The Annals of Statistics
1975, Vol. 3, No. 4, 884–896

UNBOUNDED EXPECTED UTILITY

BY PETER C. FISHBURN

The Pennsylvania State University

Let P be a convex set of finitely additive probability measures defined on a Boolean algebra of subsets of a set X of consequences. Axioms are specified for a preference relation $\succ$ on P which are necessary and sufficient for the existence of a real-valued utility function u on X for which expected utility $E(u, p)$ is finite for all p in P and for which $p \succ q$ iff $E(u, p) > E(u, q)$, for all p and q in P. A slightly simpler set of axioms yields the same results when the algebra is a Borel algebra and every measure in P is countably additive. The axioms allow P to contain nonsimple probability measures without necessarily implying that the utility function u is bounded.

1. Introduction. Given a preference relation $\succ$ on a set P of probability measures defined on an algebra of subsets of a set X of consequences, this paper addresses the question of whether there exists a real-valued utility function u on X for which expected utility $E(u, p) = \int u(x)\, dp(x)$ exists and is finite for all $p \in P$ and for which

$$(1) \qquad p \succ q \quad \text{if and only if} \quad E(u, p) > E(u, q)$$

for all $p, q \in P$. The primary purpose of the paper is to present general axioms for $\succ$ on P which are necessary and sufficient for the existence of u as specified whenever P is any set of finitely additive probability measures defined on an appropriate Boolean algebra of subsets of X, subject only to elementary closure conditions on P. Secondarily, we note how these axioms simplify to provide necessary and sufficient conditions for the finite expected-utility representation when every $p \in P$ is a countably additive measure defined on a Borel algebra (σ-field) of subsets of X. As will be observed, the present formulation subsumes most of the special cases treated by others. In addition, it is designed to deal with P sets which include nonsimple measures without necessarily implying that the utility function u is bounded. The possibility of unbounded u in the presence of nonsimple measures, discussed some time ago by Menger [10], distinguishes the present theory from all others that I am aware of except for the important contributions by DeGroot [4] and Ledyard [9]. I shall return to these momentarily.

The initial axiomatization of (1), by von Neumann and Morgenstern [12], was designed for the case in which P is the set P_s of all simple probability measures on X (those which assign probability 1 to a finite subset of X). Despite this, their axioms, or an equivalent set of axioms [6, 8], can be applied to any

Received June 1974; revised January 1975.

AMS 1970 *subject classifications.* Primary 62C05, 90A10; Secondary 06A75, 60A05, 90D35.

Key words and phrases. Expected utility, unbounded utility, convex sets of probability measures, preference axioms.

set of probability measures which is closed under finite convex combinations with $\lambda p + (1 - \lambda)q \in P$ whenever $0 \leqq \lambda \leqq 1$ and $p, q \in P$. However, these axiomatizations imply (1) only when p and q are in P_s. Thus, when $P_s \subseteqq P$, the axioms of von Neumann and Morgenstern yield (1) for all $p, q \in P_s$ but do not necessarily imply (1) when either p or q is not a simple measure. In this case u can be unbounded when X has an infinite number of indifference classes.

The next significant development for (1), by Blackwell and Girshick [3], took P equal to the set P_d of all discrete probability measures on X (those which are countably additive and assign probability 1 to a countable subset of X). To obtain (1) for all $p, q \in P_d$ they used an independence or dominance axiom based on countable convex combinations of measures and noted in this case that u must be bounded. Later extensions of this approach by Arrow [1] and Fishburn [5, 6], designed to handle nondiscrete measures in the context of (1) with $P_d \subseteqq P$, also require u to be bounded.

This work left open the possibility of specifying axioms which imply (1) for all $p, q \in P$ but do not necessarily imply that u is bounded when $P_s \subset P$ but P_d is not included in P. This possibility was realized by DeGroot [4] and Ledyard [9]. Both authors first apply the preference axioms to a convex set P_b of bounded probability measures. With $\precsim$ a preference-or-indifference relation on the consequence set X (read $x \precsim y$ as "x is not preferred to y"), p is *bounded* if and only if $p(\{x : x_1 \precsim x \precsim x_2\}) = 1$ for some $x_1, x_2 \in X$. Since only bounded measures are used in this initial step, (1) is obtained for all $p, q \in P_b$ without requiring u to be bounded. Having obtained u on X in a bounded-measures context, they then use appropriate axioms to extend (1) to unbounded measures that are integrable with respect to u in the sense that $E(u, p)$ exists and is finite for each such measure.

This paper presents an alternative to the two-step approach adopted by DeGroot and Ledyard. Throughout, we shall work with a fixed set P of measures which, in a specific decision situation, would be determined by particular aspects of that situation. Some measures in P will be bounded while others may be unbounded. The axioms for $\succ$ on P are applied in one step. Since the axioms are necessary and sufficient for the finite expected-utility representation, they show precisely what must be true of $\succ$ on P to obtain this representation for the particular P under consideration. Thus, instead of going through a first step to determine which measures are integrable with respect to u, we specify what must be true of $\succ$ on P so that all measures in P will in fact be integrable with respect to an appropriate u. Depending on the particular characteristics of P, the axioms may or may not entail boundedness of u.

The practical import of not requiring u to be bounded when P contains nonsimple measures is most easily illustrated in the traditional monetary setting where say $X = [0, \infty)$ and each $x \in X$ represents a potential net wealth of an individual. For the sake of argument, suppose that preference increases in x. Then if (1) is to apply and if $P_d \subseteqq P$, u must be bounded above. However, some of the more attractive functional forms for u which agree fairly well with some individuals' preferences, such as power functions $u(x) = (x + a)^b$ with $a \geqq 0$ and $0 < b < 1$, or log functions $u(x) = \log(x + a)$ with $a > 0$, are unbounded above. Hence, if such a u is used for (1), then P must be restricted to avoid contradiction with finite $E(u, p)$. In particular, P must not include all discrete distributions although it may contain a variety of nonsimple discrete distributions as well as various nondiscrete distributions such as lower-truncated normals or gammas. Indeed, if P is not allowed to contain any measure with infinite mean (which is a very sensible restriction in the net wealth setting), then [2, 7] any increasing u which is concave or risk-averse on $[x, \infty)$ for some $x \in X$ has finite $E(u, p)$ for every $p \in P$. Additional comments in the net wealth setting on the aspect of unbounded u are given by Ryan [11] and Arrow [2].

The rest of the paper is organized as follows. The next section formulates the minimal requirements that will be imposed on the Boolean algebra on which probability measures are defined and on the set P of measures to which $\succ$ is applied in the axioms. Section 3 then presents four axioms which are necessary for (1) with finite expected utilities and notes that they imply (1) *if* u is bounded. Section 4 shows that a fifth necessary axiom is required for those cases in which u might be unbounded. Since Ledyard's axioms are set within the Boolean algebra framework, I shall comment further on his theory at the end of Section 4. Section 5 then shows how the axioms simplify when the algebra is a Borel algebra and every measure in P is countably additive. Further comparisons with DeGroot's theory are made at the end of Section 5.

2. Notation and structure. Our axioms specify the behavior of a binary relation $\succ$ ("is preferred to") on a set P of probability measures defined on a Boolean algebra $\mathscr{A}$ of subsets of the set X of consequences. The basic structure for $(\mathscr{A}, P, \succ)$ that is used in the axioms and derivations will be set forth in this section.

First, each singleton subset $\{x\}$ from X will be assumed to be in $\mathscr{A}$ and each one-point measure will be assumed to be in P. We distinguish between consequence $x \in X$ and the measure which assigns probability 1 to x by letting x^* denote the latter. That is, $x^* \in P$ and $x^*(\{x\}) = 1$ for every $x \in X$. Then $x^* \succ y^*$ indicates that consequence x is preferred to consequence y. We shall also write $x^* \succ y^*$ as $x \succ y$.

The use of one-point measures permits us to define utilities of consequences in an unambiguous way. Similarly, for the purpose of ensuring that expected utilities are well defined, we shall require all preference intervals from X to be in $\mathscr{A}$. Writing $x \prec y$ iff $y \succ x$, $x \succsim y$ iff not $(y \succ x)$, and $x \precsim y$ iff $y \succsim x$, a subset A of X is a *preference interval* iff $z \in A$ whenever $x, y \in A$, $x \precsim z$ and $z \precsim y$. In special cases the following notation is used for the designated preference intervals: $(-\infty, x) = \{y \in X : y \prec x\}$, $(-\infty, x] = \{y \in X : y \precsim x\}$, $(x, y] = \{z \in X : x \prec z \text{ and } z \precsim y\}$, $(x, \infty) = \{y \in X : y \succ x\}$, and so forth. It will be assumed later that $\succ$ on P is an asymmetric weak order, in which case all preference intervals from X will be in $\mathscr{A}$ iff $\mathscr{A}$ contains each lower interval A having the property that $x \in A$ when $y \in A$ and $x \precsim y$, and contains each upper interval B having the property that $y \in B$ when $x \in B$ and $x \precsim y$. In some cases, such as increasing preferences on $X = [0, \infty)$, the preference intervals have a natural structure, but this will not be true in general. If $\mathscr{A}$ is taken to be the power set of X then all preference intervals will of course be in $\mathscr{A}$.

Besides inclusion of all one-point measures in P, we shall assume that P is closed under finite convex combinations and that P is closed under the formation of conditional probability measures on preference intervals with positive measure. Both assumptions facilitate the orderly computation of expected utilities as discussed in [6, pages 134–137]. The convex closure assumption says that $\lambda p + (1 - \lambda)q \in P$ when $\lambda \in [0, 1]$ and $p, q \in P$. The convex combination $\lambda p + (1 - \lambda)q$ is the measure which assigns probability $\lambda p(A) + (1 - \lambda)q(A)$ to each $A \in \mathscr{A}$. The conditional probability closure assumption says that if $p \in P$, if A is a preference interval from X, and if $p(A) > 0$, then $p_A \in P$ where $p_A(B) = p(A \cap B)/p(A)$ for all $B \in \mathscr{A}$. If $A = [x, y] = \{z \in X: x \precsim z \precsim y\}$ then $p_A = p_{[x,y]}$.

These structural impositions on $(\mathscr{A}, P, \succ)$ are summarized by the following axiom.

Axiom 0. *The set $\mathscr{A}$ is a Boolean algebra of subsets of X which contains $\{x\}$ for each $x \in X$ and contains every preference interval from X. The set P is a set of finitely-additive probability measures defined on $\mathscr{A}$ such that: $x^* \in P$ for each $x \in X$; $\lambda p + (1 - \lambda)q \in P$ whenever $\lambda \in [0, 1]$ and $p, q \in P$; and $p_A \in P$ whenever $p \in P$, A is a preference interval from X, and $p(A) > 0$.*

The smallest P which satisfies Axiom 0 is the set P_s of all simple probability measures on $\mathscr{A}$. The largest P which satisfies the axiom is of course the set of all finitely-additive probability measures on $\mathscr{A}$. Unless X is finite, a huge number of admissible P sets lie between these extremes.

3. Four preference axioms. Our first four axioms for $\succ$ on P are taken from Fishburn [5, 6]. As explained in this section, they do most of the work towards implying (1) for all $p, q \in P$ with finite $E(u, p)$ for all $p \in P$. A fifth axiom which may be needed for (1) is discussed in the next section. The relation $\succ$ on P is an *asymmetric weak order* iff it is asymmetric ($p \succ q \Rightarrow$ not $(q \succ p)$) and negatively transitive ($p \succ q \Rightarrow p \succ r$ or $r \succ q$), for all $p, q, r \in P$. When $\succ$ on P is an asymmetric weak order, the preference-indifference relation $\succsim$ on P, defined by $p \succsim q$ iff not $(q \succ p)$, is a 'weak order' or 'complete preorder,' being reflexive, transitive and complete ($p \neq q \Rightarrow p \succsim q$ or $q \succsim p$).

Axiom 1. *The binary relation $\succ$ on P is an asymmetric weak order.*

Axiom 2. *If $p, q, r \in P$, $p \succ q$ and $0 < \lambda < 1$, then $\lambda p + (1 - \lambda)r \succ \lambda q + (1 - \lambda)r$.*

AXIOM 3. *If $p, q, r \in P$, $p \succ q$ and $q \succ r$, then $\alpha p + (1-\alpha)r \succ q$ and $q \succ \beta p + (1-\beta)r$ for some α and β strictly between zero and one.*

AXIOM 4. *If $p, q \in P$, $A \in \mathscr{A}$ and $p(A) = 1$, then $p \succsim q$ if $x^* \succ q$ for all $x \in A$, and $q \succsim p$ if $q \succ x^*$ for all $x \in A$.*

Since these axioms are discussed in detail elsewhere [5, 6], we shall mention here only that Axiom 2 is an independence or linearity assumption which preserves preference under similar combinations, Axiom 3 is an Archimedean-type assumption which facilitates the derivation of real-valued (as opposed to vector-valued) utilities, and Axiom 4 is a dominance principle which says, for example, that if $p(\{x : x^* \succ q\}) = 1$ then $p \succsim q$. All four axioms are easily seen to be necessary for (1) with finite expected utilities, given Axiom 0. In addition, Axiom 4 is superfluous if $P = P_s$, as can be seen from the following basic lemma [6, Theorem 8.4].

LEMMA 1. *Suppose Axiom 0 holds. Then Axioms 1, 2 and 3 hold if and only if there exists $v : P \to Re$ such that, for all $\lambda \in [0, 1]$ and all $p, q \in P$,*

$$p \succ q \quad \text{iff} \quad v(p) > v(q)\,, \tag{2}$$

$$v(\lambda p + (1-\lambda)q) = \lambda v(p) + (1-\lambda)v(q)\,. \tag{3}$$

Moreover, if v satisfies (2) and (3) for all $\lambda \in [0, 1]$ and all $p, q \in P$, then $v' : P \to Re$ satisfies (2) and (3) in place of v for all $\lambda \in [0, 1]$ and all $p, q \in P$, if and only if there exist real numbers $a > 0$ and b such that $v'(p) = av(p) + b$ for all $p \in P$.

Given v on P as in Lemma 1, it is natural (though not essential) to define $u : X \to Re$ by $u(x) = v(x^*)$ for all $x \in X$. Then, by the linearity property (3), $E(u, p) = v(p)$ for all $p \in P_s$. Hence, using (2), Axioms 1 through 3 yield (1) whenever p and q are simple measures. However, as shown by examples in [5, 6], Axioms 1 through 3 do not imply $E(u, p) = v(p)$ when p is not simple. The next lemma shows that our further concerns can focus on the question of whether $E(u, p) = v(p)$ for all $p \in P$.

LEMMA 2. *Suppose Axiom 0 holds and $v : P \to Re$ satisfies (2) and (3) for all $\lambda \in [0, 1]$ and all $p, q \in P$. Then, defining $u : X \to Re$ by $u(x) = v(x^*)$ for all $x \in X$, there exists $u' : X \to Re$ which satisfies (1) for all $p, q \in P$ and for which $E(u', p)$ exists and is finite for all $p \in P$, if and only if $E(u, p) = v(p)$ for all $p \in P$.*

PROOF. Suppose $E(u, p) = v(p)$ for all $p \in P$. Then $E(u, p)$ is finite for all $p \in P$ and (1) follows from (2). Conversely, suppose $u' : X \to Re$ gives $E(u', p)$ finite for all $p \in P$ along with $p \succ q$ iff $E(u', p) > E(u', q)$, for all $p, q \in P$. Then $E(u', \cdot)$ is linear on P; that is, $E(u', \lambda p + (1-\lambda)q) = \lambda E(u', p) + (1-\lambda)E(u', q)$. Consequently, defining $v'(p) = E(u', p)$, (2) and (3) hold for v'. Therefore, by the last part of Lemma 1, there exist $a > 0$ and b such that $v'(p) = av(p) + b$ for all $p \in P$. Hence $av(p) + b = v'(p) = E(u', p) = aE(u, p) + b$ so that $v(p) = E(u, p)$ for all $p \in P$. □

Henceforth it is to be understood that v *satisfies* (2) *and* (3) *and that* $u(x) = v(x^*)$ *for all* $x \in X$. The effect of Axiom 4 on the question of whether $E(u, p) = v(p)$ will now be noted.

THEOREM 1. *Suppose Axioms* 0 *through* 4 *hold. Then, for all* $p \in P$:

(a) $E(u, p) = v(p)$ *if there exists* $A \in \mathscr{A}$, *such that* $p(A) = 1$ *and both* $\inf\{u(x) : x \in A\}$ *and* $\sup\{u(x) : x \in A\}$ *are finite*;

(b) $v(p) \geqq E(u, p)$ *if there exists* $A \in \mathscr{A}$ *such that* $p(A) = 1$ *and* $\inf\{u(x) : x \in A\}$ *is finite*;

(c) $E(u, p) \geqq v(p)$ *if there exists* $A \in \mathscr{A}$ *such that* $p(A) = 1$ *and* $\sup\{u(x) : x \in A\}$ *is finite*;

(d) $E(u, p)$ *is well defined and finite.*

PROOF. The proof of part (a) is essentially the same as the proof given for Theorem 10.1 on page 142 of [6]. With $c = \inf\{u(x) : x \in A\}$ and $d = \sup\{u(x) : x \in A\}$ and c and d finite, it follows from standard definitions of expectations that $E(u, p)$ is the supremum of $\{E(f_i, p) : i = 1, 2, \cdots\}$ where $f_1, f_2, \cdots$ is a sequence of simple functions on $B = \{x : c \leqq u(x) \leqq d\}$ which converge uniformly from below to u on B. (A simple function has only a finite number of distinct values.) Axiom 4 leads to $c \leqq v(p) \leqq d$, and appropriate partitions of B with the use of the closure properties in Axiom 0 leads to $v(p) = E(u, p)$.

For part (b) suppose that $A \in \mathscr{A}$, $p(A) = 1$, $\inf\{u(x) : x \in A\}$ is finite and, to avoid case (a) (which gives the desired result), suppose also that $p((y, \infty)) > 0$ for all $y \in X$. Letting $E(u, p_{(-\infty,y]}) = 0$ when $p_{(-\infty,y]}$ is undefined (because of $p((-\infty, y]) = 0$), the expectation definition for this case leads to

$$E(u, p) = \sup\{p((-\infty, y])E(u, p_{(-\infty,y]}) + p((y, \infty))u(y) : y \in X\}\,. \tag{4}$$

Since the sum in braces in (4) does not decrease as $u(y)$ increases [i.e., $\int_X \min\{u(x), u(y)\}\, dp(x)$ does not decrease as $u(y)$ increases], (4) can also be written as

$$E(u, p) = \lim\{p((-\infty, y])E(u, p_{(-\infty,y]}) + p((y, \infty))u(y)\} \tag{4'}$$

with the understanding that the limit is taken as $u(y)$ approaches $\sup u(X)$, whether finite or infinite. Now for each $y \in X$, $p = p_{(y,\infty)}$ if $p((y, \infty)) = 1$, and $p = p((-\infty, y])p_{(-\infty,y]} + p((y, \infty))p_{(y,\infty)}$ if $p((y, \infty)) < 1$. Hence, letting $v(p_{(-\infty,y]}) = 0$ by convention when $p((-\infty, y]) = 0$, (3) yields

$$v(p) = p((-\infty, y])v(p_{(-\infty,y]}) + p((y, \infty))v(p_{(y,\infty)}) \quad \text{for all} \quad y \in X\,. \tag{5}$$

Now when $p((-\infty, y]) > 0$, $p_{(-\infty,y]}(A \cap (-\infty, y]) = 1$ and therefore $v(p_{(-\infty,y]}) = E(u, p_{(-\infty,y]})$ by part (a) of the theorem. Moreover, since $x \succ y$ for every $x \in (y, \infty)$, and since $p_{(y,\infty)}((y, \infty)) = 1$, Axiom 4 along with (2) yields $v(p_{(y,\infty)}) \geqq v(y^*) = u(y)$, for all $y \in X$. Comparison of (4) and (5) then shows that $v(p) \geqq E(u, p)$, as was to be proved.

The proof of part (c) is similar to the proof of part (b) (using the other half of Axiom 4) and will be omitted.

Finally, for part (d), let p be in P and let z be a consequence in X. If $p((-\infty, z]) = 1$ then $p = p_{(-\infty,z]}$ and $v(p) \leqq E(u, p) \leqq v(z^*)$ by part (c); if $p((z, \infty)) = 1$ then $p = p_{(z,\infty)}$ and $v(z^*) \leqq E(u, p) \leqq v(p)$ by part (b); otherwise, $p = p((-\infty, z])p_{(-\infty,z]} + p((z, \infty))p_{(z,\infty)}$ and $p((-\infty, z])v(p_{(-\infty,z]}) + p((z, \infty))v(z^*) \leqq E(u, p) \leqq p((-\infty, z])v(z^*) + p((z, \infty))v(p_{(z,\infty)})$. □

Clearly, if u is bounded, then $E(u, p) = v(p)$ for all $p \in P$ under Axioms 0 through 4. Moreover, as shown by part (a) of Theorem 1, if u is unbounded but if each $p \in P$ has a corresponding $A \in \mathscr{A}$ such that $p(A) = 1$ and $\inf u(A)$ and $\sup u(A)$ are finite, then $E(u, p) = v(p)$ for all $p \in P$. The problem in obtaining $E(u, p) = v(p)$ therefore arises in the setting of Axioms 0 through 4 only when $p((-\infty, y]) < 1$ for all $y \in X$ and u is unbounded above, or when $p([y, \infty)) < 1$ for all $y \in X$ and u is unbounded below. We examine this case further in the next section.

4. A fifth preference axiom. The following example shows that $E(u, p) = v(p)$ is not guaranteed by Axioms 0 through 4.

EXAMPLE 1. Let $X = \{1, 2, \cdots\}$, let $\mathscr{A}$ equal the set of all finite subsets of X and their complements, let P_0 equal the set of all one-point measures on $\mathscr{A}$ plus the measure q for which $q(\{i\}) = 2^{-i}$ for $i = 1, 2, \cdots$, and let P be the smallest set of probability measures on $\mathscr{A}$ which includes P_0 and satisfies Axiom 0 under the assumption that $1 \prec 2 \prec 3 \prec \cdots$. With q^i the measure which has $q^i(\{k\}) = 0$ for $k < i$ and $q^i(\{k\}) = 2^{i-1-k}$ for $k \geqq i$, so that $q^1 = q$, q^2 is the conditional of q on $\{2, 3, \cdots\}$, and so forth, P is the set of all measures which can be expressed in the form $\lambda_0 p' + \sum \{\lambda_i q_i : i \in I\}$ such that I is a finite subset of X, $\lambda_i \geqq 0$ for each $i \in I \cup \{0\}$, $\sum \lambda_i = 1$, and $p' \in P_s$, because such measures must be in P and the set of all such measures is closed under finite convex combinations and the formation of conditional measures on preference intervals in X. It is easily checked that P_0 is a maximal linearly independent subset of P, and that for every $p \in P$ which is not in P_0 there is an essentially unique representation of the form

$$\alpha_1 p + \textstyle\sum_{i=2}^{n} \alpha_i p_i = \sum_{j=1}^{m} \beta_j r_j \,, \tag{6}$$

where $n \geqq 1$, $m \geqq 1$, $\alpha_i > 0$ and $\beta_j > 0$ for all i and j, $\sum \alpha_i = \sum \beta_j = 1$, and $p_2, \cdots, p_n, r_1, \cdots, r_m \in P_0$.

Now define $v(i^*) = u(i) = i$ for each $i \in X$, define $v(q) = 3$, extend v linearly by (6) to all of P and take $p \succ p'$ iff $v(p) > v(p')$, for all $p, p' \in P$. Then, by construction and Lemma 1, Axioms 0 through 3 hold. Since $E(u, q) = 2$, $v(q) > E(u, q)$ and it is easily checked that $v(q^i) > E(u, q^i)$ for all i. Therefore, using the λ representation of the preceding paragraph, $v(p) \geqq E(u, p)$ for all $p \in P$. Then, if $p(A) = 1$, $p' \in P$ and $x^* \succ p'$ for all $x \in A$, $v(p) \geqq E(u, p)$ and $E(u, p) > v(p')$ since $u(x) = v(x^*) > v(p')$ for all $x \in A$, so that $v(p) > v(p')$, or $p \succ p'$. And if $p(A) = 1$, $p' \in P$ and $p' \succ x^*$ for all $x \in A$, then $v(p') > E(u, p) \geqq v(p)$, or $p' \succ p$. Hence Axiom 4 holds. However, $v(q^i) > E(u, q^i)$ for each $i \in X$ so that it is not true that $v(p) = E(u, p)$ for all $p \in P$.

To be more explicit about an important aspect of Example 1, a little algebraic manipulation shows that $v(q^{i+1}) - v(i^*) = 2^i + 2$ and hence that $q(\{i+1, i+2, \cdots\})[v(q^{i+1}) - v(i^*)] = 1 + 2^{-i+1}$, which does not approach zero as i gets large. In effect, the failure of the latter quantity to go to zero as $i \to \infty$ is the reason that it is not true that $E(u, q) = v(q)$. In more general terms, we do not get the desired result when u is unbounded above and $p((y, \infty)) > 0$ for all $y \in X$ unless $\inf p((y, \infty))[v(p_{(y,\infty)}) - u(y)] = 0$.

To formulate an axiom which will accomplish this, we shall let P^+ be the subset in P of measures which are bounded below but have 'upper preference tails', and let P^- be the measures in P which are bounded above but have 'lower preference tails'. Formally,

$$P^+ = \{p \in P : p([x, \infty)) = 1 \text{ for some } x \in X \text{ and } p((y, \infty)) > 0 \text{ for all } y \in X\},$$

$$P^- = \{p \in P : p((-\infty, x]) = 1 \text{ for some } x \in X \text{ and } p((-\infty, y)) > 0 \text{ for all } y \in X\}.$$

AXIOM 5. (a) *If $p \in P^+$, if $p_1 \succ p_0$ with $p_0, p_1 \in P_s$, and if $p((-\infty, y]) > 0$ for some $y \in X$, then there is a $y \in X$ such that $p((-\infty, y])p_1 + p((y, \infty))y^* \gtrsim p((-\infty, y])p_0 + p((y, \infty))p_{(y,\infty)}$.*

(b) *If $p \in P^-$, if $p_1 \succ p_0$ with $p_0, p_1 \in P_s$, and if $p([y, \infty)) > 0$ for some $y \in X$, then there is a $y \in X$ such that $p((-\infty, y))p_{(-\infty,y)} + p([y, \infty))p_1 \gtrsim p((-\infty, y))y^* + p([y, \infty))p_0$.*

Since Axiom 5 is not as elegant as one might desire and since it has not been discussed in detail elsewhere, it deserves a few words. We shall focus on Axiom 5(a) since interpretations of Axiom 5(b) are similar.

Note first that the hypotheses of Axiom 5(a) require $p((-\infty, y]) > 0$ for some $y \in X$, which is automatic when p is countably additive and $\mathscr{A}$ is a Borel algebra. The reason for requiring $p((-\infty, y]) > 0$ in the finitely-additive context is as follows. If P^+ contains a measure p for which $p((y, \infty)) = 1$ for all $y \in X$, then Theorem 1(b) and (4) require u to be bounded above, and if $u(y) < \sup u(X)$ for all $y \in X$ then $\sup u(X) = E(u, p) > u(y)$ for all $y \in X$ so that $p \succ y^*$ for all $y \in X$. But if $p \succ y^*$ for all y and if $p((y, \infty)) = 1$ for all y then it is impossible to get $p((-\infty, y])p_1 + p((y, \infty))y^* \gtrsim p((-\infty, y])p_0 + p((y, \infty))p_{(y,\infty)}$, and hence the conclusion of the axiom fails.

Since it is easily seen that Axiom 5(a) is implied by previous axioms when u is bounded above, suppose in this paragraph that u is unbounded above. Then, with $p \in P^+$, $p((-\infty, y]) > 0$ for some $y \in X$, and p_1 preferred to p_0, Axiom 5(a) says that Gamble 1 is preferred or indifferent to Gamble 2 for some $y \in X$, where:

Gamble 1 gives p_1 with probability $p((-\infty, y])$,
or y with probability $p((y, \infty)) = 1 - p((-\infty, y])$

Gamble 2 gives p_0 with probability $p((-\infty, y])$,
or $p_{(y,\infty)}$ with probability $p((y, \infty))$.

In a manner of speaking, this says that the relative advantage of getting $p_{(y,\infty)}$, which gives something preferred to y with probability 1, over getting y as a sure thing, can be outweighted by getting p_1 instead of p_0, taking account of the mixing probabilities $p((-\infty, y])$ and $p((y, \infty))$, at least for some y for which $u(y)$ is very large. Now if $p((y, \infty)) \to \alpha$ with $0 < \alpha < 1$ then, as seen from Theorem 1 (b) and (4), u must be bounded above under our earlier axioms. Hence, under the temporary presumption of unbounded u, $p((y, \infty)) \to 0$ as $u(y)$ gets large. Then a preference for Gamble 1 over Gamble 2 may not seem unreasonable for some y with $p((y, \infty))$ very near to zero, since in this case p_1 is almost a sure thing under Gamble 1, p_0 is almost a sure thing under Gamble 2, and if in fact the conditioning event with probability $p((y, \infty))$ does occur then, although Gamble 1 gives y and Gamble 2 gives something better than y, y may be 'high enough' in the preference order to provide a suitably attractive 'prize'. But note that this must hold regardless of how close together $E(u, p_1)$ and $E(u, p_0)$ are so long as $E(u, p_1) > E(u, p_0)$.

Despite these arguments, I would hesitate to say that Axiom 5 seems as intuitively reasonable (apart from a host of psychological problems that I shall not pursue here) as Axioms 1 through 4. Nevertheless, if we are interested in 'justifying' the expected utility model with finite expected utilities in the context of Axiom 0 without necessarily requiring utility to be bounded, then Axiom 5 or something similar to it must be adopted so long as P contains unbounded measures.

LEMMA 3. *Suppose Axioms* 0 *through* 4 *hold. Then* $E(u, p) = v(p)$ *for all* $p \in P$ *if and only if Axiom* 5 *holds.*

PROOF. If u is bounded above then Axiom 5 (a) is easily seen to hold, and $E(u, p) = v(p)$ for all $p \in P^+$ by Theorem 1 (a). Assume then that u is unbounded above and P^+ is not empty. To verify the necessity of Axiom 5 (a) suppose that $p \in P^+$, $p((-\infty, y]) > 0$ for some $y \in X$, $p_1 \succ p_0$ with $p_0, p \in P_s$, and $E(u, q) = v(q)$ for all $q \in P$. Then (3) and (4′) imply $\lim \{p((-\infty, y])\, E(u, p_{(-\infty,y]}) + p((y, \infty))u(y)\} = v(p) = p((-\infty, y])E(u, p_{(-\infty,y]}) + p((y, \infty))v(p_{(y,\infty)})$, so that $\lim \{p((y, \infty))[u(y) - v(p_{(y,\infty)})]\} = 0$. Hence, since $v(p_1) - v(p_0) > 0$ and $p((-\infty, y]) > 0$ for some y, there is a $y \in X$ with $p((-\infty, y]) > 0$ such that

$$p((-\infty, y])[v(p_1) - v(p_0)] + p((y, \infty))u(y) > p((y, \infty))v(p_{(y,\infty)}) \, .$$

Then transposition of $p((-\infty, y])v(p_0)$ and the use of (2) and (3) imply the conclusion of Axiom 5 (a).

On the other hand, with u unbounded above and P^+ not empty, suppose Axioms 0 through 5 (a) hold and $p \in P^+$ has $p((-\infty, y]) > 0$ for some y. Then, since the positive difference $v(p_1) - v(p_0)$ can be made arbitrarily small by the choice of p_0 and p_1 from P_s, Axiom 5 (a) in conjunction with (2), (3), and $p_{(\infty,y)} \gtrsim y^*$ yields $\inf \{p((y, \infty))[v(p_{(y,\infty)}) - u(y)]\} = 0$. Then, since $E(u, p_{(-\infty,y]}) = v(p_{(-\infty,y]})$ by Theorem 1 (a), $v(p) = E(u, p)$ follows from (4) and (5).

In a similar fashion, Axiom 5(b) is necessary for the finite expected utility representation, and Axioms 0 through 4 along with Axiom 5(b) imply $E(u, p) = v(p)$ for all $p \in P^-$.

Finally, letting P' be all measures in P which are in neither P^+ nor P^-, if $p \in P$ does not fall under the condition in part (a) of Theorem 1 then it can be written as a convex combination of measures in P' and P^+, or in P' and P^-, or in P^- and P^+, and $E(u, p) = v(p)$, given Axioms 0 through 5, then follows from (3), the linearity of $E(u, \cdot)$, Theorem 1 (a), and the results for P^- and P^+ established in the present proof. □

As a general consequence of the preceding lemmas, Theorem 1, and the observed necessity of the preference axioms for the finite expected utility model, we obtain the following summary theorem.

THEOREM 2. *Suppose Axiom* 0 *holds. Then there exists* $u : X \to Re$ *for which* $E(u, p)$ *is well defined and finite for all* $p \in P$ *and such that* (1) *holds for all* $p, q \in P$, *if and only if Axioms* 1 *through* 5 *hold.*

Before going on to countably-additive measures, we shall briefly compare Ledyard's approach [9] to Theorem 2. As mentioned previously, Ledyard considers first the case in which every $p \in P$ is bounded. He also assumes most of our Axiom 0 (except for the conditional measures part) and Axioms 1 through 3. Instead of using an axiom like Axiom 4 in his main development, he defines distribution functions for the measures $p, q, \cdots$ based on u on X [viz., $F_p(r) = p(\{x \in X : u(x) \leq r\})$], defines a pseudo-metric ρ on the distribution functions [viz., $\rho(F_p, F_q) = \int |F_p(r) - F_q(r)| \, dr$], and then assumes that $\{p \in P : p \succ q\}$ and $\{p \in P : q \succ p\}$ are contained in the topology on P induced by ρ, for each $q \in P$. His interpretation of this axiom is that "whenever two measures imply almost the same distribution on the indifference classes of X, their utility is almost the same" (page 797). Although this delivers the desired conclusions (including continuity of v on P with respect to ρ), it seems a bit roundabout in comparison with Axiom 4.

Late in his paper, Ledyard suggests the assumption of finite expected utilities as an alternative to the boundedness assumption for measures in P and shows how his previous theory would be modified in this setting. In this context he does not require an additional axiom like our Axiom 5 because the contingency that Axiom 5 was designed to handle, as illustrated by Example 1, is taken care of by the continuity of v on P with respect to ρ. Thus, Ledyard's topological axiom applied in the context of Example 1 would require $v(q) = 2$ and hence $E(u, q) = v(q)$.

5. Countable additivity. To conclude our discussion we first note how Theorem 2 can be modified when all measures are countably additive, and then compare this modification to DeGroot's theory [4].

If $\mathscr{A}$ is a Borel algebra and every measure in P is countably additive then, as discussed in Chapter 10 of [6], Axiom 4 can be replaced in Theorem 1 by

AXIOM 4′. *If* $p \in P$, $A \in \mathscr{A}$, $p(A) = 1$ *and* $y \in X$, *then* $p \gtrsim y^*$ *if* $x^* \gtrsim y^*$ *for all* $x \in A$, *and* $y^* \gtrsim p$ *if* $y^* \gtrsim x^*$ *for all* $x \in A$.

In addition, the rather cumbersome Axiom 5 can be replaced in Theorem 2 by the more obvious

AXIOM 5′. *If* $p_0 \in P_s$ *and* $p \in P$, *then* $p_{(-\infty,y]} \gtrsim p_0$ *for some* $y \in X$ *if* $p \succ p_0$, *and* $p_0 \gtrsim p_{[y,\infty)}$ *for some* $y \in X$ *if* $p_0 \succ p$.

This simply says that if $p \succ p_0$ then some 'upper truncation' of p will be at least as good as p_0, and if $p_0 \succ p$ then p_0 will be at least as good as some 'lower truncation' of p. Examples show that Axiom 5′ is not a necessary condition for (1) in the finitely-additive context, but it is necessary in the countably-additive context. For in the latter context, if p is bounded above then the first part of Axiom 5′ holds when (1) holds for all $p, q \in P$, and if p is unbounded above then, given (1) with finite expected utilities,

$$E(u, p) = \lim E(u, p_{(-\infty,y]}) \tag{7}$$

so that $E(u, p) > E(u, p_0)$ implies $E(u, p_{(-\infty,y]}) \geqq E(u, p_0)$ for some $y \in X$. The necessity of the latter part of Axiom 5′ follows in like manner.

To verify the sufficiency of Axiom 5′ in the countably-additive setting it will suffice to show that $E(u, p) \geqq v(p)$ when $p \in P^+$ and that $v(p) \geqq E(u, p)$ when $p \in P^-$, for then, by Theorem 1, $E(u, p) = v(p)$ for all $p \in P^- \cup P^+$, and $E(u, p) = v(p)$ then follows for all $p \in P$ as in the final paragraph of the proof of Lemma 3. Given $p \in P^+$, (7) follows from (4) or (4′) in the present context. Contrary to the desired result, if $v(p) > E(u, p)$ then $v(p) > \alpha > E(u, p_{(-\infty,y]})$ for some real number α and all $y \in X$; then, since $u(x_2) > \alpha > u(x_1)$ for some $x_1, x_2 \in X$, there is a $p_0 \in P_s$ with $E(u, p_0) = \alpha$. But this gives $p \succ p_0$ and $p_0 \gtrsim p_{(-\infty,y]}$ for all $y \in X$, by Lemma 1 and Theorem 1 (a), thus contradicting the first part of Axiom 5′. In similar fashion, $v(p) \geqq E(u, p)$ for $p \in P^-$ follows from the latter part of Axiom 5′. Thus we obtain

THEOREM 3. *Suppose Axiom* 0 *holds,* $\mathscr{A}$ *is a Borel algebra and every measure in* P *is countably additive. Then there exists* $u : X \to Re$ *for which* $E(u, p)$ *is well defined and finite for all* $p \in P$ *and such that* (1) *holds for all* $p, q \in P$, *if and only if Axioms* 1, 2, 3, 4′ *and* 5′ *hold.*

In DeGroot's theory, $\mathscr{A}$ is taken to be a Borel algebra of subsets of X which contains every singleton subset of X and every closed preference interval in X. P_b is then defined to be the set of all countably-additive and bounded probability measures on $\mathscr{A}$. In addition to axioms for $\succ$ on P_b which are equivalent to Axioms 1, 2 and 3 applied to P_b, he uses an assumption (U_3, page 106) which guarantees that the utility function u on X is measurable with respect to $\mathscr{A}$ (which is accomplished by our Axiom 0) along with an axiom (U_4, page 108) which states that $p \sim \beta x_2^* + (1 - \beta)x_1^*$ when $p([x_1, x_2]) = 1$ and $\beta = \int_{[x_1,x_2]} \alpha(x)\, dp(x)$, where $\alpha(x)$ for $x \in [x_1, x_2]$ is defined as a number in [0, 1] for which $x^* \sim \alpha(x)x_2^* + [1 - \alpha(x)]x_1^*$. The latter axiom, which has a fairly straightforward interpretation but seems to me to be a bit less transparent than Axiom 4 or 4′, is used instead of something like Axiom 4′ to obtain the finite expected-utility representation for $\succ$ on P_b.

To complete his theory, DeGroot then defines P_e to be the set of all countably-additive measures on $\mathscr{A}$ which are integrable with respect to u on X as obtained in the P_b setting. His versions of Axioms 1, 2 and 3 are then extended to apply to P_e, and two final assumptions are introduced. The first of these (U_5, page 110) says that $p \gtrsim q$ if $p, q \in P_e$ and $p([x, \infty)) = q((-\infty, x]) = 1$ for some $x \in X$. This is proposed in a spirit similar to Axiom 4′ and follows directly from that axiom and the transitivity of $\gtrsim$. Moreover, DeGroot's U_5 clearly implies Axiom 4′ for $P = P_e$ so that U_4 in the preceding paragraph becomes superfluous at this point. Thus, U_5 and Axiom 4′ for $P = P_e$ are almost equivalent.

DeGroot's final assumption (U_6, page 112) serves much the same purpose as Axiom 5′ but is a bit more awkward to state. Half of U_6 says that if $y_1 \lesssim y_2 \lesssim y_3 \lesssim \cdots$ is such that for every $x \in X$ there is some y_i such that $x \lesssim y_i$, and if $p, q \in P_e$ and $p_{(-\infty, y_i]} \lesssim q$ for all i larger than some n, then $p \lesssim q$. The other half of U_6 is the dual of this.

Based on these comparisons we see that the axioms used in Theorem 3 for countably-additive measures are very similar to the assumptions made by DeGroot to arrive at the finite expected-utility representation for P_e. Thus, apart from some minor differences concerning the nature of $\mathscr{A}$ and the statements of the axioms, the difference between our two approaches lies in the two-step versus the one-step procedure as discussed in the introduction.

REFERENCES

[1] Arrow, K. J. (1958). Bernoulli utility indicators for distributions over arbitrary spaces. Technical report No. 57, Department of Economics, Stanford University.

[2] Arrow, K. J. (1974). The use of unbounded utility functions in expected-utility maximization: response. *Quart. J. Econ.* **88** 136–138.

[3] Blackwell, D. and Girshick, M. A. (1954). *Theory of Games and Statistical Decisions.* Wiley, New York.

[4] DeGroot, M. H. (1970). *Optimal Statistical Decisions.* McGraw-Hill, New York.

[5] Fishburn, P. C. (1967). Bounded expected utility. *Ann. Math. Statist.* **38** 1054–1060.

[6] Fishburn, P. C. (1970). *Utility Theory for Decision Making.* Wiley, New York.

[7] Fishburn, P. C. (1975). Unbounded utility functions in expected utility theory. *Quart. J. Econ.* (forthcoming).

[8] Herstein, I. N. and Milnor, J. (1953). An axiomatic approach to measurable utility. *Econometrica* **21** 291–297.

[9] Ledyard, J. O. (1971). A pseudo-metric space of probability measures and the existence of measurable utility. *Ann. Math. Statist.* **42** 794–798.

[10] Menger, K. (1934). Das Unsicherheitsmoment in der Wertlehre. *Zeitschrift für Nationalo-ekonomie* **5** 459–485. (Trans. (1967) by W. Schoellkopf as The Role of Uncertainty in Economics, *Essays in Mathematical Economics* (211–231), M. Shubik, ed. Princeton Univ. Press.)

[11] Ryan, T. M. (1974). The use of unbounded utility functions in expected-utility maximization: comment. *Quart. J. Econ.* **88** 133–135.

[12] von Neumann, J. and Morgenstern, O. (1944). *Theory of Games and Economic Behavior.* Princeton Univ. Press.

Room 310 L, Business Administration Bldg.
The Pennsylvania State University
University Park, Pennsylvania 16802