d), and is to have subjective probability 1/2. Second, use $\succ$ on $\mathscr{C} \times \{E\} \times \mathscr{C}$ to scale the utilities of consequences in a utility-difference comparison mode[4] that gives $cEd \succsim c'Ed'$ iff $u(c) - u(c') \geq u(d') - u(d)$, with u unique up to a positive affine transformation. Third, use u to scale probabilities for other events according to the subjective expected utility model: e.g., if $c \succ d \succ e$ and $d \sim cAe$, then

$$P(A) = \frac{u(d) - u(e)}{u(c) - u(e)}.$$

Finally, define subjective probabilities for conditional events using the more general acts. For example, if $cAd \sim (c'Be)Ad$, and if $u(c') \neq u(e)$ and $P(A) > 0$, then

$$P(B|A) = \frac{u(c) - u(e)}{u(c') - u(e)}.$$

The latter ratio of utility differences is formed on the basis of equality between the subjective expected utilities of cAd and $(c'Be)Ad$, which respectively are $P(A)u(c) + P(\bar{A})u(d)$ and

$$P(A \cap B)u(c') + P(A \cap \bar{B})u(e) + P(\bar{A})u(d) =$$
$$= P(A)P(B|A)u(c') + P(A)[1 - P(B|A)]u(e) + P(\bar{A})u(d).$$

Ramsey's axioms postulate the existence of an ethically neutral event, assume the effective equivalence of all such events, and state several conditions for $\succ$ on $\mathscr{C} \times \{E\} \times \mathscr{C}$ that are similar in many ways to Debreu's axioms that we shall state shortly. These lead to u on $\mathscr{C}$ as noted above, with $u(\mathscr{C})$ either a real interval or a dense subset of a real interval. In other words, $\mathscr{C}$ must be infinite and give arbitrarily fine gradations in utility.

Independence-consistency conditions as well as indifference-existence axioms are then needed to obtain a unique P. For example, if A is neither certain nor impossible, then it is presumed that there are consequences $c \succ d \succ e$ such that $d \sim cAe$. Moreover, if both $d \sim cAe$ and $d' \sim c'Ae'$, then the ratios $[u(d) - u(e)]/[u(c) - u(e)]$ and $[u(d') - u(e')]/[u(c') - u(e')]$ are assumed to be equal. Ramsey does not indicate how such an equality might be expressed in a preference axiom. He does, however, note that his consistency conditions lead to the following "fundamental laws of probable belief" for nonnull events:

$$P(A) + P(\bar{A}) = 1$$
$$P(A|B) + P(\bar{A}|B) = 1$$

$$P(A \cap B) = P(A)P(B|A)$$
$$P(A \cap B) + P(A \cap \bar{B}) = P(A).$$

Finite additivity, which Ramsey doesn't mention, follows from the last of these: $P(A + B) = P((A + B) \cap A) + P((A + B) \cap \bar{A}) = P(A) + P(B)$.

With proper account taken of null events, one might attribute to Ramsey the subjective expected utility representation

$$(c_1A_1d_1)B_1e_1 \succ (c_2A_2d_2)B_2e_2 \quad \text{iff}$$
$$P(A_1 \cap B_1)u(c_1) + P(\bar{A}_1 \cap B_1)u(d_1) + P(\bar{B}_1)u(e_1) >$$
$$> P(A_2 \cap B_2)u(c_2) + P(\bar{A}_2 \cap B_2)u(d_2) + P(\bar{B}_2)u(e_2),$$

even though he does not fully axiomatize this model. By identifying cAd with $(cAd)Sc$, we obtain the specialization mentioned earlier:

$$cAd \succ c'Bd' \quad \text{iff} \quad P(A)u(c) + P(\bar{A})u(d) > P(B)u(c') + + P(\bar{B})u(d').$$

In a related investigation, Luce (1958) examines comparisons between acts such as cAd and $(cAd)Be$ on the basis of binary preference probabilities $\rho(x, y)$ where, with x and y acts of the type indicated, $\rho(x, y)$ is the probability that x is judged strictly preferred to y. Although this approach lies outside of the present review, Luce's paper offers an interesting alternative to Ramsey's deterministic-preference approach.

5.2. *Equally-likely states*

A system of axioms that is very similar to Ramsey's initial axioms has been presented by Debreu (1959). Debreu assumes that $\mathscr{C}$ is a connected (in the relative usual topology) subset of a finite-dimensional Euclidean space and applies $\succ$ to $\mathscr{C} \times \mathscr{C}$, where $cd \in \mathscr{C} \times \mathscr{C}$ can be thought of as an even-chance act between c and d. Hence, if E is a Ramsey (ethically neutral) event, then cd yields c if E and d if $\bar{E}$.

In addition to his structural axiom for $\mathscr{C}$ and weak order, Debreu assumes for all $c, \ldots, d^* \in \mathscr{C}$ that

(1) If $cd' \succsim c'd^*$ and $c'd \succsim c^*d'$ then $dc \succsim d^*c^*$;

(2) $\{cd \in \mathscr{C} \times \mathscr{C}: cd \succ c'd'\}$ and $\{cd \in \mathscr{C} \times \mathscr{C}: c'd' \succ cd\}$ are open in $\mathscr{C} \times \mathscr{C}$.

The first of these is an independence-cancellation axiom that is necessary for the desired representation. The interchange between c and d in the conclusion of (1) yields $cd \sim dc$, which captures the even-chance or ethically neutral notion. Axiom (2) is a continuity condition that is related to (B3) in the Herstein-Milnor axioms.

Debreu proves that his axioms imply the existence of a continuous u on $\mathscr{C} \times \mathscr{C}$ for which

$$\forall cd, c'd' \in \mathscr{C} \times \mathscr{C}\colon\ cd \succ c'd' \quad \text{iff} \quad u(cd) > u(c'd'),$$

$$\forall c, d \in \mathscr{C}\colon\ u(cd) = \tfrac{1}{2}u(cc) + \tfrac{1}{2}u(dd).$$

In other words, $cd \succ c'd'$ iff $u(c) + u(d) > u(c') + u(d')$, where $u(c)$ is defined as $u(cc)$, with u unique up to a positive affine transformation.

As several people have noted, Debreu's approach generalizes easily to a set S of n states that are, by way of assumption, equally likely to obtain. For example, one can begin with an axiomatization of additive conjoint measurement[5] for $\succ$ on $\mathscr{C}^n$ when $S = \{s_1, \ldots, s_n\}$ to get

$$c_1 \ldots c_n \succ d_1 \ldots d_n \quad \text{iff} \quad \sum_{i=1}^{n} u_i(c_i) > \sum_{i=1}^{n} u_i(d_i),$$

with each u_i defined on $\mathscr{C}$. The further assumption that $c \sim d$ whenever d is obtained from c by a permutation of c's components then leads, for example, to

$$c_1 \ldots c_n \succ d_1 \ldots d_n \quad \text{iff} \quad \sum_{i=1}^{n} \frac{1}{n} u(c_i) > \sum_{i=1}^{n} \frac{1}{n} u(d_i)$$

when we define $u(c)$ as $\Sigma_i u_i(c)$. A somewhat different route to this equally-likely states model, which is based on preferences but assumes that utilities are given, is discussed by Milnor (1954).

In addition, Chernoff (1954) and Maskin (1979) derive a similar model using a choice-function approach. Chernoff takes utilities as given, but alludes to the von Neumann-Morgenstern approach as one way of measuring utilities; Maskin assumes explicitly that u on $\mathscr{C}$ is based on the von Neumann-Morgenstern axioms applied to $\succ$ on $\mathscr{M}(\mathscr{C})$. The choice-function approach uses a function F on a set of 'decision problems' such that $F(\mathscr{A})$ is a non-empty subset of an act set $\mathscr{A}$ for each such $\mathscr{A}$ in a specified domain. Chernoff and Maskin then specify conditions on the choice function that lead to $F(\mathscr{A})$

as the acts in $\mathscr{A}$ that have the greatest expected utility under an equally-likely states model.

5.3. *Equally-spaced utilities*

In contrast to the equally-likely states model, Davidson and Suppes (1956) axiomatize an equally-spaced utilities model. Empirical research related to their model is discussed in Davidson *et al.* (1957).

Davidson and Suppes take $\mathscr{C}$ and S as basic: $\mathscr{C}$ is finite, and $\mathscr{S}$ contains a Ramsey event E that will have $P(E) = 1/2$. Their primitive relations are $\succ$ on $\mathscr{C}$ and a subset $\approx$ of $\mathscr{C}^2 \times \mathscr{S} \times \mathscr{C}^2$ whose generic member will be written as $cd \approx_A c'd'$ and interpreted as indifference between the Ramsey acts cAd and $c'Ad'$. They assume that $\succ$ is a linear order (prior indifference can be divided out), and require $c \neq d$ and $c' \neq d'$ whenever $cd \approx_A c'd'$. These restrictions prohibit comparisons between consequences and genuine gambles in $\approx$, thus avoiding the potential distortion of a utility of gambling factor.

To signify that S obtains, it is assumed that $cd \approx_S c'd' \Rightarrow c = c'$. The Ramsey event E is to satisfy a typical inversion axiom ($c \neq d \Rightarrow cd \approx_E dc$) along with

(1) If c and d are adjacent in $\succ$, if c^* and d^* are adjacent in $\succ$, and if $\{c \succ d, c^* \succ d^*, c \neq d^*, c^* \neq d\}$, then $cd^* \approx_E c^*d$.

By c and d adjacent in $\succ$, we mean that $c \neq d$ and no e lies between c and d in preference. Axiom (1) is an equal-spacing axiom. By their model, $cd^* \approx_E c^*d \Rightarrow \frac{1}{2}u(c) + \frac{1}{2}u(d^*) = \frac{1}{2}u(c^*) + \frac{1}{2}u(d) \Rightarrow u(c) - u(d) = u(c^*) - u(d^*)$. The restrictions in the hypotheses of (1) allow the conclusion that $u(c) - u(d) = u(d) - u(e)$ when $c \succ d \succ e$ and $c \succ d' \succ e$ for no $d' \neq d$, provided that there are at least five consequences.

Their other axioms involve a generic $\approx_A$. One provides the following connection between $\approx$ and $\succ$: $\{cd \approx_A c^*d^*,\ d \neq d^*,\ c \succ c^*\} \Rightarrow d^* \succ d$. Another posits the existence of $c, c^*, d, d^* \in \mathscr{C}$ for which $cd \approx_A c^*d^*$ with $c \neq c^*$ and $d \neq d^*$. This nonnecessary (apart from its uniqueness implication) condition leads to the definition of $P(A)$ as

$$P(A) = \frac{u(d^*) - u(d)}{u(d^*) - u(d) + u(c) - u(c^*)}.$$

The other $\approx_A$ axioms supply $\approx_A$ with properties that are necessary for the representation, including symmetry, transitivity, and Ramsey-type equality between different u ratios that could be used to define $P(A)$.

The Davidson-Suppes axioms imply that there is a utility function u on $\mathscr{C}$ and a possibly nonadditive probability function P on $\mathscr{C}$ such that u preserves $\succ$ on $\mathscr{C}$, $A \subseteq B \Rightarrow P(A) \leqslant P(B)$, and

$$cd \approx_A c^*d^* \quad \text{iff} \quad u(c) \neq u(d), u(c^*) \neq u(d^*) \quad \text{and} \quad P(A)u(c) + P(\bar{A})u(d) = P(A)u(c^*) + P(\bar{A})u(d^*).$$

In addition, if $\mathscr{C}$ has at least five consequences, then P is unique and u is unique up to a positive affine transformation. According to the definition of $P(A)$ displayed above, all subjective probabilities are rational numbers because of equally-spaced utilities.

5.4. *Pfanzagl's theory*

Because Pfanzagl's (1967, 1968) theory contains technical features that are not easily summarized, I shall describe its salient aspects in a modified format based on $\mathscr{C}$, $\mathscr{S}$, and $\succ$ on the set

$$\mathscr{P} = \{(cAd)B(c'Ad')\colon\ c, d, c', d' \in \mathscr{C}; A, B \in \mathscr{S}\}$$

of *Pfanzagl acts*. Act $(cAd)B(c'Ad')$ yields

c if $A \cap B$ obtains

d if $\bar{A} \cap B$ obtains

c' if $A \cap \bar{B}$ obtains

d' if $\bar{A} \cap \bar{B}$ obtains,

and is an obvious generalization of $(cAd)Be$ used by Ramsey (1931) and Luce (1958).

Pfanzagl's $\mathscr{C}$ is a nondegenerate real interval, and $\mathscr{S}$ is a Boolean algebra with proper ideal $\mathscr{N}$ of null events identified by (for example)

$$B \in \mathscr{N} \quad \text{iff} \quad (cAd)B(c'Ad') \sim (c^*Ad^*)B(c'Ad') \text{ for all } A \in \mathscr{S} \text{ and all } c, \ldots, d^* \in \mathscr{C}.$$

The algebra is not presumed to contain a Ramsey event E.

It is assumed (Pfanzagl, 1967) that $\succ$ on $\mathscr{P}$ is an asymmetric weak order. According to the above interpretation of Pfanzagl acts, we assume that

(1) $(cAd)B(c'Ad') \sim (c'Ad')\bar{B}(cAd)$;

(2) $(cAd)B(c'Ad') \sim (d\bar{A}c)B(d'\bar{A}c')$;

(3) $(cAd)B(c'Ad') \sim (cBc')A(dBd')$;

(4) $(cAd)B(cAd) \sim (cAd)B'(cAd)$;

(5) $(cAc)B(cAc) \sim (cA'c)B'(cA'c)$,

for all $c, d, c', d' \in \mathscr{C}$ and all $A, A', B, B' \in \mathscr{C}$. Items (1) through (3) illustrate Pfanzagl's "lack of illusion" principle, which says that the presentational format should not affect preferences between acts that have the same consequences for the same events. Item (4) allows us to define $\succ$ unambiguously as an asymmetric weak order on Ramsey acts cAd, and (5) provides a similar service for $\succ$ on $\mathscr{C}$. We write $e \sim cAd$ iff $(eA'e)B(eA'e) \sim (cAd)B'(cAd)$ for all $A', B, B' \in \mathscr{S}$, and so forth.

Pfanzagl assumes that $c \succ d$ iff $c > d$, and that for every Ramsey act cAd there is a (necessarily unique) *certainty equivalent* $\mu_A(cd) \in \mathscr{C}$ for which $cAd \sim \mu_A(cd)$. Moreover, if $e \succ cAd$ then there is a $c' > c$ such that $e \succ c'Ad$. These are parts of his continuity conditions.

Certainty equivalents are also postulated for *conditional acts* of the form $c(A|B)d$ with $B \notin \mathscr{N}$. In Pfanzagl's terms, this act leads to c if A obtains and to d if $\bar{A}$ obtains, all under the condition that B obtains. Its certainty equivalent is denoted as $\mu_{A|B}(cd)$. It is presumed in effect that $\mu_{A|B}(cd)$ is the amount in $\mathscr{C}$ that would make $(cAd)Be$ indifferent to $\mu_{A|B}(cd)Be$. This suggests the independence axiom

(6) $(cAd)B(c''Ad'') \succ (c'Ad')B(c''Ad'')$ iff

$(cAd)B(c^*Ad^*) \succ (c'Ad')B(c^*Ad^*)$,

for all $c, \ldots, d^* \in \mathscr{C}$ and all $A, B \in \mathscr{S}$, which is similar to part of Savage's sure-thing principle in the next section. Axiom (6) says that if the same Ramsey act appears in the same position of two Pfanzagl acts, then this Ramsey act can be replaced by any other Ramsey act without changing the preference (or indifference) status between the Pfanzagl acts. With cBe increasing in preference as c increases when $B \in \mathscr{S} \setminus \mathscr{N}$, and with $(cAd)Be \sim c'Be$ for some $c' \in \mathscr{C}$, (6) then implies that there is a unique $c' = \mu_{A|B}(cd)$ such that $(cAd)Be \sim c'Be$ for all $e \in \mathscr{C}$.

In addition to the order, continuity and independence conditions suggested above, we presume

(7) If $c \neq d$ and $cAd \sim cBd$ then $c'Ad' \sim c'Bd'$;

(8) If $c > d, cAd \succ cBd$ and $c' > d'$ then $c'Ad' \succ c'Bd'$,

for all $c, \ldots, d' \in \mathscr{C}$ and all $A, B \in \mathscr{S}$. These axioms provide order for qualitative probabilities. As suggested by earlier comments, (7) covers $P(A) = P(B)$, and (8) deals with $P(A) > P(B)$. Axioms for consistent conditional probabilities based on any fixed nonnull conditioning event (cf. Ramsey) can be expressed in a similar format.

The foregoing sketch indicates the principal types of axioms used by Pfanzagl. His scaling procedure is based on the certainty equivalents defined above. Taking $\mu_A = \mu_{A|S}$, and letting B be any nonnull event, his primary axioms imply:

1. $\mu_{A|B}(cc) = \mu_{S|B}(cd) = \mu_{C|B}(cd) = c$ when $B \subseteq C$.
2. $\mu_{A|B}(cd) = \mu_{\bar{A}|B}(dc)$.
3. $\mu_{A|B}(cd)$ increases in c if $A \cap B \notin \mathscr{N}$, increases in d if $\bar{A} \cap B \notin \mathscr{N}$, and is continuous in both c and d.
4. If $c \neq d$ and $\mu_{A|B}(cd) = \mu_{C|B}(cd)$ then $\mu_{A|B}(c'd') = \mu_{C|B}(c'd')$ for all $c', d' \in \mathscr{C}$.
5. $(cAd)Be \sim \mu_{A|B}(cd)Be$.
6. $\mu_A(\mu_{B|A}(cd), \mu_{B|\bar{A}}(c^*d^*)) = \mu_B(\mu_{A|B}(cc^*), \mu_{A|\bar{B}}(dd^*))$.
7. $\mu_B(\mu_{A|B}(cd), d) = \mu_{A \cap B}(cd)$.

These properties then imply that there are real-valued functions u_A on $\mathscr{C}$ for $A \in \mathscr{S} \setminus \mathscr{N}$ that are increasing and continuous and which satisfy

$$u_B(\mu_{A|B}(cd)) = u_{A \cap B}(c) + u_{\bar{A} \cap B}(d) + k$$

whenever $A \cap B, \bar{A} \cap B \in \mathscr{S} \setminus \mathscr{N}$, with k a constant that can depend on A and B. Moreover, the u_A are unique up to simultaneous (same $a > 0$) positive affine transformations.

The u_A functions provide an intermediate step in Pfanzagl's derivation. His final step, which involves the decomposition of $u_A(c)$ into $P(A)$ times

$u(c)$, where $u = u_S$, requires the existence of independent events. Given events A and B such that none of $A, B, \bar{A}$ and $\bar{B}$ is in $\mathscr{N}$, we say that B is *independent* of A iff

$$\mu_{B|A}(cd) = \mu_{B|\bar{A}}(cd) \quad \text{for all} \quad c, d \in \mathscr{C}.$$

This says that $(cBd)Ae \sim (cBd)\bar{A}e$ and leads to the equality $P(B|A) = P(B|\bar{A})$ of conditional subjective probabilities.

Let $\mathscr{S}' = \{A \in \mathscr{S}: A \notin \mathscr{N}, \bar{A} \notin \mathscr{N}\}$ and assume

(9) For every $A \in \mathscr{S}'$ there is a $B \in \mathscr{S}'$ that is independent of A.

Then there is a unique probability measure P on $\mathscr{S}$ with $P(B) = 0$ iff $B \in \mathscr{N}$, such that, for all $A \in \mathscr{S} \setminus \mathscr{N}$ and all $B \in \mathscr{S}$,

$$u(\mu_{B|A}(cd)) = \frac{P(A \cap B)}{P(A)} u(c) + \left[1 - \frac{P(A \cap B)}{P(A)}\right] u(d).$$

Among other things, this leads to the representation

$$cAd \succ c'Bd' \quad \text{iff} \quad P(A)u(c) + P(\bar{A})u(d) > P(B)u(c') + + P(\bar{B})u(d')$$

for all Ramsey acts.

An interesting aspect of Pfanzagl's derivation is the 'simultaneous' measurement of utility and probability by way of his decomposition $u_A(c) = P(A)u(c) + k$. In other words, the u_A functions, which are related by positive affine transformations for events in $\mathscr{S} \setminus \mathscr{N}$, involve both the subjective probabilities and the holistic utilities. Related 'simultaneous' derivations are found in several other theories, including Suppes (1956), Fishburn (1967) and Luce and Krantz (1971). In contrast to this, Ramsey (1931) first obtains u on the basis of E and then defines P on the basis of u, whereas Savage (1954) reverses this procedure. As noted in the next section, Savage first constructs P on $\mathscr{S}$ and then obtains u from P.

6. SAVAGE

The inspiration for Savage's theory came primarily from the works of Ramsey and von Neumann-Morgenstern, and from de Finetti's (1937) seminal contributions to personal probability.[6] Luce and Raiffa (1957, Chapter 13), Luce and Suppes (1965, pp. 298–299), and Fishburn (1975) summarize Savage's approach, Arrow (1966) gives a partial derivation of his results, and Fishburn (1970, Chapter 14) presents a complete derivation of Savage's representation-uniqueness theorem. Here I shall outline the main aspects of Savage's theory, discuss structural restrictions that have motivated later axiomatizations, and mention other Savage-type theories.

6.1. *The basic theory*

Savage's basic primitive sets are $\mathscr{C}$ and S, with $\mathscr{S}$ and $\mathscr{F}$ respectively constructed as 2^S and $\mathscr{C}^S$ (or some large subset of $\mathscr{C}^S$). In comparison to Ramsey, his major conceptual innovation was the definition of acts as functions from S into $\mathscr{C}$ and the application of $\succ$ to the Savage act set $\mathscr{F}$.

Savage uses seven axioms, including asymmetric weak order for $\succ$ on $\mathscr{F}$ and a nontriviality condition. To describe the other five, we define f_A as the restriction of Savage act f to $A \in \mathscr{S}$ and define his null event set by

$$\mathscr{N} = \{A \in \mathscr{S}\colon \quad f \sim g \quad \text{whenever} \quad f_{\bar{A}} = g_{\bar{A}}\}.$$

Preference statements involving consequences in $\mathscr{C}$ are defined in the obvious ways from $\succ$ on $\mathscr{F}$ using constant acts $f \equiv c$ [$f(s) = c$ for all $s \in S$] or constant restricted acts $f_A \equiv c$ [$f_A(s) = c$ for all $s \in A$]. In addition, $f_A \succ g_A$ means that $f' \succ g'$ whenever $f'_A = f_A$, $g'_A = g_A$ and $f'_{\bar{A}} = g'_{\bar{A}}$. It should be noted that the following axioms never compare f_A and g_B when $A \neq B$. The axioms apply to all $f, g, f', g' \in \mathscr{F}$, all $A, B \in \mathscr{S}$, and all $c, d, c', d' \in \mathscr{C}$.

(1) If $f_A = f'_A, g_A = g'_A, f_{\bar{A}} = g_{\bar{A}}$, and $f'_{\bar{A}} = g'_{\bar{A}}$, then $f \succ g$ iff $f' \succ g'$;

(2) If $A \in \mathscr{S} \backslash \mathscr{N}, f_A \equiv c$ and $g_A \equiv d$, then $f_A \succ g_A$ iff $c \succ d$;

(3) If $c \succ d, f_A \equiv c, f_{\bar{A}} \equiv d, g_B \equiv c, g_{\bar{B}} \equiv d$, and similarly for c', d', f' and g', then $f \succ g$ iff $f' \succ g'$;

(4) If $f \succ g$ then there is a finite partition $\mathscr{A}$ of S such that, $\forall A \in \mathscr{A}$, $(f'_A \equiv c, f'_{\bar{A}} = f_{\bar{A}}) \Rightarrow f' \succ g$ and $(g'_A \equiv c, g'_{\bar{A}} = g_{\bar{A}}) \Rightarrow f \succ g'$;

(5) If $f_A \succ g(s)$ for all $s \in A$ (i.e., $f_A \succ h_A$ when $h_A \equiv g(s)$) then $f_A \succsim g_A$; if $g(s) \succ f_A$ for all $s \in A$ then $g_A \succsim f_A$.

Axioms (1) and (2) explicate Savage's sure-thing principle:[7] (1) says that $\succ$ is independent of states that have identical consequences for the two acts, and is similar to (A2) and (B2) (section 4) and Pfanzagl's (6) in the preceding section; (2) aligns consequence preferences with 'restricted act' preferences on nonnull events. Independence axiom (3) is designed to make $\succ^*$ on $\mathscr{S}$ an unambiguous asymmetric weak order when this qualitative probability relation is defined by

$$A \succ^* B \quad \text{iff} \quad f \succ g \quad \text{whenever} \quad c \succ d, f_A = c, f_{\bar{A}} = d,$$
$$g_B = c \quad \text{and} \quad g_{\bar{B}} = d.$$

It is closely related to (7) and (8) in Pfanzagl's theory.

Axiom (4) is a continuity condition that prohibits any consequence c from being 'infinitely desirable' or 'infinitely undesirable', and whose partitioning feature plays a major role in the derivation of Savage's probability measure on $\mathscr{S}$. In fact, weak order, nontriviality, and (1) through (4) imply that there is a unique P on $\mathscr{S}$ such that, for all $A, B \in \mathscr{S}$,

$$A \succ^* B \quad \text{iff} \quad P(A) > P(B);$$
$$0 < \lambda < 1 \Rightarrow P(C) = \lambda P(B) \quad \text{for some} \quad C \subseteq B.$$

The latter condition says that events are 'continuously divisible' and obviously implies that S is uncountable. On the other hand, Savage's $\mathscr{C}$ can contain as few as two consequences.

Measure P is then used to construct lotteries from simple acts, and it is shown that the axioms, through (4), imply the von Neumann-Morgenstern axioms for $\succ$ defined on the set $\mathscr{M}(\mathscr{C})$ of lotteries in the obvious way. This yields a utility function u on $\mathscr{C}$ that is unique up to a positive affine transformation and establishes Savage's representation for simple acts.

Axiom (5) is a dominance axiom that is similar to (A4) in section 4. Although Savage was not aware that (5) implies that u is bounded when he wrote his book, we later proved (Fishburn, 1970, p. 206) that this is the case. It then follows, as Savage proved, that, for any $f, g \in F$,

$$f \succ g \quad \text{iff} \quad \int_S u(f(s))\, dP(s) > \int_S u(g(s))\, dP(s).$$

6.2. *Structural restrictions*

Although it may be too much to expect that people's indifference relations or even their preference relations will always be transitive (Luce, 1956;

Tversky, 1969; Fishburn, 1970), or that they will generally satisfy sure-thing principles and other independence axioms (Allais, 1953; Ellsberg, 1961; MacCrimmon, 1968; Slovic and Tversky, 1974), many writers argue that these conditions are appealing requirements for an ideally coherent and consistent individual. From the normative viewpoint, theorists have been more concerned with nonnecessary structural and existential assumptions.

One of the structural features of Savage's theory that has been criticized is axiom (4) and its implication that events are continuously divisible, which prevents his theory from being applied directly to situations in which S is countable. Other theories that also give the nice uniqueness properties for (u, P) obtained by Savage were designed in part for finite or arbitrary state spaces. These include the primary theories in the preceding section as well as the majority of theories discussed later. However, all of these pay a structural price somewhere else in order to obtain uniqueness properties.

For example, Davidson and Suppes (1956) require equally-spaced utilities, and others, including Ramsey (1931), Pfanzagl (1967) and Suppes (1956), require $\mathscr{C}$ to be infinite with $u(\mathscr{C})$ dense in a nondegenerate real interval. Another route to uniqueness with finite or arbitrary S (see section 8) is to construct lotteries from random devices or scaling probabilities that are causally independent of S. Savage (1954, pp. 33 and 38), in fact, alluded to this route. For example, with S arbitrary, replace it by $S^* = S \times T$, where T is a set of auxiliary states that are identified with the points on the perimeter of a wheel-of-fortune device. Then P^* on $\mathscr{S}^*$ is obtained from Savage's theory, and P on $\mathscr{S}$ is defined from P^* by $P(A) = P^*(A \times T)$.

Another criticism of Savage's structure is his use of 2^S as the set of events. Even when S has a nice structure, such as $S = [0, \infty)$, 2^S will contain subsets of S that are impossible to visualize. This and the fact that Savage's probability measure need not be countably additive are discussed by Savage (1954, pp. 42–43) and Dubins and Savage (1965).

The most objectionable aspect of Savage's structure for some theorists appears to be his use of constant acts ($f \equiv c$) and the associated $\mathscr{F}$ as a large subset of $\mathscr{C}^S$. In virtually any realistic problem that is formulated in the Savage mode, some consequences will be incompatible with some states or events, as is "carry an umbrella on a bright, sunny day" with 'rain'. In fact, the natural set of consequences that could occur under one state may be disjoint from the set that could occur under another state. This difficulty did not greatly bother Savage since he felt that the preference comparisons required by his axioms were conceptually reasonable. However, others have taken exception to this opinion and have formulated theories (Fishburn, 1970, section 13.2; 1973; Luce and Krantz, 1971; Richter, 1975) that substantially alleviate the "constant act problem".

6.3. *Savage-type theories*

Other axioms for $\succ$ on $\mathscr{F} \subseteq \mathscr{C}^S$ have been presented by several writers. I have already cited Richter's (1975) polynomial-ring approach for the one-way Savage representation with finite S and $\mathscr{C}$. Elsewhere (Fishburn, 1975) I have discussed simpler conditions for $\succ$ as a strict partial order on $\mathscr{F}$ without, however, giving axioms that are sufficient for the one-way representation. The other theories mentioned in this section assume that $\succ$ is an asymmetric weak order.

Stigum (1972) takes S finite with $n \geqslant 3$ states, lets $\mathscr{C}$ be the set of all nonnegative vectors in a finite-dimensional Euclidean space – as is often done in consumer preference theory, and represents $\mathscr{F}$ as the product set $\mathscr{C}^n$ with each coordinate corresponding to a particular state. He assumes that $\succ$ on $\mathscr{C}^n$ is preserved by a utility function that is continuous, increasing and strictly quasi-concave, then observes that an independence axiom that is tantamount to Savage's axiom (1) or to Debreu's (1960) independence axiom for additive conjoint measurement yields $u_1, \ldots, u_n$ on $\mathscr{C}$ such that

$$c_1 \ldots c_n \succ d_1 \ldots d_n \quad \text{iff} \quad \sum_{i=1}^{n} u_i(c_i) > \sum_{i=1}^{n} u_i(d_i),$$

for all $c_1 \ldots c_n$, $d_1 \ldots d_n \in \mathscr{C}^n$, with each u_i continuous, increasing and strictly quasi-concave, and with $u_i(\mathbf{0}) = 0$ for each i. Within this setting he then states four conditions that are necessary and sufficient for the existence of u on $\mathscr{C}$ (continuous, increasing, strictly concave) and P on 2^S such that

$$c_1 \ldots c_n \succ d_1 \ldots d_n \quad \text{iff} \quad \sum_{i} P(s_i)u(c_i) > \sum_{i} P(s_i)u(d_i)$$

for all pairs of acts. Stigum's first three conditions for this form are similar to Savage's (2), (3) and a finite-states version of (4). His fourth condition is a type of marginal rates of substitution axiom that makes connections between different states.

Narens (1976) considers tradeoffs between uncertainty and utility in two models, the first of which applies $\succ$ to $\mathscr{C} \times \mathscr{S}_0$. A generic element cA in $\mathscr{C} \times \mathscr{S}_0$ yields c if A obtains and 'nothing' otherwise. Narens assumes that every $c \in \mathscr{C}$ is preferred to 'nothing', and, using axioms that are related to Pfanzagl's (1967) and Savage's (1954) but involve some new twists due to the structure of $\mathscr{C} \times \mathscr{S}_0$, he obtains a positive u on $\mathscr{C}$ and a probability measure P on $\mathscr{S}$ such that

$$cA \succ dB \quad \text{iff} \quad u(c)P(A) > u(d)P(B),$$

for all cA, $dB \in \mathscr{C} \times \mathscr{S}_0$, with u unique up to a similarity transformation and P unique. Uniqueness up to a similarity transformation rather than a positive affine transformation is explained by the omission of 'nothing' from $\mathscr{C}$ and the representation. If we let $z =$ 'nothing', then the foregoing could be written as

$$cAz \succ dBz \quad \text{iff} \quad P(A)u(c) + P(\bar{A})u(z) > P(B)u(d) + P(\bar{B})u(z)$$

with $u(z)$ fixed at 0. Explicit inclusion of z and removal of the restriction $u(z) = 0$ would then give u unique up to a positive affine transformation.

Narens's second model falls more directly into the Savage tradition. As in the first model, $\mathscr{C}$ and S are nonempty sets, $\mathscr{S}$ is a Boolean algebra of subsets of S, and each $c \in \mathscr{C}$ is presumed to be better than 'nothing'. In the second model, $\succ$ is applied to a set $\mathscr{G} \subseteq \cup\{\mathscr{C}^A : A \in \mathscr{S}_0\}$ such that each $f_A \in \mathscr{G}$ is a 'gamble' or *simple* measurable act – of the Savage type under the interpretation that it yields 'nothing' if $\bar{A}$ obtains. I shall denote the constant act $f_S \equiv c$ as c and the conditioned-on-A constant act $f_A \equiv c$ as c_A. Narens uses the following four axioms in addition to asymmetric weak order and an Archimedean condition. In the axioms, events are always in $\mathscr{S}_0$ and, as before, $A + B$ always indicates that $A \cap B = \emptyset$.

(1) If $c_A \succ c_B$ then $d_A \succ d_B$. If $c_A \succ d_A$ then $c_B \succ d_B$;

(2) If $c \succ d$ then $c_A \sim d$ for some A. For every $g \in \mathscr{G}, g \sim c$ for some $c \in \mathscr{C}$;

(3) There exists $c \in \mathscr{C}$ such that for all A, B, D: $c_A \succ c_B$ iff $c_{A+D} \succ c_{B+D}$. For all $c \in \mathscr{C}$ and all A, B, if $c_A \succ c_B$ then $c_D \sim c_B$ for some $D \subset A$;

(4) For all $c \in \mathscr{C}$, all A, B, C, D, and all $f_C, g_D \in \mathscr{G}$: if $c_A \sim f_C$ and $c_B \sim g_D$ then $c_{A+B} \sim f_C + g_D$.

In the final axiom, $f_C + g_D$ indicates that $C \cap D = \emptyset$ with $h = f_C + g_D$ the act in $\mathscr{G}$ defined on $C \cup D$ that equals f_C on C and g_D on D.

Axiom (1) is an independence axiom for qualitative probability ($c_A \succ c_B \Rightarrow d_A \succ d_B$) and for consequence preferences ($c_A \succ d_A \Rightarrow c_B \succ d_B$). The latter part of (1) is reasonable only if every event in $\mathscr{S}_0$ has positive probability. Axiom (2) is a tradeoff condition. Given $c \succ d$, we imagine shrinking A to a point where c_A becomes indifferent to d. The second part of (2) provides a certainty equivalent for every act.

The first part of (3) posits a valued consequence and a nontrivial additive qualitative probability structure. The second part of (3) is like the first part of (2). The final axiom calls for the preservation of indifference between a (restricted) consequence and acts under disjoint unions. It is similar in spirit to the Herstein-Milnor independence axiom (B2) in section 4, or to (A2) in the same section.

Although f_A has different interpretations for Savage and for Narens, the latter's axioms are very similar in their intentions to Savage's axioms, excluding his axiom (5). Narens's axioms imply that there is a positive u on $\mathscr{C}$ and a probability measure P on $\mathscr{S}$ with $P > 0$ on $\mathscr{S}_0$ such that

$$f_A \succ g_B \quad \text{iff} \quad \int_A u(f_A(s))\, \mathrm{d}P(s) > \int_B u(g_B(s))\, \mathrm{d}P(s),$$

for all $f_A, g_B \in \mathscr{G}$. (Recall that each act in $\mathscr{G}$ uses only a finite number of consequences and that 'nothing' is omitted from $\mathscr{G}$, so that u('nothing') $P(\bar{A})$ does not appear in the expected utility form for f_A.) In addition, u is unique up to a similarity transformation, and P is unique.

The final theory considered in this section is due to Roberts (1974). Like Luce (1972) (see section 9), Roberts uses an extensive measurement structure for $\mathscr{C}$ that posits a primitive binary operation $*$ on $\mathscr{C}$ such that $c*d \in \mathscr{C}$ whenever $c, d \in \mathscr{C}$. The meaning of $c*d$ is roughly 'c and d'. For example, if consequences are monetary prizes, then $c*d = c + d$. There is a subset $\mathscr{N}'$ of null consequences in $\mathscr{C}$, with $c*d = c$ or $c*d \sim c$ whenever $d \in \mathscr{N}'$. In the monetary case, $\mathscr{N}' = \{0\}$. An act $f_A \in \mathscr{G}$ assigns a null consequence to every $s \in \bar{A}$ and a nonnull consequence in $\mathscr{C} \setminus \mathscr{N}'$ to each $s \in A$, for $A \in \mathscr{S} \setminus \mathscr{N}$.

Apart from $*$ on $\mathscr{C}$, Roberts's theory is quite similar to Narens's (1976) for $\succ$ on *his* $\mathscr{G}$. With $*$, Roberts provides a Savage-type theory with an extensive measurement structure. The binary operation $*$ on $\mathscr{C}$ extends naturally to $\mathscr{G}$ with the definition $(f*g)(s) = f(s)*g(s)$ for each $s \in S$. Roberts proves that his axioms imply u on $\mathscr{C}$ and probability measure P on $\mathscr{S}$ such that

$$u(c) = 0 \quad \text{iff} \quad c \in \mathscr{N},$$

$$u(c*d) = u(c) + u(d),$$

$$P(A) = 0 \quad \text{iff} \quad A \in \mathscr{N},$$

and

$$f \succ g \quad \text{iff} \int_S u(f(s))\,\mathrm{d}P(s) > \int_S u(g(s))\,\mathrm{d}P(s),$$

for all simple measurable $f, g \in \mathscr{G}$. For this representation, u is unique up to a similarity transformation, and P is unique. As Luce (1972) notes, $u(c*d) = u(c) + u(d)$ is a very restrictive and often unrealistic conclusion. For example, it says that utility for money is linear in the amount, or $u(c) = kc$, when $\mathscr{C}$ is a set of monetary prizes.

7. SUPPES

As an alternative to Savage's approach, Suppes (1956) developed a theory based on a Ramsey event in conjunction with Savage acts. This enabled Suppes to adopt an arbitrary S in contrast to Savage's continuously divisible structure, but also forced $\mathscr{C}$ to be infinite in contrast to Savage's arbitrary $\mathscr{C}$. Like Savage, Suppes assumes that all constant acts are in $\mathscr{F}$ and appears to take $\mathscr{S} = 2^S$. He does not use a nontriviality axiom, but we shall presume that $\succ$ is nonempty for expositional purposes.

Suppes's basic primitives are $\mathscr{C}$, S and $\succ$ on $\mathscr{F} \times \mathscr{F}$, where $\mathscr{F}$ is a subset of $\mathscr{C}^S$ that contains all constant acts. He interprets $fg \in \mathscr{F} \times \mathscr{F}$ as an even-chance gamble that yields f or g. The Ramsey event E implicit in this interpretation (f if E, g if $\bar{E}$) is supposed to be independent of the state that obtains, and there need be no $A \in \mathscr{S}$ for which $P(A) = 1/2$. Needless to say, "$fg \sim gf$ for all $f, g \in \mathscr{F}$" is one of Suppes's axioms. He assumes also that $\succ$ on $\mathscr{F} \times \mathscr{F}$ is an asymmetric weak order, and extends $\succ$ to $\mathscr{F}$ by the definition $f \succ g$ iff $ff \succ gg$. The typical independence axiom "$f \succ g$ iff $fh \succ gh$" ensures that $\succ$ on $\mathscr{F}$ is an asymmetric weak order.

Suppes's other axioms consist of an additivity condition, a continuity-denseness axiom, an Archimedean axiom, two dominance conditions, and an existential continuity-midpoint axiom. We write these in the order indicated, with $f(s)^*$, $g(s)^*$, and so forth denoting constant acts with the designated consequences. An approximate meaning of fgL^nfh in axiom (3) is that h is between f and g with h one unit from f when the preference interval from f to g is divided into 2^n equal units. The axioms apply to all $f, g, f', g', f'', g'' \in \mathscr{F}$.

(1) If $fg' \succsim f'g''$ and $f'g \succsim f''g'$ then $fg \succsim f''g''$;

(2) If $fg \succ f'g'$ and $g \succ g'$, then there is an $h \in \mathscr{F}$ such that $g \succ h \succ g'$ and $fg \succ f'h$;

(3) If $f \succ g$ and $f' \succ g'$, then there is an $h \in \mathscr{F}$ and a positive integer n such that fgL^nfh and $f'h \succsim fg'$;

(4) If $f(s)^*g(s)^* \succsim f'(s)^*g'(s)^*$ for all $s \in S$, then $fg \succsim f'g'$;

(5) There is an $h \in \mathscr{F}$ such that $h(s)^* \succsim f(s)^*$ and $h(s)^* \succsim g(s)^*$ for every $s \in S$;

(6) There is an $h \in \mathscr{F}$ such that $f(s)^*g(s)^* \sim h(s)^*$ for every $s \in S$.

Only (1) and (4) place no structural demands on $\mathscr{F}$.

Axiom (1) is Suppes's additivity or independence-cancellation axiom which, in the presence of $fg \sim gf$, is tantamount to Debreu's (1959) axiom (1) in section 5. Axiom (4) is a dominance condition that is related to Savage's axiom (5). It says that if fg is as good as $f'g'$ for each state, then fg is holistically as good as $f'g'$. Its companion, axiom (5) above, says that for every fg there is an h that dominates both f and g. If $\mathscr{F}$ were taken as $\mathscr{C}^S$, then (5) would be unnecessary since we could define $h(s)$ as the more preferred of $f(s)$ and $g(s)$.

Axiom (6) is a midpoint axiom in the sense that $c^*d^* \sim e^*$ indicates that e^* is midway in preference between c^* and d^* according to the even-chance model. Axiom (2) posits an act h that is between g and g' ($g \succ h \succ g'$) which, when substituted for g' in $fg \succ f'g'$, does not change the direction of preference. Given $f \succ g$ and $f' \succ g'$, axiom (3) essentially prohibits the preference differential between f and g from being infinitely greater than that between f' and g'.

The foregoing axioms imply a real valued function ϕ on $\mathscr{F}$ such that

$$fg \succ f'g' \quad \text{iff} \quad \phi(f) + \phi(g) > \phi(f') + \phi(g')$$

for all $f, g, f', g' \in \mathscr{F}$, with ϕ unique up to a positive affine transformation. Then, defining $u(c) = \phi(c^*)$, it is shown that there is a probability measure P on $\mathscr{S}$ such that

$$\phi(f) = \int_S u(f(s))\,\mathrm{d}P(s)$$

for all $f \in \mathscr{F}$ that are *bounded* in the sense that $\{s: a \leqslant u(f(s)) \leqslant b\} = S$ for some real a and b. Suppes thus obtains Savage's representation for bounded acts.

Unlike the case for Savage, there is nothing in Suppes's theory that forces u to be bounded. Moreover, P need not be unique. For example, if $\mathscr{F}$ consists only of constant acts, which is permitted by the axioms, then *every* probability measure P on $\mathscr{S}$ will satisfy $\phi(f) = \int u(f)\,\mathrm{d}P$.

Fishburn (1967, section 5; 1970, p. 189) presents a related Ramsey-Savage theory that applies $\succ$ to $\mathscr{F} \times \mathscr{F}$ with S finite and $\mathscr{F} = \mathscr{C}^S$. He assumes that $\mathscr{C}$ is a connected and separable topological space, that $\succ$ on $\mathscr{F} \times \mathscr{F}$ is an asymmetric weak order, and that Debreu's axioms (1) and (2) in section 5 hold for $\succ$ applied to $\mathscr{F} \times \mathscr{F}$ rather than $\mathscr{C} \times \mathscr{C}$. These axioms imply a continuous ϕ on $\mathscr{F}$ for which $fg \succ f'g'$ iff $\phi(f) + \phi(g) > \phi(f') + \phi(g')$, with ϕ unique up to a positive affine transformation.

Two more axioms then imply a unique probability measure P on $\mathscr{S} = 2^S$ that satisfies $\phi(f) = \int u(f(s))\,\mathrm{d}P(s)$ for all acts when we define u on $\mathscr{C}$ by $u(c) = \phi(c^*)$. The first of these is a simple nontriviality axiom. The second is a sure-thing principle for each state. It says that if $t \in S$ and if $\{f(s), g(s)\} = \{f'(s), g'(s)\}$ for all $s \in S\backslash\{t\}$, then

$$f(t)^* g(t)^* \succsim f'(t)^* g'(t)^* \Rightarrow fg \succsim f'g';$$

$$f(t)^* g(t)^* \succ f'(t)^* g'(t)^* \text{ and } t \text{ nonnull} \Rightarrow fg \succ f'g',$$

where t is null iff $fg \sim f'g'$ whenever $\{f(s), g(s)\} = \{f'(s), g'(s)\}$ for all $s \neq t$.

8. LOTTERY THEORIES

In contrast to the use of a single extraneous Ramsey event as a measuring device (Suppes, 1956; Debreu, 1959), a number of writers have proposed the use of lottery sets based on extraneous random devices or given scaling probabilities for measuring utilities and subjective probabilities. It appears from Arrow (1951) and Suppes (1956) that Herman Rubin made early contributions to this approach, but I do not know to what extent the theories reviewed in this section were anticipated by Rubin.[9]

As we proceed, it will become evident that the structure provided by lotteries allows more flexibility in the basic decision structure than is usually granted by theories that do not use extraneous measurement devices. This attractive feature of some lottery-based theories has not, however, engendered their general acceptance. For example, Krantz *et al.* (1971, section 8.6) criticize theories for individual decisions under uncertainty that rely on extraneous scaling probabilities. Their main objection is that these probabilities are assumed *a priori* and are not derived from preferences. This type of criticism can be at least partially avoided by combining extraneous random events in an augmented state space as was suggested for Savage's theory in section 6.

8.1. *Anscombe and Aumann*

Anscombe and Aumann's (1963) aim was to define subjective probabilities for the states in a finite S on the basis of lotteries. Their consequence set $\mathscr{C}$ contains at least two nonindifferent elements. They assume that the individual has a preference relation $\succ_0$ on the set $\mathscr{M}(\mathscr{C})$ of consequence lotteries, and another preference relation $\succ$ on the set $\mathscr{M}(\mathscr{H})$ of lotteries defined on horse lotteries in $\mathscr{H} = [\mathscr{M}(\mathscr{C})]^S$. Each relation is assumed to satisfy the von Neumann-Morgenstern axioms on its lottery set. They assume further that $\mathscr{C}$ contains a most-preferred consequence c_1 and a least-preferred consequence c_0 with $c_1 \succ_0 c_0$, and that $h_1 \succ h_0$ when these are constant horse lotteries with $h_1(s) = c_1$ and $h_0(s) = c_0$ for all $s \in S$. Here and later, c is identified with the lottery in $\mathscr{M}(\mathscr{C})$ that assigns probability 1 to c, and $h \in \mathscr{H}$ is identified with the lottery in $\mathscr{M}(\mathscr{H})$ that assigns probability 1 to h.

The foregoing assumptions imply the existence of linear, order-preserving u_0 on $\mathscr{M}(\mathscr{C})$ and u on $\mathscr{M}(\mathscr{H})$, which Anscombe and Aumann normalize so that $u_0(c_1) = 1$, $u_0(c_0) = 0$, $u(h_1) = 1$ and $u(h_0) = 0$. Two more axioms are specified for all $h, h', h^1, \ldots, h^m \in \mathscr{H}$, all $t \in S$, and all nonnegative $\lambda_1, \ldots, \lambda_m$ for which $\Sigma\lambda_i = 1$:

(1) If $h(s) = h'(s)$ for all $s \in S\backslash\{t\}$, and $h(t) \succsim_0 h'(t)$, then $h \succsim h'$;

(2) If $h(s) = \Sigma_{i=1}^{m} \lambda_i h^i(s)$ for all $s \in S$, then $h \sim \Sigma_{i=1}^{m} \lambda_i h^i$.

Axiom (1) is a form of sure-thing principle that interconnects $\succ_0$ and $\succ$ in the expected manner. Axiom (2) says that the mixture of horse lotteries $\Sigma\lambda_i h^i$ is indifferent to the horse lottery formed by first applying the same probability mixture to each state. Whatever s obtains, h and $\Sigma\lambda_i h_i$ have the same probabilities for the consequences.

With u_0 and u scaled as indicated, it then follows that there is a unique probability distribution P on S such that

$$u(h) = \sum_S P(s)u_0(h(s))$$

for all $h \in \mathscr{H}$. In addition,

$$u\left(\sum_i \lambda_i h^i\right) = \sum_i \sum_S \lambda_i P(s) \sum_{\mathscr{C}} [h^i(s)(c)]u_0(c)$$

for $\Sigma\lambda_i h^i \in \mathscr{M}(\mathscr{H})$, and

$$u(f) = \sum_S P(s)u_0(f(s))$$

for Savage acts f. Thus, for all $f, g \in \mathscr{F}$, where $\mathscr{F}$ is viewed as a specialization of $\mathscr{H}$, we have

$$f \succ g \quad \text{iff} \quad \sum_S P(s)u_0(f(s)) > \sum_S P(s)u_0(g(s)).$$

Ferreira (1972) generalizes the theory of Anscombe and Aumann to accommodate an infinite state space. With $\mathscr{S} = 2^S$, he adopts Jensen's (1967) linear utility axioms for $\succ_0$ on $\mathscr{M}(\mathscr{C})$ and $\succ$ on $\mathscr{M}(\mathscr{H})$, assumes (2) above along with $h \succ h'$ iff $x \succ_0 y$ when $h(s) = x$ and $h'(s) = y$ for all $s \in S$, and uses the following interconnecting dominance axiom (3) and monotone continuity condition (4):

(3) If $h(s) \succsim_0 h'(s)$ for all $s \in S$, then $h \succsim h'$;

(4) If $h^1 \succsim h^2 \succsim \ldots$, $\{s\colon h^i(s) \neq h(s)\} \downarrow \emptyset$, and $h^i \succsim h'$ for $i = 1, 2, \ldots$, then $h \succsim h'$.

Here $\{s: h^i(s) \neq h(s)\} \downarrow \emptyset$ means that for every s there is a smallest i such that $h^i(s) \neq h(s)$ with $h^j(s) = h(s)$ for all $j > i$. When the h^i monotonically converge to h in this way and they do not increase in preference with $h^i \succsim h'$ for all i, then (4) requires $h \succsim h'$. This axiom, like Fishburn's (1972) condition for countable additivity, leads to a countably additive P.

Ferreira does not assume that $\mathscr{C}$ has a best and a worst consequence. His axioms, along with nontriviality, imply the existence of linear u_0 on $\mathscr{M}(\mathscr{C})$, linear u on $\mathscr{M}(\mathscr{H})$, and a unique countably additive probability measure P on 2^S such that

$$u(h) = \int_S u_0(h(s))\,\mathrm{d}P(s)$$

for all $h \in \mathscr{H}$. This requires of course that u and u_0 be properly aligned. Under the normalizations of Anscombe and Aumann with $c_1 \succ_0 c_0$, we have $P(A) = u(h_A)$ when $h_A(s) = c_1$ for all $s \in A$ and $h_A(s) = c_0$ for all $s \in \bar{A}$. Ferreira proves also that every $h \in \mathscr{H}$ is bounded in the sense that $P\{s: a \leqslant u_0(h(s)) \leqslant b\} = 1$ for some finite a and b, and that if there is a denumerable partition of S such that $P(A) > 0$ for each A in the partition, then u_0 must be bounded (cf. Blackwell and Girshick, 1954).

8.2. *Pratt, Raiffa and Schlaifer*

The approach used by Pratt, Raiffa and Schlaifer (1964, 1965) is designed to simultaneously measure utilities and subjective probabilities in a direct and intuitively appealing way on the basis of 'canonical lotteries'. Like Anscombe and Aumann (1963), they assume that $\mathscr{C}$ has a best consequence c_1 and a worst consequence c_0, and that $\mathscr{S} = 2^S$. Because their primitive relation $\succsim$ is not presumed to be complete, it will be convenient to maintain it as primitive rather than $\succ$, and to define $\sim$ and $\succ$ by $x \sim y$ iff ($x \succsim y$ and $y \succsim x$), and $x \succ y$ iff ($x \succsim y$ and not $y \succsim x$).

As in their paper (1964), I shall assume initially that S is finite and let L denote the set of all functions x from $S \times [0, 1]^2$ into $\mathscr{C}$. The function x is a *canonical lottery* that yields consequence $x(s, \alpha, \beta)$ if s obtains and if a random device that selects a point in the unit square $[0, 1]^2$ according to the uniform probability distribution on the square (independent of s) selects point (α, β).

They apply $\succsim$ to L and use the first component of points in $[0, 1]^2$ to scale utilities of consequences, with the second used to scale subjective probabilities of events as suggested by axiom (2) that follows. The relation $\succsim$ on L is extended to $\mathscr{C}$ in the natural way: $c \succsim d$ iff $x \succsim y$ when $x(s,\alpha,\beta) = c$ and $y(s,\alpha,\beta) = d$ for all $(s,\alpha,\beta) \in S \times [0,1]^2$. A *generalized interval* in $[0,1]^2$ is a rectangle with sides parallel to the axes, or else a finite union of such rectangles.

The first two axioms stated by Pratt, Raiffa and Schlaifer explicate their basic use of canonical lotteries. The first axiom applies to all $c, d \in \mathscr{C}$, all $A \subseteq S$, all generalized intervals $I_1, I_2 \subseteq [0,1]^2$, and all $x_1, x_2 \in L$ defined by

$$x_i(s,\alpha,\beta) = \begin{cases} c & \text{if } (s,\alpha,\beta) \in A \times I_i \\ d & \text{otherwise.} \end{cases}$$

We let $\mu(I)$ denote the area of I.

(1) If $c \succ d$ and $\mu(I_1) = \mu(I_2)$, then $x_1 \sim x_2$;
if $A = S, c \succ d$ and $\mu(I_1) > \mu(I_2)$, then $x_1 \succ x_2$.

The first part of this axiom indicates that the random device for $[0,1]^2$ is viewed as independent of S by the individual, and that its probability distribution over $[0,1]^2$ is believed to be uniform. Thus the present authors provide a behavioral basis for their extraneous scaling probabilities that relates to the suggestion made for Savage's theory in the second paragraph of this section. The second part of axiom (1) is an obvious monotonicity assumption.

The second axiom uses the following special canonical lotteries:

$$\begin{aligned} x_A:&\quad x_A(s,\alpha,\beta) = c_1 \quad \text{if} \quad s \in A, = c_0 \text{ otherwise;} \\ x(\lambda)_1:&\quad x(\lambda)_1(s,\alpha,\beta) = c_1 \quad \text{if} \quad 0 \leqslant \alpha \leqslant \lambda, = c_0 \text{ otherwise;} \\ x(\lambda)_2:&\quad x(\lambda)_2(s,\alpha,\beta) = c_1 \quad \text{if} \quad 0 \leqslant \beta \leqslant \lambda, = c_0 \text{ otherwise.} \end{aligned}$$

According to prior interpretations, x_A yields c_1 iff A obtains, and $x(\lambda)_i$ yields c_1 with probability λ, with c_0 the outcome otherwise.

(2) There exist $c_0, c_1 \in \mathscr{C}$ with $c_1 \succ c_0$, $u\colon \mathscr{C} \to [0,1]$, and $P\colon \mathscr{S} \to [0,1]$ such that:

$\forall c \in \mathscr{C}, \quad c \sim x(u(c))_1$;

$\forall A \in \mathscr{S}, \quad x_A \sim x(P(A))_2$.

The u part of axiom (2) leads to $c_1 \sim x(1)_1$ and $c_0 \sim x(0)_1$, hence $u(c_1) = 1$ and $u(c_0) = 0$, with the utilities of other consequences falling between 0 and 1. The second part identifies $P(A)$ as the 'probability' at which you are indifferent between getting the preferred c_1 if A obtains and getting c_1 if an event with area $P(A)$ in $[0, 1]^2$ obtains for the extraneous random device.

In addition to (1) and (2), it is assumed that $\succsim$ is transitive and that $\sim$ is preserved under certain types of substitutions: if corresponding 'pieces' of two canonical lotteries are respectively indifferent, then the whole canonical lotteries will be indifferent. It is then proved that u and P as posited in axiom (2) are unique and that, for all Savage-type acts x and y of the form

$$x(s, \alpha, \beta) = c_s \quad \text{for each } s \in S \quad \text{and all } (\alpha, \beta) \in [0, 1]^2,$$

$$y(s, \alpha, \beta) = d_s \quad \text{for each } s \in S \quad \text{and all } (\alpha, \beta) \in [0, 1]^2$$

we have

$$x \succsim y \quad \text{iff} \quad \sum_S P(s)u(c_s) \geqslant \sum_S P(s)u(d_s).$$

A variety of related interesting results are also proved by the authors, but I shall not go into these here.

Pratt, Raiffa and Schlaifer (1965, Chapter 8) extend the form $\Sigma P(s)u(c_s)$ to the more general $\int u(f(s))\,dP(s)$ under three more axioms (dominance, general substitutability, continuity) in a manner similar to Ferreira's (1972) extension of Anscombe and Aumann (1963). Their dominance axiom says that if $x(s, \alpha, \beta) \succsim y(s, \alpha, \beta)$ for all (s, α, β), then $x \succsim y$. General substitutability asserts that if x depends only on s, y depends only on s and α, and if, for all $(s, \alpha) \in S \times [0, 1]$, $y(s, \alpha) = c_1$ if $\alpha \leqslant u(x(s))$ and $= c_0$ otherwise, then $x \sim y$. Their continuity axiom essentially says that if x is bounded in preference by two special sequences of canonical lotteries, each of which converges to y, then $x \sim y$.

8.3. *Preferences on horse lotteries*

We now consider theories for preferences on $\mathscr{H}$ or on a structure similar to $\mathscr{H}$ rather than on $\mathscr{M}(\mathscr{C})$ and $\mathscr{M}(\mathscr{H})$ as in the Anscombe and Aumann approach. Horse lotteries in $\mathscr{H}$ are very similar to canonical lotteries in L as defined above. For example, suppose for each $s \in S$ that $x \in L$ assigns a finite number of consequences to the points in $[0, 1]^2$ such that each

$\{(\alpha,\beta): x(s,\alpha,\beta)=c\}$ is Lebesgue measurable with measure $\mu_s(c)$. Then x corresponds to the horse lottery h for which $h(s)=\mu_s$ for each $s\in S$. Hence the structure discussed below can be viewed as a variant of that used by Pratt, Raiffa and Schlaifer.

We begin with S finite and $\mathscr{H} = [\mathscr{M}(\mathscr{C})]^S$, then consider generalizations. Fishburn (1967) shows that Jensen's (1967) linear utility axioms for $\succ$ on $\mathscr{H}$ imply the existence of linear u_s on $\mathscr{M}(\mathscr{C})$ for each s such that

$$h \succ h' \quad \text{iff} \quad \sum_S u_s(h(s)) > \sum_S u_s(h'(s))$$

for all $h,h'\in\mathscr{H}$, with the u_s unique up to simultaneous ($u_s \to au_s + b_s$, $a>0$) positive affine transformations. With $\succ$ assumed to be nonempty, s is defined to be null iff $h\sim h'$ whenever $h(t)=h'(t)$ for all $t\neq s$. Preferences on $\mathscr{M}(\mathscr{C})$ are defined on the basis of constant horse lotteries in the usual manner, and the following is assumed for all $h,h'\in\mathscr{H}$ and all $t\in S$:

(1) If $h(s)=h'(s)$ for all $s\in S\backslash\{t\}$, and $h(t)\succsim h'(t)$, then $h\succsim h'$; if, in addition, t is not null and $h(t)\succ h'(t)$ then $h\succ h'$.

This sure-thing principle is similar to those mentioned earlier. It then follows that there is a linear u on $\mathscr{M}(\mathscr{C})$ and a unique probability measure P on 2^S such that $P(s)=0$ iff s is null, and

$$h \succ h' \quad \text{iff} \quad \sum_S P(s)u(h(s)) > \sum_S P(s)u(h'(s))$$

for all $h,h'\in\mathscr{H}$. This representation gives a Savage-type representation under the usual identification conventions.

A closely related theory is presented by Myerson (1979) with S and $\mathscr{C}$ finite. Myerson uses another version of von Neumann-Morgenstern type axioms for $\succ$ on $\mathscr{H}$ to obtain the initial representation in the preceding paragraph. His sure-thing axiom says that if $h(t)=h'(t')$, $h''(t)=h''(t')$, $h(s)=h''(s)$ for all $s\neq t$, and $h'(s)=h''(s)$ for all $s\neq t'$, then $h\succsim h''$ iff $h'\succsim h''$. His axioms yield the representation at the end of the preceding paragraph and imply that if $\succ$ is not empty then $P(s)>0$ for every state.

Fishburn (1970, section 13.2) relaxes the strong structure presumed above by not requiring all consequences to be relevant under every state.

He assumes that S is finite, lets $\mathscr{C}_s$ denote the consequences that can occur if s obtains, and applies $\succ$ to

$$\mathscr{H}' = \{h \text{ on } S\colon\ h(s) \in \mathscr{M}(\mathscr{C}_s) \text{ for each } s \in S\}.$$

Jensen's linear utility axioms for $\succ$ on $\mathscr{H}'$ are presumed to hold, and null states and $\succ$ on constant acts and consequences that can occur under every state are defined in the usual ways. Fishburn assumes also that

(1) There are $c_0, c_1 \in \mathscr{C}_s$ for all s, such that $c_1 \succ c_0$;

(2) If s and t are not null, $h \in \mathscr{H}'$, and x and y are in both $\mathscr{M}(\mathscr{C}_s)$ and $\mathscr{M}(\mathscr{C}_t)$, then ($h$ with $h(s)$ replaced by x) $\succ$ (h with $h(s)$ replaced by y) iff (h with $h(t)$ replaced by x) $\succ$ (h with $h(t)$ replaced by y).

According to (1), there need be no more than two Savage-type constant acts since, apart from c_0 and c_1, the different $\mathscr{C}_s$ can be disjoint. Axiom (2) is a sure-thing/substitution axiom for the $\mathscr{H}'$ context. The cited axioms imply that there is a utility function u on $\cup_S \mathscr{C}_s$ and a probability measure P on 2^S such that $h \succ h'$ iff $\Sigma P(s)u(h(s)) > \Sigma P(s)u(h'(s))$, where $u(h(s)) = \Sigma\, [h(s)(c)]\, u(c)$. In addition, P is unique, $P(s) = 0$ iff s is null, and the restriction of u on the union of the $\mathscr{C}_s$ for nonnull s is unique up to a positive affine transformation.

The standard horse lottery approach is generalized to arbitrary S in Fishburn (1969; 1970, section 13.3), with $\mathscr{C}$ and S arbitrary and $\mathscr{S} = 2^S$. Jensen's axioms are again used for $\succ$ on $\mathscr{H}$ along with a nontriviality axiom and the following sure thing and dominance axioms:

(1) If A is not null, $h(s) = x$ and $h'(s) = y$ for all $s \in A$, and $h_{\bar{A}} = h'_{\bar{A}}$, then $h \succ h'$ iff $x \succ y$;

(2) If $h(s) \succ h'$ for all s, then $h \succsim h'$; if $h' \succ h(s)$ for all s, then $h' \succsim h$.

We then obtain linear u on $\mathscr{M}(\mathscr{C})$ and unique P on $\mathscr{S}$ such that

$$h \succ h' \quad \text{iff} \quad \int_S u(h(s))\, \mathrm{d}P(s) > \int_S u(h'(s))\, \mathrm{d}P(s)$$

for all horse lotteries h and h'. Moreover, every h is bounded, i.e. $P\{s\colon a \leqslant u(h(s)) \leqslant b\} = 1$ for some finite a and b, $P(A) = 0$ iff A is null, and u is bounded if there is a denumerable partition of S such that $P(A) > 0$ for every A in the partition.

Fishburn (1972) generalizes the preceding results by (i) using an arbitrary mixture set $\mathscr{M}$ instead of $\mathscr{M}(\mathscr{C})$, (ii) using an arbitrary Boolean algebra $\mathscr{S}$ of subsets of S in place of 2^S, and (iii) replacing the dominance axiom (2) by the formally weaker axiom that has h' as the constant horse lottery $h'(s) = x$ for all $s \in S$. The representation of the preceding paragraph is obtained, without using axiom (2), for simple measurable horse lotteries. With axiom (2), and calling h u-measurable iff $\{s\colon u(h(s)) \in I\}$ is in $\mathscr{S}$ for every real interval I, it is shown that every u-measurable h is bounded and that the representation holds for all such h. Another axiom for countable additivity (cf. Ferreira, 1972) implies that P is countably additive when $\mathscr{S}$ is presumed to be a σ-algebra.

In a related theory, Fishburn (1975, section 3) considers $\succ$ on $\mathscr{H}$ that is not presumed to be an asymmetric weak order. He uses seven axioms, the last two of which are somewhat awkward, and shows that these axioms imply the existence of linear u on $\mathscr{M}(\mathscr{C})$ and finitely additive P on 2^S such that the one-way representation

$$h \succ h' \Rightarrow \int_S u(h(s))\,\mathrm{d}P(s) > \int_S u(h'(s))\,\mathrm{d}P(s)$$

holds for simple horse lotteries h and h'. This has not been generalized for $h, h' \in \mathscr{H}$ that can assign an infinite number of lotteries in $\mathscr{M}(\mathscr{C})$ to the different states in S.

8.4. *Towards generalized structures*

To conclude this section, we mention three theories that use mixture sets in different ways than the theories discussed above. The basic primitive sets in the first two of these are an act set $\mathscr{F}$ and a state set S. Consequences as used previously are replaced by act-state pairs $fs \in \mathscr{F} \times S$. More generally, we can view the act-event pair $fA \in \mathscr{F} \times \mathscr{S}_0$ as $\{fs\colon\ s \in A\}$ and interpret it as "whatever might happen if you do f and A obtains". This description is purposefully vague in contrast to the precise conceptualization of consequences, although one can always write fs as $f(s)$ and call it a consequence. In fact, consequences in preceding theories can be visualized in this way. However, since the ensuing theories do not depend on constant acts or presume that $\mathscr{F} \times \{s\}$ has entities identical to some elements in $\mathscr{F} \times \{t\}$ when $t \neq s$, and since the notation $u(fs)$ is more suggestive than $u(f(s))$ in conveying the idea that states as well as acts can contribute to utility, I shall use the new notation here.

Fishburn (1973, 1974) applies $\succ$ to $\mathscr{M} \times \mathscr{S}_0$, where $\mathscr{M}$ is an arbitrary mixture set that can best be visualized as the set $\mathscr{M}(\mathscr{F})$ of mixed acts or lotteries on acts, and $\mathscr{S}$ is a Boolean algebra of subsets of S. We interpret $xA \in \mathscr{M} \times \mathscr{S}_0$ as whatever might happen if you do x (which yields act f with probability $x(f)$) and A obtains, and $xA \succ yB$ indicates that the individual would rather 'do' x under the assurance that A obtains than 'do' y under the assurance that B obtains. There may of course be conceptual difficulties in trying to compare act-event pairs that are based on different events (Fishburn, 1974, pp. 27–28), but it is presumed that the individual is able to make such comparisons.

In addition to asymmetric weak order for $\succ$ on $\mathscr{M} \times \mathscr{S}_0$ and a nontriviality condition, the present theory uses the following four axioms for all $x, y, z, w \in \mathscr{M}$ and all $A, B \in \mathscr{S}_0$:

(1) If $xA \sim zB$ and $yA \sim wB$ then $(\frac{1}{2}x + \frac{1}{2}y)A \sim (\frac{1}{2}z + \frac{1}{2}w)B$;

(2) $\{\alpha\colon (\alpha x + (1-\alpha)y)A \succ zB\}$ and $\{\beta\colon zB \succ (\beta x + (1-\beta)y)A\}$ are open subsets of $[0, 1]$;

(3) If $xA \succsim xB$ then $xA \succsim x(A+B) \succsim xB$;

(4) If $A \cap B = \emptyset$ then $xA \succ xB$ and $yB \succ yA$ for some $x, y \in \mathscr{M}$.

Axioms (1) and (2) are generalizations of the Herstein-Milnor axioms (B2) and (B3) respectively, axiom (3) is an averaging condition, and (4) is a special structural condition which says that no event in $\mathscr{S}_0$ is dominated by a disjoint event in $\mathscr{S}_0$ in the sense that $xA \succsim xB$ for all $x \in \mathscr{M}$. Although (4) will sometimes fail in realistic situations, something like it is needed in the absence of things like constant acts to obtain a nice representation for preferences.

The axioms in the preceding paragraph imply u on $\mathscr{M} \times \mathscr{S}_0$ that is linear in its first component – $u(\lambda x + (1-\lambda)y, A) = \lambda u(xA) + (1-\lambda)u(yA)$ – and nonnegative real numbers $P(A|A+B)$ and $P(B|A+B)$ that sum to 1 for each pair of disjoint A and B in $\mathscr{S}_0$, such that

$$xA \succ yB \quad \text{iff} \quad u(xA) > u(yB)$$

and

$$u(x, A+B) = P(A|A+B)u(xA) + P(B|A+B)u(xB)$$

for all $x, y \in \mathscr{M}$ and all $A, B \in \mathscr{S}_0$, with P unique and u unique up to a positive affine transformation. However, P does not necessarily behave like a probability measure, and in order to guarantee such behaviour we need to use another structural condition:

(5) If $A, B, C \in \mathscr{S}_0$ are mutually disjoint and $xA \sim xB$ for some $x \in \mathscr{M}$, then exactly two of yA, yB and yC are indifferent for some $y \in \mathscr{M}$.

If this 'linear independence' axiom fails then, as shown in Fishburn (1973, 1974), the unique P can be nonadditive. When (5) holds, it follows that each $P(\ |A)$ for $A \in \mathscr{S}_0$ is a finitely additive probability measure on $\{A \cap B: B \in \mathscr{S}\}$, that $P(A|C) = P(A|B)P(B|C)$ whenever $A \subseteq B \subseteq C$ and $A, B, C \in \mathscr{S}_0$, and that if $\{A_1, \ldots, A_n\}$ is a measurable partition of $A \in \mathscr{S}_0$, then

$$u(xA) = \sum_{i=1}^{n} P(A_i|A)u(xA_i).$$

Fishburn (1973, 1974) uses the dominance axiom

(6) If $xs \succ yB$ for all $s \in A$, then $xA \succsim yB$;

if $xA \succ ys$ for all $s \in B$, then $xA \succsim yB$,

under the assumption that $\{s\} \in \mathscr{S}$ for every $s \in S$, to extend the preceding form to

$$u(xA) = \int_A u(xs)\mathrm{d}P(s|A)$$

for all xA that are u-measurable and bounded. The extended form is the same as

$$u(xA) = \frac{1}{P(A)} \int_A u(xs)\,\mathrm{d}P(s),$$

provided that $P(A) = P(A|S)$ is positive. A number of criticisms of the theory just outlined as well as those mentioned in the next two paragraphs have been offered by Pratt (1974).

Balch (1974) proposes an alternative to the theory just presented that applies $\succ$ to $\mathscr{M}(\mathscr{F} \times \mathscr{S}_0)$ rather than to $\mathscr{M}(\mathscr{F}) \times \mathscr{S}_0$. This simplifies several technical aspects of the preceding theory but introduces mixtures $\lambda(xA) + (1-\lambda)(yB)$ based on different events that Fishburn sought to avoid. Balch notes that axioms similar to weak order and (1) through (4) above lead to $u(\lambda(xA) + (1-\lambda)(yB)) = \lambda u(xA) + (1-\lambda)u(xB)$ plus a representation like that stated prior to (5), and shows that, when $x(A+B) \sim \lambda(xA) + (1-\lambda)(xB)$, $\lambda = P(A|A+B)$.

Finally, Balch and Fishburn (1974) outline an approach to a conditional theory of subjective expected utility that begins with a primitive set $\mathscr{F}$ of acts and a state set S_f for each $f \in \mathscr{F}$, with $\mathscr{S}_f$ a Boolean algebra of subsets of S_f for each f. This theory was motivated by the Luce-Krantz theory discussed in the next section and by the practical difficulty of constructing a manageable set S of Savage states in many situations (cf. $\mathscr{C}^{\mathscr{F}}$ in section 2). The tailor-made state set S_f for each act precludes the powerful structure enjoyed by Savage and others, and leads to strong structural conditions in the Balch-Fishburn theory that I shall not recount here.[10] Their theory applies $\succ$ to $\mathscr{M}(\mathscr{C}^*)$, where $\mathscr{C}^*$ is the set of all act-event pairs fA for which $f \in \mathscr{F}$ and $A \in \mathscr{S}_f \backslash \{\emptyset\}$. Their axioms include the Herstein-Milnor axioms for $\succ$ on $\mathscr{M}(\mathscr{C}^*)$, an averaging condition like (3) above for each f, and the special structural conditions. With $u(fA)$ defined from the order-preserving linear u on $\mathscr{M}(\mathscr{C}^*)$ in the usual way, the representation includes

$$u(f, A+B) = P_f(A|A+B)u(fA) + P_f(B|A+B)u(fB)$$

for disjoint A and B in $\mathscr{S}_f \backslash \{\emptyset\}$. The P_f have the usual probability properties. Because states and events are conditioned on acts, each act carries its own event probabilities.

9. LUCE AND KRANTZ

The theory of Luce and Krantz (1971), which is discussed also in Krantz *et al.* (1971, Chapter 8), was designed to provide a qualitative theory of conditional decisions that admits a conditional subjective expected utility representation and enjoys a level of generality comparable to Savage's (1954) unconditional theory. Its basic primitive sets are S and $\mathscr{C}$; constructed primitive sets are a Boolean algebra $\mathscr{S}$ of subsets of S, a subset $\mathscr{N} \subseteq \mathscr{S}$ of null events that contains $\emptyset$, and a set $\mathscr{G}$ of conditional acts $f_A, g_B, \ldots$ that map nonnull events $A, B, \ldots$ in $\mathscr{S}$ into $\mathscr{C}$. I shall let $f_A + g_B$ denote the function from $A + B$ into $\mathscr{C}$ that equals f_A on A and g_B on B.

Luce and Krantz apply $\succ$ to $\mathscr{G}$, and assume that $\mathscr{G}$ is closed under disjoint unions and nonnull restrictions, so that, for all $A, B \in \mathscr{S} \setminus \mathscr{N}$ and all $f_A, g_B \in \mathscr{G}$,

(1) $f_A + g_B \in \mathscr{G}$; if $B \subseteq A$ then the restriction of f_A to B is in $\mathscr{G}$.

As with $A + B$, $f_A + g_B$ always signifies that $A \cap B = \emptyset$.

Their basic representation involves u on the *acts* in $\mathscr{G}$ (unique up to a positive affine transformation) and a finitely additive probability measure P on $\mathscr{S}$ (unique) such that u preserves $\succ$, $P(A) = 0$ iff $A \in \mathscr{N}$, and, for all $f_A, g_B \in \mathscr{G}$,

$$u(f_A + g_B) = P(A|A + B)u(f_A) + P(B|A + B)u(g_B).$$

The authors view at least some acts in $\mathscr{G}$ as natural objects of choice, while others – formed from disjoint unions and restrictions of natural acts – are artificial. Depending on what one views as the natural acts, the + in $f_A + g_B$ has several interpretations, as we shall discuss shortly. Unlike Balch and Fishburn (1974), all acts conditioned on A have the same $P(\ |A)$ values on subevents in A. Thus, at least for natural acts, the individual's choice can delimit the states that are relevant to that act but cannot influence the conditional probabilities in the relevant event. In addition, the Luce-Krantz state set S differs from Savage's. This difference along with other important points is best illustrated by an example.

9.1. *An example*

An individual plans to travel from New York to Boston by either air (A) or bus (B). Several carrier options are available for each mode of travel. If he goes by air – act f_A^i for airline i – then the states relevant to this mode and over which he has no control comprise event A. If he goes by bus – act g_B^j for bus line j – then the states relevant to the bus mode and over which he has no control comprise event B. We shall assume that $A \cap B = \emptyset$ and that there are interdependencies between A and B due to factors such as the weather. One state in each of A and B will obtain, but only one of the two obtaining states will be relevant to the mode he actually chooses. The natural acts are the f_A^i and g_B^j. The composite act $f_A^i + g_B^j$, which indicates that he will either take airline i or bus line j, is an 'artificial' act.

The Luce-Krantz state set is $S = A + B$. Savage's state set would be $A \times B$:

one and only one Savage state $(s_A, s_B) \in A \times B$ will obtain. If P^* is Savage's probability measure, and if singleton events are in $\mathscr{S}$ for Luce and Krantz, then their conditional probabilities will be

$$P(s_A|A) = \sum_B P^*(s_A, s_B),$$

$$P(s_B|B) = \sum_A P^*(s_A, s_B).$$

Interdependencies between A and B preclude calculation of P^* on the basis of $P(\ |A)$ and $P(\ |B)$.

Suppose for the moment that $\mathscr{S} = \{\emptyset, A, B, S\}$ with $A + B = S$. Assuming that A and B are nonnull, the Luce-Krantz representation uses only $P(A) = P(A|S)$ and $P(B) = P(B|S) = 1 - P(A)$ in addition to $P(\emptyset) = 0$ and $P(S) = 1$, and both $P(A)$ and $P(B)$ are positive. The representation written above gives

$$u(f_A^i + g_B^j) = P(A)u(f_A^i) + P(B)u(g_B^j),$$

and the axioms of their theory imply that $f_A^i \succ g_B^j$ for some i and j, and $g_B^k \succ f_A^m$ for some k and m. Hence, it makes no sense to view $P(A)$ as the probability that the individual will choose the A mode independent of i and j when he is confronted with $f_A^i + g_B^j$. Even if there were only one available airline and only one available bus line, and we assume that $f_A \nsim g_B$, it still makes little sense to view $P(A)$ as the probability he will fly, given an open choice of either flying or riding, since then $P(A)$ would presumably be either 0 or 1 (Balch, 1974). Moreover, $P(A)$ is not the probability that some $s_A \in A$ obtains independent of what the individual does, for this probability is unity, as is the probability that some $s_B \in B$ obtains.

Luce and Krantz favor the interpretation that $P(A)$ is the individual's subjective probability (not necessarily 1/2) for an extraneous event E, with $P(B)$ his probability for $\bar{E}$. Accordingly, $f_A^i + g_B^j$ is viewed as an act in which he 'gets' f_A^i if E and g_B^j if $\bar{E}$, much as we interpreted Ramsey acts in section 5. We can think of E as an event that lies behind the Luce-Krantz system that is available for structural-scaling purposes although it (or more general extraneous event sets) does not appear explicitly in their axioms. Thus, we shall think of $P(A)$ in this way when f_A and g_B are natural acts although it has no particular relationship to Savage's P^*.

To consider restrictions of natural acts, and disjoint unions formed from such restrictions, we partition A into its subset A_0 of 'foul weather' states and its subset A_1 of 'fair weather' states, and let f_0^i and f_1^i be respectively the restrictions of f_A^i to A_0 and to A_1. Similar notations apply to B. Assuming that these new events are in $\mathscr{S}$, we then encounter the following types of expressions in the representation:

$$\begin{aligned} u(f_0^i + f_1^m) &= P(A_0|A)u(f_0^i) + P(A_1|A)u(f_1^m), \\ u(f_0^i + g_0^j) &= P(A_0|A_0 + B_0)u(f_0^i) + P(B_0|A_0 + B_0)u(g_0^j), \\ u(f_0^i + g_1^j) &= P(A_0|A_0 + B_1)u(f_0^i) + P(B_1|A_0 + B_1)u(g_1^j). \end{aligned}$$

In the first expression, $f_0^i + f_1^m$ says "fly on airline i in foul weather and on airline m in fair weather", and $P(A_0|A)$ is the probability of 'foul weather'. The latter conclusion assumes that E is independent of A and B, and we maintain this assumption below.

In the second expression, $f_0^i + g_0^j$ says "fly i if E, take bus line j if $\bar{E}$, given foul weather", with $P(A_0|A_0 + B_0) = Pr(E)$. Act $f_0^i + g_1^j$ in the third expression is more complex. It says "fly i if E and go by bus line j if $\bar{E}$, given foul weather in the first instance and fair weather in the second instance". On breaking down the conditional probability $P(A_0|A_0 + B_1)$, we obtain

$$\begin{aligned} P(A_0|A_0 + B_1) &= \frac{P(A_0)}{P(A_0) + P(B_1)} \\ &= \frac{P(A_0|A)P(A|S)}{P(A_0|A)P(A|S) + P(B_1|B)P(B|S)} \\ &= \frac{Pr(\text{'foul'})Pr(E)}{Pr(\text{'foul'})Pr(E) + Pr(\text{'fair'})Pr(\bar{E})} \\ &= \frac{Pr(\text{'foul'})Pr(E)}{Pr(E)[2Pr(\text{'foul'}) - 1] + 1 - Pr(\text{'foul'})} \end{aligned}$$

If $Pr(E) = \frac{1}{2}$, then $P(A_0|A_0 + B_1) = Pr(\text{'foul'})$; if $Pr(\text{'foul'}) = \frac{1}{2}$, then $P(A_0|A_0 + B_1) = Pr(E)$. In other cases, $P(A_0|A_0 + B_1)$ depends both on E and the state of the weather.

9.2. *Axioms*

Luce and Krantz use nine axioms in their basic system. These include axiom (1), weak order, an axiom for $\mathscr{N}$, an Archimedean axiom like those used by Krantz *et al.* (1971) for additive measurement, and an independence axiom for nonnull disjoint A and B which uses $+$ to imply that if the preference 'difference' between f_A^1 and f_A^2 is as large as that between f_A^3 and f_A^4, and if $g_B^i \sim f_A^i$ for $i = 1, 2, 3, 4$, then the preference 'difference' between g_B^1 and g_B^2 is as large as that between g_B^3 and g_B^4. In the last of these, $f_A^1 + h_B \succsim f_A^2 + h'_B$ and $f_A^3 + h_B \sim f_A^4 + h'_B$ indicate that the preference 'difference' between f_A^1 and f_A^2 is as large as that between f_A^3 and f_A^4. In terms of their representation, the stated ($\succsim$, $\sim$) pair translates into

$$\rho[u(f_A^1) - u(f_A^2)] \geqslant (1-\rho)[u(h'_B) - u(h_B)] =$$
$$= \rho[u(f_A^3) - u(f_A^4)],$$

where $\rho = P(A|A+B)$.

Their other four basic axioms are, for all $A, B \in \mathscr{S}\setminus\mathscr{N}$ and all $f_A, f'_A, f''_A, f_{A+B}, g_B \in \mathscr{G}$:

(2) $\mathscr{S}\setminus\mathscr{N}$ contains at least three mutually disjoint events; $\succ$ is not empty;

(3) $h_A \sim g_B$ for some $h_A \in \mathscr{G}$; if $f'_A + g_B \succsim f_{A+B} \succsim f''_A + g_B$, then $h_A + g_B \sim f_{A+B}$ for some $h_A \in \mathscr{G}$;

(4) If $f_A \sim g_B$ then $f_A + g_B \sim f_A$;

(5) $f'_A \succ f''_A$ iff $f'_A + g_B \succ f''_A + g_B$.

Axiom (4) is an obviously necessary averaging condition, and axiom (5) is a necessary sure-thing (additivity, cancellation) axiom. Unlike (4), (5) relies on the presumption that the probability of A given $A + B$ does not depend on the particular A-conditioned act under consideration.

Axioms (1), (2) and (3) are the nonnecessary structural conditions used by Luce and Krantz. Axiom (2) is their nontriviality condition. Its first part can be weakened to posit only two nonnull disjoint events (as in our example) if they strengthen their independence assumptions after the fashion of Debreu's (1959) axiom (1) in section 5 or Suppes's (1956) axiom (1) in section 7. The first part of axiom (3) implies that u has the same range for acts conditioned on each nonnull event, and is therefore quite strong. The second part of (3) is an existential intermediate-value axiom that involves P as well as u. Although Luce and Krantz acknowledge the strength of their structural conditions, they show that these conditions are less demanding than those used by Savage (1954).

It is proved in Krantz *et al.* (1971, Chapter 8) that the nine axioms in the basic Luce-Krantz system imply the representation set forth in the paragraph after (1) above.

9.3. *Utilities for consequences*

The basic Luce-Krantz representation has u on $\mathscr{G}$ only and not on $\mathscr{C}$. Two more axioms, based on constant conditional acts, are then used to obtain the representation

$$f_A \succ g_B \quad \text{iff} \quad \frac{1}{P(A)} \int_A u(f_A(s))\,\mathrm{d}P(s) > \frac{1}{P(B)} \int_B u(g_B(s))\,\mathrm{d}P(s),$$

for all simple nonnull-measurable acts in $\mathscr{G}$. With c_A the function from A into $\mathscr{C}$ that assigns c to every $s \in A$, the new axioms are: for all $c \in \mathscr{C}$,

(6) $c_A \in \mathscr{G}$ for some $A \in \mathscr{S}\backslash\mathscr{N}$;

(7) If $A, B \in \mathscr{S}\backslash\mathscr{N}$ and $c_A, c_B \in \mathscr{G}$, then $c_A \sim c_B$.

Structural axiom (6) invokes the presence of some – but by no means all – constant acts. Axiom (7), which is implied by the new representation, asserts that the consequences embody all valued aspects of the situation.

In section 6, I discussed Narens's second model using notation that is similar to the notation in the preceding paragraph. However, we have seen that the formulations of Narens (1976) and Luce and Krantz (1971) are substantially different, and I hope that the notational similarities will not blur these differences.

9.4. *Extensive measurement*

Luce (1972) shows how the Luce-Krantz theory can be modified to accommodate an extensive measurement operation $*$ on conditional acts or on consequences. Like in Roberts's (1974) theory at the end of section 6, if * is initially defined on $\mathscr{C}$, it is extended to $\mathscr{G}$ in a natural way. In Luce's case, we define $f_A * g_B$ by

$$(f_A * g_B)(s) = \begin{cases} f_A(s) & \text{if} \quad s \in A \setminus B \\ f_A(s) * g_B(s) & \text{if} \quad s \in A \cap B \\ g_B(s) & \text{if} \quad s \in B \setminus A, \end{cases}$$

with $c*d$ interpreted as 'both c and d'. It is assumed that $\mathscr{G}$ is closed under * and under nonnull restrictions. This replaces axiom (1), with $f_A * g_B$ now used instead of $f_A + g_B$ as previously interpreted. Luce states ten axioms that he feels are somewhat easier to understand than the Luce-Krantz axioms, and shows that they imply an order-preserving u on $\mathscr{G}$ (unique up to a similarity transformation) and a unique probability measure P on $\mathscr{S}$ with $P(A) = 0$ iff $A \in \mathscr{N}$ such that, for all $f_A, g_B \in \mathscr{G}$,

$$u(f_A * g_B) = u(f_{A \setminus B})P(A \setminus B | A \cup B) + u(g_{B \setminus A})P(B \setminus A | A \cup B) + \\ + [u(f_{A \cap B}) + u(g_{A \cap B})]P(A \cap B | A \cup B).$$

In this representation, $f_{A \setminus B}$ is the restriction of f_A to $A \setminus B$, and so forth, with $u(f_C)P(C|D) = 0$ whenever $C \in \mathscr{N}$.

10. MONO-SET THEORIES

The final representation type that we shall consider was first discussed extensively by Jeffrey (1965a). Its initial axiomatization was given by Bolker (1967) on the basis of mathematics developed in Bolker (1966). Recently, Jeffrey (1978) modified Bolker's axioms to accommodate null events, and Domotor (1978) axiomatized a finite S version of the mono-set representation.

Mono-set theories apply $\succ$ to part of a Boolean algebra $\mathscr{S}$, which I shall view as a set of events from some S. An ideal $\mathscr{N}$ of null events may be presumed. Events in $\mathscr{S}$ are both uncertain – except for null events and their complements – and valued, so that subjective probabilities and utilities apply to the same entities. The basic mono-set representation consists of real valued functions M and P on $\mathscr{S}$ such that

(i) M is a signed measure, with values unrestricted in sign and $M(A+B) = M(A) + M(B)$;

(ii) P is a probability measure with $P(A) = 0$ if and only if $A \in \mathscr{N}$;

(iii) for all $A, B \in \mathscr{S} \setminus \mathscr{N}$,

$$A \succ B \quad \text{iff} \quad \frac{M(A)}{P(A)} > \frac{M(B)}{P(B)}.$$

Viewing $U(A) = M(A)/P(A)$ as the conditional subjective expected utility of event A, additivity gives

$$U(A+B) = P(A|A+B)U(A) + P(B|A+B)U(B).$$

Under suitable regularity conditions, the Radon-Nikodym Theorem yields $u = \mathrm{d}M/\mathrm{d}P$ with $M(A) = \int_A u(s)\mathrm{d}P(s)$, hence

$$U(A) = \frac{1}{P(A)} \int_A u(s)\mathrm{d}P(s).$$

Thus the mono-set representation bears at least a superficial resemblance to the conditional theory of Luce and Krantz (1971). The main conceptual difference between the two is the distinction between consequences and states in the Luce-Krantz theory and their explicit formation of conditional acts.

10.1. *Interpretations*[11]

Each $s \in S$ in a mono-set theory ideally specifies all aspects of the situation that are of concern to the individual, including acts, consequences, Savage states, and so forth. It is understood that, whatever happens, the individual believes that one and only one $s \in S$ will 'obtain'. Although this approach blurs the often useful distinctions among acts, consequences and other entities that appear in other theories (Bolker, 1967, p. 335), Jeffrey (1965a, 1965b, 1974) argues that it is somehow unnatural to break things up in a Savage manner, and finds the holistic mono-set viewpoint more appealing.

According to Jeffrey, an act is any event $A \in \mathscr{S}$ that the individual thinks he has the power to make happen. Alternatively, an act is a collection of states that covers all possibilities under one or more specific courses of action, where the individual believes it within his power to follow some course of action thus specified. We might visualize atomic acts as the 'minimal' events that the individual feels he can bring about (e.g., all states in which he carries his umbrella, or continues smoking, etc.), and composite acts – including S if it is not the only atomic act – as unions of two or more atomic acts.

Bolker and Jeffrey provide several related interpretations of $A \succ B$. One is that you would prefer to hear that some member of A will occur than to hear that some member of B will occur. Another is that $A \succ B$ "means that a random member of A is preferred to a random member of B, where 'random' means 'selected from A, or B, according to our subject's own estimate, conscious or not, of the actual conditional probabilities'" (Bolker, 1967, p. 335). While this seems reasonable for some events, for other events it could pose problems that relate to our discussion of Luce and Krantz (1971) and to Balch's (1974) criticism of their closure axiom. For example, if A and B are singleton states, or if A and B are atomic acts, then $A \succ B$ is easily visualized. However, when composite acts are involved, matters may be confused by the individual's ability and desire to make happen some subevent in the composite act. In Bolker's theory, $A \succ B$ requires $A \succ A + B \succ B$, but if one can make A happen then why wouldn't one have $A \sim A + B$? In other words, why wouldn't one have $P(A|A+B) = 1$ instead of $P(A|A+B) < 1$ (as required by Bolker) when A is an atomic act and $A \succ B$?

Luce and Krantz avoid this type of problem by invoking extraneous events for composites like $f_A + g_B$ when f_A and g_B are natural acts, but Bolker and Jeffrey do not advocate a similar approach. Hence, when A and B are disjoint Bolker-Jeffrey atomic acts, $P(A|A+B)$ has a different interpretation than was used in the preceding section: presumably, it is the individual's subjective probability that he would do A, given that he would do either A or B. Thus, as long as $0 < P(A|A+B) < 1$ when A is not indifferent to B, there is a strong suggestion of probabilistic choice in the theory. While I see nothing objectionable *per se* in this, it does tend to differentiate mono-set theories from those discussed in previous sections.

10.2. *Bolker's theory*

Bolker's (1967) event set $\mathscr{S}$ is a complete (closed under arbitrary unions and intersections) atom free Boolean algebra, and hence a σ-algebra. His null event set is effectively $\{\emptyset\}$, and $\succ$ is applied to $\mathscr{S}_0$.

Bolker assumes that $\succ$ is a continuous asymmetric weak order. With all designated events in $\mathscr{S}_0$, continuity in the present setting says that if $A_n \uparrow A$ (hence $\cup A_n = A$) or $A_n \downarrow A$ (hence $\cap A_n = A$), and if $B \succ A \succ C$, then $B \succ A_n \succ C$ for all large n. Because of the powerful structure adopted by Bolker, he needs only two more axioms:

(1) If $A \succ B$ then $A \succ A + B \succ B$; if $A \sim B$ then $A \sim A + B \sim B$;

(2) If A, B and C are mutually disjoint, $A \sim B \nsim C$, and $A + C \sim B + C$, then $A + D \sim B + D$ for all D disjoint from A and B.

Axiom (1) is an averaging axiom, and (2), which Bolker terms *impartiality*, is an independence condition that is related to axiom (5) in the preceding section. Bolker's representation and the hypotheses of (2) imply $P(A) = P(B)$, which requires $A + D \sim B + D$ to avoid a contradiction. To avoid the trivial case, we shall assume also that $\succ$ is not empty.

The representation implied by these axioms consists of (i) through (iii) above, plus: P and M are countably additive; $\mathscr{N} = \{\emptyset\}$ so that $\mathscr{S} \setminus \mathscr{N} = \mathscr{S}_0$; and $M(A) = \int_A u(s)\,\mathrm{d}P(s)$ with $u = \mathrm{d}M/\mathrm{d}P$ for each $A \in \mathscr{S}_0$.

Because of the mono-set structure, Bolker's uniqueness theorem differs from those in preceding theories. When (M, P) satisfies Bolker's representation, so does (M', P') if and only if there are real numbers a, b, c, d such that $-d/c$ is not in the interval of values $\{M(A)/P(A)\colon A \in \mathscr{S}_0\}$; $ad > bc$; $cM(S) + d = 1$; $M'(A) = aM(A) + bP(A)$ for all $A \in \mathscr{S}_0$; $P'(A) = cM(A) + \mathrm{d}P(A)$ for all $A \in \mathscr{S}_0$.

The first condition requires $c = 0$ if $\{M(A)/P(A)\colon A \in \mathscr{S}_0\}$ is the entire real line, and in this case $d = 1$ with P unique and $U = M/P$ unique up to a positive ($a > 0$) linear transformation. More generally, with $U' = M'/P'$, the foregoing gives

$$U'(A) = \frac{aU(A) + b}{cU(A) + d}$$

and U is said to be unique up to a *fractional linear transformation* subject to $-d/c \notin \{U(A)\colon A \in \mathscr{S}_0\}$ – so that the denominator does not vanish, and $ad > bc$ – so that U' increases as U increases.

10.3. *Jeffrey's theory*

Jeffrey (1978) modifies Bolker's theory in a manner consistent with Jeffrey (1965). A main structural difference between the two theories lies in the nature of $\mathscr{S}$ and its null event set $\mathscr{N}$. Jeffrey assumes that $\mathscr{S}$ is a σ-algebra, $\mathscr{N}$ is a proper σ-ideal of $\mathscr{S}$, every 'disjoint subset' of $\mathscr{S}\setminus\mathscr{N}$ is countable, and every event in $\mathscr{S}\setminus\mathscr{N}$ is the union of two disjoint events in $\mathscr{S}\setminus\mathscr{N}$. Jeffrey applies $\succ$ to all of $\mathscr{S}$ and assumes that $\succ$ is an asymmetric weak order with $A \succ S \succ \bar{A}$ for some A (nontriviality), $S \sim \emptyset$, and $A \sim B$ whenever $(A\setminus B) \cup (B\setminus A) \in \mathscr{N}$. His other three axioms are very similar to Bolker's continuity, averaging (1) and impartiality (2) axioms.

Jeffrey's representation and uniqueness theorems are essentially the same as Bolker's with the exception of his treatment of null events. Jeffrey has

$$A \succ B \quad \text{iff} \quad U(A) > U(B), \text{for } all\ A, B \in \mathscr{S},$$

with

$$U(A) = \begin{cases} \dfrac{1}{P(A)} \displaystyle\int_A u(s)\,\mathrm{d}P(s) & \text{if} \quad P(A) > 0 \\ U(S) & \text{if} \quad P(A) = 0. \end{cases}$$

In addition, $P(A) = 0$ iff $A \in \mathscr{N}$.

10.4. *A finite-states version*

Domotor (1978) provides necessary and sufficient conditions for a Bolker-type representation when $\mathscr{S}$ is a finite Boolean algebra. He then considers a nonstandard representation for arbitrary Boolean algebras. I shall comment only on the finite version.

With S finite, Domotor assumes that $\mathscr{S}$ is a Boolean algebra with $\succ$ an asymmetric binary relation on $\mathscr{S}_0$. He then adopts a condition referred to as *projectivity*, which is a sort of super-independence axiom. In terms of probabilities and utilities, projectivity essentially says that if

$$\sum_{i=1}^{n} P(A_i)P(B_i)P(C_i)P(D_i)\,[U(B_i) - U(A_i)]\,[U(D_i) - U(C_i)] = 0,$$

and if $U(B_i) \geqslant U(A_i)$ for $i = 1, \ldots, n$ and $U(D_i) \geqslant U(C_i)$ for $i = 1, \ldots, n-1$, then $U(C_n) \geqslant U(D_n)$. Like Richter's (1975) axiom, projectivity has little intuitive appeal although it illuminates the mathematical structure.

Domotor subsequently proves that the conditions of the preceding paragraph are necessary and sufficient for the existence of a utility function U on $\mathscr{S}_0$ and a probability measure P on $\mathscr{S}$ with P positive on $\mathscr{S}_0$ such that, for all nonnull events (as a primitive, the null events form a proper ideal) $A, B \in \mathscr{S}$,

$$A \succ B \quad \text{iff} \quad U(A) > U(B),$$
$$U(A+B) = P(A|A+B)U(A) + P(B|A+B)U(B).$$

A modest variant of the axioms that explicitly involves $\mathscr{N}$ replaces positive P on $\mathscr{S}_0$ by $P(A)=0$ iff $A \in \mathscr{N}$. Because of finiteness, the strong uniqueness results obtained by Bolker do not apply to the finite case.

11. SUMMARY

As we have seen, theories of subjective expected utility exhibit both differences and similarities in their primitives, axioms, and representation and uniqueness theorems. All theories apply a preference relation $\succ$ or a preference-indifference relation $\succsim$ to a set of entities often referred to as acts, and several theories use more than one such relation. There is always an ordering axiom and one or more independence or averaging conditions that serve to separate utility from subjective probability. Most theories also use an Archimedean or continuity condition to ensure that the representing functions will be real valued. Other prominent axiom types involve nontriviality conditions, additivity postulates, and dominance assertions.

Virtually by definition, a representation for subjective expected utility involves a utility function and a probability function that combine in an expectational form. In most but not all cases the probability function is a finitely additive measure, and several theories use more than one measure conditioned on different events. In all but the mono-set theories, utility and probability are defined on different sets.

Many of the more structured multi-set theories give rise to unique subjective probabilities and utilities that are unique up to a positive affine transformation: uniqueness of utilities up to a similarity transformation can arise when a binary operation is used or when the utility of a worst consequence is set at zero. Less restrictive conditions on probabilities and utilities arise in some more loosely-structured theories and when the preference relation is not a weak order. Mono-set theories have slightly different uniqueness characteristics.

Differences among theories arise mainly from different structures to which the preference relations are applied, but can also arise within the same structure owing to different preference axioms such as weak order versus partial order. New theories have been motivated in part by earlier structural assumptions that are perceived to be too inflexible to accommodate realistic aspects of decision making under uncertainty, or that create interpretational problems. However, many of the newer formulations, including those based on horse lotteries, combinations of conditional acts, and mono-set structures, have introduced new interpretational difficulties.

Table I presents a brief and approximate summary of aspects of the theories discussed in sections 5 through 10. The table provides one expression for the act space used in each theory, along with notes on special features and uniqueness. In the final column, 'usual' means that P is unique and u is unique up to a positive affine transformation, 'u similarity' means that u is unique up to a similarity transformation, and 'arbitrary' indicates something less restrictive than the usual type of uniqueness. I have not stated in the table that all theories in section 8 use something like extraneous scaling probabilities along with one or more applications of von Neumann-Morgenstern type axioms.

I shall conclude with a few opinions on the question of which theory or theories are most satisfactory for a general normative treatment of decision making under uncertainty. My primary criteria in this regard are ability to characterize a diversity of situations in a realistic format, simplicity and intuitive appeal of the preference axioms, interpretability of structural conditions, and nice representation-uniqueness features that can be easily connected to methods for assessing utilities and subjective probabilities.

The following theories seem unsatisfactory as general treatments on the basis of the foregoing criteria: all in section 5 (restricted act spaces); in section 6, Richter (1975) (restricted acts, loose structure, nonintuitive axiom), Stigum (1972) (specialized structure, axioms), Narens's (1976) first theory (restricted acts), Roberts (1974) (extensive type structure); in section 7, Fishburn (1967) (specialized act space); in section 8, Anscombe and Aumann (1963) (finite S, use of two relations), Ferreira (1972) (use of two relations – but fairly appealing), Fishburn (1972), Myerson (1979) and Fishburn (1970) for finite restrictions, and Fishburn (1975), Balch (1974) and Balch and Fishburn (1974) for nonintuitive axioms or interpretational problems; in section 9, Luce (1972) (extensive type structure); in section 10, Domotor (1978) (finite S, nonintuitive axiom).

Among the other theories, I feel that all have shortcomings but that these

TABLE I

Section & Source	Act Space[a]	Restrictions on Sets	Special Features	Uniqueness Status
5. Ramsey (1931)	$\mathscr{C}^3 \times \mathscr{S}^2$	$\mathscr{C}$ infinite	ethically neutral E	usual
Debreu (1959)	$\mathscr{C}^2$	$\mathscr{C}$ infinite	implicit E, topology	usual for u
Davidson & Suppes (1957)	$\mathscr{C}^4 \times \mathscr{S}$	$\mathscr{C}$ finite S finite	E, two relations, equally-spaced u, P nonadditive	usual for u, P unique if $\lvert\mathscr{C}\rvert \geqslant 5$
Pfanzagl (1968)	$\mathscr{C}^4 \times \mathscr{S}^2$	$\mathscr{C}$ is interval	$\mathscr{N}$, conditional acts, certainty equivalents	usual
6. Savage (1954)	$\mathscr{C}^S$	S uncountable	bounded u, continuously divisible S	usual
Richter (1975)	in $\mathscr{C}^S$	$\mathscr{C}$ finite S finite	partial order, polynomial ring	arbitrary
Stigum (1972)	$\mathscr{C}^S$	$\mathscr{C}$ vector space S finite	continuous concave u, topology	usual
Narens (1976)	$\mathscr{C} \times \mathscr{S}_0$		tradeoff structure	P unique &
Narens (1976)	in $\mathscr{G}$		tradeoff structure	u similarity
Roberts (1974)	in $\mathscr{G}$		binary operation	like Narens
7. Suppes (1956)	in $\mathscr{C}^S \times \mathscr{C}^S$	$\mathscr{C}$ infinite	implicit E	usual for u but not P
Fishburn (1967)	$\mathscr{C}^S \times \mathscr{C}^S$	$\mathscr{C}$ infinite S finite	implicit E, topology	usual
8. Anscombe & Aumann (1963)	$\mathscr{M}(\mathscr{H})$ & $\mathscr{M}(\mathscr{C})$	S finite	two relations, u bounded	usual
Ferreira (1972)	$\mathscr{M}(\mathscr{H})$ & $\mathscr{M}(\mathscr{C})$		two relations, countable additivity	usual
Pratt *et al.* (1964, 1965)	$\mathscr{C}^{S \times [0,1]^2}$	S finite in (1964)	$\succsim$ not complete, u bounded	usual

Table 1 (*continued*)

Section & Source	Act Space[a]	Restrictions on Sets	Special Features	Uniqueness Status
Fishburn (1967)	$\mathscr{H}$	S finite		usual
Myerson (1979)	$\mathscr{H}$	$\mathscr{C}$ finite S finite	positive state probabilities	usual
Fishburn (1970)	in $\mathscr{H}$	S finite	different consequences for different states	P unique, u partly affine
Fishburn (1969, 1972)	$\mathscr{H}$ or like $\mathscr{H}$		general theory	usual
Fishburn (1975)	$\mathscr{H}$		partial order	arbitrary
Fishburn (1973, 1974)	$\mathscr{M} \times \mathscr{S}_0$		act-event pairs, conditional flavor	usual
Balch (1974)	$\mathscr{M}(\mathscr{F} \times \mathscr{S}_0)$			usual
Balch & Fishburn (1974)	$\mathscr{M}(\mathscr{C}^*)$		states for each act, special structures	usual (many P's)
9. Luce & Krantz (1971)	in $\mathscr{G}$		conditional acts, special combinations	usual
Luce (1972)	in $\mathscr{G}$		binary operation, conditional theory	P unique & u similarity
10. Bolker (1967)	$\mathscr{S}_0$	S infinite	mono-set, no atoms, countable additivity	fractional linear trans.
Jeffrey (1978)	$\mathscr{S} \setminus \mathscr{N}$	S infinite	mono-set, null events, countable add.	same as Bolker
Domotor (1978)	$\mathscr{S}_0$	S finite	mono-set, axiom of projectivity	arbitrary

[a] $\mathscr{C}$ (consequences), S (states), $\mathscr{S}$ (subsets of S), $\mathscr{S}_0 = \mathscr{S} \setminus \{\emptyset\}$, $\mathscr{N}$ (null events), $\mathscr{F}$ (primitive acts), $\mathscr{G} = \{\mathscr{C}^A : A \in \mathscr{S}_0\}$, $\mathscr{M}$ (mixtures, $\mathscr{H} = [\mathscr{M}(\mathscr{C})]^S$.

differ in degree. My main objection to Bolker (1967) and Jeffrey (1978) is the difficulty in sorting out the decisional aspects within their mono-set format. For this reason, I find some of the others more attractive.

Savage's (1954) theory is suitable for a wide variety of situations, its axioms are elegant and intuitively sensible, and its representation-uniqueness result is easily connected to assessment techniques. Despite its use of constant and other specialized acts, and its implication of continuously divisible events, I regard it as one of the best. Narens's (1976) second theory has a similar appeal, and I would rate Suppes (1956), which replaces Savage's continuously divisible events by an infinite consequence set and an ethically neutral event, as a close contender.

If one does not object to the direct use of extraneous scaling probabilities, then several theories in section 8 are quite satisfactory. Pratt, Raiffa and Schlaifer (1965) may be the best of these. Their format allows general application with simple, interpretable axioms that tie in very closely with assessment. Fishburn (1969, 1972) uses a structure that is simpler in certain respects but does not tie in to assessment as directly as the structure used by Pratt *et al.* While Fishburn (1973, 1974) attempts to relax some of the structural restrictions of others, he has a few less-intuitive axioms and some potential interpretational problems in conditional preference comparisons.

Finally, although the theory of Luce and Krantz (1971) is appealing in its conditional approach and has reasonably straightforward axioms, I feel that it encounters serious interpretational difficulties. These arise from their combinations and restrictions of conditional acts, which allow them to avoid extraneous probabilities while relaxing Savage's structure. However, there is room for question as to whether the latter positive features outweigh the interpretational problems in their approach.

ACKNOWLEDGMENTS

I am indebted to Duncan Luce and John Pratt for past discussions that have materially influenced the nature of this review, and to Luce, Pratt, Zoltan Domotor and Mike DeFazio for comments and suggestions on an earlier draft of the paper.

Bell Laboratories, New Jersey

NOTES

[1] See, for example, Allais (1953), Edwards (1954), Davidson, Suppes and Siegel (1957), Ellsberg (1961), Luce and Suppes (1965), Becker and McClintock (1967), Tversky (1967b), MacCrimmon (1968), Rapoport and Wallsten (1972), Slovic and Tversky (1974), Kahneman and Tversky (1979), and Grether and Plott (1979).

[2] See also Tversky (1967a), Shapiro (1979) and Richter and Shapiro (1978). By 'one-way' we mean that the representation is not an "if and only if" model for preferences but only seeks to attribute greater utility to one act than to a second act when the first is preferred to the second.

[3] Other renditions of the von Neumann-Morgenstern theory are presented by Marschak (1950), Friedman and Savage (1952), Luce and Raiffa (1957), and Fishburn and Roberts (1978).

[4] Axiom systems to measure utility differences that are not based on even-chance gambles or an 'ethically neutral' event are presented by Suppes and Winet (1955), Scott and Suppes (1958), Suppes and Zinnes (1963, pp. 34–38), Fishburn (1970, Chapter 6) and Krantz *et al.* (1971, Chapter 4). See also Pfanzagl (1959, 1968). While Ramsey talks about utility differences, he could just as well have talked about sums, or the simple representation of cEd by $u(c) + u(d)$ or $\frac{1}{2}u(c) + \frac{1}{2}u(d)$.

[5] See, e.g., Debreu (1960), Luce and Tukey (1964), Fishburn (1970) and Krantz *et al.* (1971).

[6] See Savage (1967) and the commentaries following his article for additional discussions of personal (subjective) probability.

[7] Various criticisms and defenses of this principle are presented by Savage (1954, pp. 101–103), Ellsberg (1961), Raiffa (1961), MacCrimmon (1968) and Slovic and Tversky (1974), among others.

[8] See Krantz *et al.* (1971) for a thorough discussion of extensive measurement.

[9] See, for example, Rubin (1949). As far as I know, Rubin's papers on this topic remain unpublished.

[10] Commentaries following Balch and Fishburn (1974), by David Krantz and Duncan Luce, Richard Jeffrey, Ethan Bolker, and John Pratt, go into this further. See also Spohn (1977).

[11] In addition to the references given here, see Sneed (1966) and Schick (1967) for comments on Jeffrey's approach.

REFERENCES

Allais, M., 1953, 'Le Comportement de l'Homme Rationnel Devant le Risque: Critique des Postulats et Axiomes de l'École Americaine', *Econometrica* **21**, 503–546.

Anscombe, F. J. and Aumann, R. J., 1963, 'A Definition of Subjective Probability', *Annals of Mathematical Statistics* **34**, 199–205.

Arrow, K. J., 1951, 'Alternative Approaches to the Theory of Choice in Risk-Taking Situations', *Econometrica* **19**, 404–437.

Arrow, K. J., 1966, 'Exposition of the Theory of Choice under Uncertainty', *Synthese* **16**, 253–269.

Balch, M., 1974, 'On Recent Developments in Subjective Expected Utility', in M. Balch, D. McFadden and S. Wu (eds.), *Essays on Economic Behavior Under Uncertainty*, North-Holland, Amsterdam, pp. 45–54.

Balch, M. and Fishburn, P. C., 1974, 'Subjective Expected Utility for Conditional Primitives', in M. Balch, D. McFadden and S. Wu (eds.), *Essays on Economic Behavior Under Uncertainty*, North-Holland, Amsterdam, pp. 57–69.

Becker, G. M. and McClintock, C. G., 1967, 'Value: Behavioral Decision Theory', *Annual Review of Psychology* **18**, 237–286.

Blackwell, D. and Girshick, M. A., 1954, *Theory of Games and Statistical Decisions*, Wiley, New York.

Bolker, E. D., 1966, 'Functions Resembling Quotients of Measures', *Transactions of the American Mathematical Society* **124**, 292–312.

Bolker, E. D., 1967, 'A Simultaneous Axiomatization of Utility and Subjective Probability', *Philosophy of Science* **34**, 333–340.

Chernoff, H., 1954, 'Rational Selection of Decision Functions', *Econometrica* **22**, 422–443.

Davidson, D. and Suppes, P., 1956, 'A Finitistic Axiomatization of Subjective Probability and Utility', *Econometrica* **24**, 264–275.

Davidson, D., Suppes, P. and Siegel, S., 1957, *Decision Making: An Experimental Approach*, Stanford University Press, Stanford, California.

Debreu, G., 1959, 'Cardinal Utility for Even-chance Mixtures of Pairs of Sure Prospects', *Review of Economic Studies* **26**, 174–177.

Debreu, G., 1960, 'Topological Methods in Cardinal Utility Theory', in K. J. Arrow, S. Karlin and P. Suppes (eds.), *Mathematical Methods in the Social Sciences, 1959*, Stanford University Press, Stanford, California, pp. 16–26.

de Finetti, B., 1937, 'La Prévision ses Lois Logiques, ses Sources Subjectives', *Annales de l'Institut Henri Poincaré* **7**, 1–68. Translated by H. E. Kyburg as 'Foresight: Its Logical Laws, Its Subjective Sources', in H. E. Kyburg and H. E. Smokler (eds.), *Studies in Subjective Probability*, Wiley, New York, 1964, pp. 97–158.

Domotor, Z., 1978, 'Axiomatization of Jeffrey Utilities', *Synthese* **39**, 165–210.

Dubins, L. E. and Savage, L. J., 1965, *How to Gamble if You Must: Inequalities for Stochastic Processes*, McGraw-Hill, New York.

Edwards, W., 1954, 'The Theory of Decision Making', *Psychological Bulletin* **51**, 380–417.

Ellsberg, D., 1961, 'Risk, Ambiguity, and the Savage Axioms', *Quarterly Journal of Economics* **75**, 643–669.

Ferreira, P. E., 1972, 'On Subjective Probability and Expected Utilities', *Annals of Mathematical Statistics* **43**, 928–933.

Fishburn, P. C., 1964, *Decision and Value Theory*, Wiley, New York.

Fishburn, P. C., 1967, 'Preference-Based Definitions of Subjective Probability', *Annals of Mathematical Statistics* **38**, 1605–1617.

Fishburn, P. C., 1969, 'A General Theory of Subjective Probabilities and Expected Utilities', *Annals of Mathematical Statistics* **40**, 1419–1429.

Fishburn, P. C., 1970, *Utility Theory for Decision Making,* Wiley, New York. Reprinted by Krieger Publishing Co., Huntington, New York, 1979.

Fishburn, P. C., 1972, 'Subjective Expected Utility with Mixture Sets and Boolean Algebras', *Annals of Mathematical Statistics*, **43**, 917–927.

Fishburn, P. C., 1973, 'A Mixture-Set Axiomatization of Conditional Subjective Expected Utility', *Econometrica* **41**, 1–25.

Fishburn, P. C., 1974, 'On the Foundations of Decision Making Under Uncertainty', in M. Balch, D. McFadden and S. Wu (eds.), *Essays on Economic Behavior Under Uncertainty*, North-Holland, Amsterdam, pp. 25–44.

Fishburn, P. C., 1975, 'A Theory of Subjective Expected Utility with Vague Preferences', *Theory and Decision* **6**, 287–310.

Fishburn, P. C., 1977, 'Expected Utility Theories: A Review Note', in R. Henn and O. Moeschlin (eds.), *Mathematical Economics and Game Theory*, Springer-Verlag, Berlin and New York, pp. 197–207.

Fishburn, P. C. and Roberts, F. S., 1978, 'Mixture Axioms in Linear and Multilinear Utility Theories', *Theory and Decision* **9**, 161–171.

Friedman, M. and Savage, L. J., 1952, 'The Expected Utility Hypothesis and the Measurability of Utility', *Journal of Political Economy* **60**, 463–474.

Gibbard, A. and Harper, W. L., 1978, 'Counterfactuals and Two Kinds of Expected Utility', in C. A. Hooker, J. J. Leach and E. F. McClennen (eds.), *Foundations and Applications of Decision Theory, Vol. I: Theoretical Foundations*, Reidel, Dordrecht, Holland, pp. 125–162.

Grether, D. M. and Plott, C. R., 1979, 'Economic Theory of Choice and the Preference Reversal Phenomenon', *American Economic Review* **69**, 623–638.

Hausdorff, F., 1957, *Set Theory* (translation), Chelsea, New York.

Hausner, M., 1954, 'Multidimensional Utilities', in R. M. Thrall, C. H. Coombs and R. L. Davis (eds.), *Decision Processes*, Wiley, New York, pp. 167–180.

Herstein, I. N. and Milnor, J., 1953, 'An Axiomatic Approach to Measurable Utility', *Econometrica* **21**, 291–297.

Jeffrey, R. C., 1965a, *The Logic of Decision*, McGraw-Hill, New York.

Jeffrey, R. C., 1965b, 'Ethics and the Logic of Decision', *Journal of Philosophy* **62**, 528–539.

Jeffrey, R. C., 1974, 'Frameworks for Preference', in M. Balch, D. McFadden and S. Wu (eds.), *Essays on Economic Behavior Under Uncertainty*, North-Holland, Amsterdam, pp. 74–79.

Jeffrey, R. C., 'Axiomatizing the Logic of Decision', in C. A. Hooker, J. J. Leach and E. F. McClennen (eds.), *Foundations and Applications of Decision Theory, Vol. I: Theoretical Foundations*, Reidel, Dordrecht, Holland, pp. 227–231.

Jensen, N. E., 1967, 'An Introduction of Bernoullian Utility Theory. I. Utility Functions', *Swedish Journal of Economics* **69**, 163–183.

Kahneman, D. and Tversky, A., 1979, 'Prospect Theory: An Analysis of Decision Under Risk', *Econometrica* **47**, 263–291.

Krantz, D. H., Luce, R. D., Suppes, P. and Tversky, A., 1971, *Foundations of Measurement*, Volume I, Academic Press, New York.

Luce, R. D., 1956, 'Semiorders and a Theory of Utility Discrimination', *Econometrica* **24**, 178–191.

Luce, R. D., 1958, 'A Probabilistic Theory of Utility', *Econometrica* **26**, 193–224.

Luce, R. D., 1972, 'Conditional Expected, Extensive Utility', *Theory and Decision* **3**, 101–106.

Luce, R. D. and Krantz, D. H., 1971, 'Conditional Expected Utility', *Econometrica* **39**, 253–271.

Luce, R. D. and Raiffa, H., 1957, *Games and Decisions*, Wiley, New York.

Luce, R. D. and Suppes, P., 1965, 'Preference, Utility, and Subjective Probability', in R. D. Luce, R. R. Bush and E. Galanter (eds.), *Handbook of Mathematical Psychology, III*, Wiley, New York, pp. 250–410.

Luce, R. D. and Tukey, J. W., 1964, Simultaneous Conjoint Measurement: A New Type of Fundamental Measurement', *Journal of Mathematical Psychology* **1**, 1–27.

MacCrimmon, K. R., 1968, 'Descriptive and Normative Implications of the Decision-Theory Postulates', in K. Borch and J. Mossin (eds.), *Risk and Uncertainty*, Macmillan, New York, pp. 3–32.

Marschak, J., 1950, 'Rational Behavior, Uncertain Prospects, and Measurable Utility', *Econometrica* **18**, 111–141.

Maskin, E., 1979, 'Decision Making under Ignorance with Implications for Social Choice', *Theory and Decision* **11**, 319–337.

Milnor, J., 1954, 'Games Against Nature', in R. M. Thrall, C. H. Coombs and R. L. Davis (eds.), *Decision Processes*, Wiley, New York, pp. 49–59.

Myerson, R. B., 1979, 'An Axiomatic Derivation of Subjective Probability, Utility, and Evaluation Functions', *Theory and Decision* **11**, 339–352.

Narens, L., 1976, Utility Uncertainty Trade-off Structures', *Journal of Mathematical Psychology* **13**, 296–322.

Pfanzagl, J., 1959, 'A General Theory of Measurement: Applications to Utility', *Naval Research Logistics Quarterly* **6**, 283–294.

Pfanzagl, J., 1967, 'Subjective Probability Derived from the Morgenstern-von Neumann Utility Concept', in M. Shubik (ed.), *Essays in Mathematical Economics*, Princeton University Press, Princeton, pp. 237–251.

Pfanzagl, J., 1968, *Theory of Measurement*, Wiley, New York.

Pratt, J. W., 1974, 'Some Comments on Some Axioms for Decision Making Under Uncertainty', in M. Balch, D. McFadden and S. Wu (eds.), *Essays on Economic Behavior Under Uncertainty*, North-Holland, Amsterdam, pp. 82–92.

Pratt, J. W., Raiffa, H. and Schlaifer, R., 1964, 'The Foundations of Decision under Uncertainty: An Elementary Exposition', *Journal of the American Statistical Association* **59**, 353–375.

Pratt, J. W., Raiffa, H. and Schlaifer, R., 1965, *Introduction to Statistical Decision Theory*, McGraw-Hill, New York.

Raiffa, H., 1961, 'Risk, Ambiguity, and the Savage Axioms: Comment', *Quarterly Journal of Economics* **75**, 690–694.

Ramsey, F. P., 1931, 'Truth and Probability', in F. P. Ramsey, *The Foundations of Mathematics and Other Logical Essays*, Harcourt, Brace and Co,. New York, Reprinted in H. E. Kyburg and H. E. Smokler (eds.), *Studies in Subjective Probability*, Wiley, New York, 1964, pp. 61–92.

Rapoport, A. and Wallsten, T. S., 1972, 'Individual Decision Behavior', *Annual Review of Psychology* **23**, 131–176.

Richter, M. K., 1975, 'Rational Choice and Polynomial Measurement Models', *Journal of Mathematical Psychology* **12**, 99–113.

Richter, M. K. and Shapiro, L., 1978, 'Revelations of a Gambler', *Journal of Mathematical Economics* **5**, 229–244.

Roberts, F. S., 1974, 'Laws of Exchange and Their Applications', *SIAM Journal on Applied Mathematics* **26**, 260–284.

Robinson, A., 1966, *Non-standard Analysis*, North-Holland, Amsterdam.
Rubin, H., 1949, 'Postulates for the Existence of Measurable Utility and Psychological Probability' (abstract), *Bulletin of the American Mathematical Society* **55**, 1050–1051.
Savage, L. J., 1954, *The Foundations of Statistics*, Wiley, New York. Second revised edition, Dover Publications, New York, 1972.
Savage, L. J., 1967, 'Difficulties in the Theory of Personal Probability', *Philosophy of Science* **34**, 305–310. Discussions also by I. Hacking (pp. 311–325) and A. Shimony (pp. 326–332).
Schick, F., 1967, Review of *The Logic of Decision* (by R. C. Jeffrey), *Journal of Philosophy* **64**, 396–400.
Scott, D. and Suppes, P., 1958, 'Foundational Aspects of Theories of Measurement', *Journal of Symbolic Logic* **23**, 113–128.
Shapiro, L., 1979, 'Necessary and Sufficient Conditions for Expected Utility Maximization: The Finite Case, with a Partial Order', *Annals of Statistics* **7**, 1288–1302.
Slovic, P. and Tversky, A., 1974, 'Who Accepts Savage's Axiom?', *Behavioral Science* **19**, 368–373.
Sneed, J. D., 1966, 'Strategy and the Logic of Decision', *Synthese* **16**, 270–283.
Spohn, W., 1977, 'Where Luce and Krantz do Really Generalize Savage's Decision Model', *Erkenntnis* **11**, 113–134.
Stigum, B. P., 1972, 'Finite State Space and Expected Utility Maximization', *Econometrica* **40**, 253–259.
Suppes, P., 1956, 'The Role of Subjective Probability and Utility in Decision Making', *Proceedings of the Third Berkeley Symposium on Mathematical Statistics and Probability, 1954–1955* **5**, 61–73.
Suppes, P. and Winet, M., 1955, 'An Axiomatization of Utility Based on the Notion of Utility Differences', *Management Science* **1**, 259–270.
Suppes, P. and Zinnes, J. L., 1963, 'Basic Measurement Theory', in R. D. Luce, R. R. Bush and E. Galanter (eds.), *Handbook of Mathematical Psychology I*, Wiley, New York, pp. 1–76.
Tversky, A., 1967a, 'A General Theory of Polynomial Conjoint Measurement', *Journal of Mathematical Psychology* **4**, 1–20.
Tversky, A., 1967b, 'Additivity, Utility, and Subjective Probability', *Journal of Mathematical Psychology* **4**, 175–201.
Tversky, A., 1969, 'Intransitivity of Preferences', *Psychological Review* **76**, 31–48.
von Neumann, J. and Morgenstern, O., 1947, *Theory of Games and Economic Behavior*, 2nd ed., Princeton University Press, Princeton.