Econometrica, Vol. 39, No. 2 (March, 1971)

CONDITIONAL EXPECTED UTILITY

BY R. DUNCAN LUCE AND DAVID H. KRANTZ[1]

This paper describes empirical laws (formulated abstractly, as axioms) that lead to simultaneous measurement of utility and subjective probability under circumstances where decisions delimit which states of nature may occur.

1. INTRODUCTION

OUR PURPOSE is to formulate a qualitative theory of conditional decisions that both admits a representation in terms of conditional expected utility and is at a level of generality comparable to Savage's [**11**] unconditional theory. The fact that conditional theories may have important advantages over unconditional ones (see Sections 4 and 5) seems not to have been generally recognized; indeed, so far as we know, such theories have only been explicitly discussed twice before. In 1964, Fishburn [**5**] suggested that the conditional formulation might be the more satisfactory, stating the sort of representation that one would want to prove, and comparing the conditional and unconditional representations in special cases. He did not, however, provide an axiom system sufficient to establish the conditional representation, and few readers seem to have realized that his proposals were novel. Later, and independently, Pfanzagl [**9**] offered an axiomatic theory for simple (two-component) gambles and two-component, non-independent compositions of simple gambles; he did not generalize his results to general conditional decisions. This we do. Very roughly, our results generalize Pfanzagl's in the same way as Savage's generalize Davidson and Suppes' [**2**] axiomatization of Ramsey's [**10**] ideas.

We present two somewhat different axiom systems that lead to (essentially) the same representation. Although the first is surely the better of the two, we provide a complete proof only for the second. The reason is that a full proof of the first result, which rests upon an unpublished version of n-dimensional additive conjoint measurement, is exceedingly long. Proofs of these two results will ultimately be given in detail (Krantz et al. [**7**]), and they are only announced here. The version of conditional expected utility that we prove, like Pfanzagl's result to which it is

[1] Work on this paper was initiated when Luce was a National Science Foundation Senior Post-doctoral Fellow at the Center for Advanced Study in the Behavioral Sciences, Stanford, California, 1966–67. During that period, which led to the formulation of Definition 1* and the proof of Theorem 1*, he had extended and stimulating discussions about expected utility with Patrick Suppes, to whom he is deeply indebted. Krantz arrived at the more satisfactory formulation given in Definition 1 and the wholly different proof of Theorem 1. He was supported in part by NSF grant GB 4947 to the University of Michigan, and greatly appreciates the hospitality of the Center for Advanced Study. Final preparation of the paper was supported in part by NSF grant GB 6536 to the University of Pennsylvania.

The first critical audience for these ideas was an informal seminar on measurement, which met irregularly at Berkeley and Stanford during 1966–67 and consisted of Ernest Adams, R. L. Causey, William Craig, John Harsanyi, A. A. J. Marley, R. D. Luce, Fred Roberts, and Patrick Suppes. Their comments helped to mold the presentation. In addition, J. Aczel, K. R. MacCrimmon, and Amos Tversky have commented critically and helpfully on an earlier draft of the material.

reduced, is faulted by having unduly restrictive assumptions, but the proof is of a more reasonable length when Pfanzagl's theorem is assumed.

2. CONDITIONAL DECISIONS

To the untutored—those not conversant with statistical decision theory—a conditional formulation of decision problems seems natural, and even the tutored have to struggle not to think in conditional terms. Suppose that you are committed to make a trip from H (here) to T (there) and that travel by commercial aircraft (A), bus (B), or your own car (C) are the feasible alternatives from which you must choose. When considering A—flying—various subevents with distinct consequences and risks must be considered: the plane may leave H and arrive at T on schedule, without any surprises; it may arrive at T some x hours late with varying consequences for the purpose of the trip, which depend upon the value of x; it may never arrive at T, which in turn breaks down into an arrival at another airport, more or less near T, or a crash, more or less fatal; and so on. Events outside the conditioning event A are irrelevant to the evaluation of that decision and, in fact, some, such as C (you driving your car from H to T) simply will not occur if you elect to fly.

Let $\mathscr{C}$ denote the set of all consequences that can arise from any of the decisions under consideration; then a particular decision conditional on the event A is simply a function from A into the set $\mathscr{C}$. For the possibilities outside A, the function is not defined. Since it is essential to make explicit the domain of conditional decisions, we write f_A for a typical decision conditional on A. Similarly, a decision involving the bus trip would be a function g_B from B into $\mathscr{C}$.

A decision problem, then, is to select one from among a set $\mathscr{D}$ of conditional decision f_A, g_B, h_C, etc. Notice that such a choice not only determines the contingencies with which various consequences arise, but it actually determines which event, A, B, or C, will occur. In other words, the decision maker controls to some extent the environment in which the consequences occur as well as the probability distribution of their occurrence. We shall suppose that between any two conditional decisions, a person prefers one to the other or he is indifferent between them, and so there is a binary ordering $\succsim$ over $\mathscr{D}$. Our goal is to axiomatize both $\mathscr{D}$ and this ordering in such a way that we can construct a utility function v over $\mathscr{C}$ and a conditional probability function P over events so that the ordering is represented by conditional expected utility, i.e., by

$$E[v(f_A)|A] = \int_A v[f_A(x)]\, dP(x|A),$$

where the integral is defined in some appropriate way.

This model differs from Savage's in two related ways. First, he supposed that all decisions have the same domain, namely, the set of all "states of nature." Second, he showed that a single (unconditional) probability function entered the calculation of all expectations. In other words, the choice of a decision does not affect the probabilities of the states of nature. Stated so baldly, the model seems strange since, almost by definition, decisions do delimit possibilities. Savage, of course, would not deny this, but he assumed that all such limitations can be encompassed by different assignments of consequences to an appropriate partition into states of nature. As this is by no means obviously true, we show in Section 5 that, for the finite case, any conditional model with the above representation can be restated in an equivalent unconditional form, and vice versa. So from one point of view, it does not matter which model one uses; however, the conditional one is usually much the simpler of the two and, in our opinion, the more natural. Moreover, it has other representations than the one stated above, and these seem of value since they admit that events may have utility independent of an arbitrary assignment of consequences to them.

3. PRIMITIVES, AXIOMS, AND REPRESENTATION THEOREMS

In this section we shall define a relational structure called a *conditional decision structure*. This consists of four undefined sets and an undefined relation (the primitives) which satisfy certain axioms. The interpretations of the primitives are as sets of events, null events, consequences, and conditional decisions, and a relation of preference between pairs of decisions. The axioms fall into two categories. The structural ones (Axioms 1, 8, and 9) are not so much constraints on preference behavior as a requirement that the total structure under consideration be sufficiently rich. The other six axioms (2–7) constitute our definition of rational preference behavior. Like Savage's rationality principles, they are mostly not transparent at first sight, but they become reasonably compelling as normative principles once their meanings are grasped.

Following this definition, we present some brief comments on the axioms; special attention is paid to their intuitive rational character. We then state the most general representation and uniqueness results as Theorem 1. Finally we indicate how the axioms must be modified so that the proof can be based on Pfanzagl's theorem, and we state the modified results as Theorem 1*.

The system has the following primitives: *events*—an algebra $\mathscr{E}$ (i.e., $\mathscr{E}$ is closed under unions and complements) of subsets of a given set X of possible chance outcomes; *null events*—a subset $\mathscr{N}$ of $\mathscr{E}$, including at least the empty set $\varnothing$, that is characterized by the axioms and that will be the events assigned probability 0 in the representation; *consequences*—an arbitrary set $\mathscr{C}$; *conditional decisions*—a set $\mathscr{D}$ of functions from non-null events (elements of $\mathscr{E} - \mathscr{N}$) into $\mathscr{C}$; and *preference ordering*—a binary relation $\succsim$ over $\mathscr{D}$.

Note that a function defined on a null event is never an element of $\mathscr{D}$. If R is null, A is non-null, $R \subset A$, and f_A is in $\mathscr{D}$, then of course, the restriction of f_A to the subset R of A is a well defined function; nevertheless, we do not consider it to be an element of $\mathscr{D}$. It seems repugnant to have to evaluate, relative to a preference ordering, decisions that are conditional on impossible events.

On the other hand, our first axiom asserts that the restriction of a conditional decision f_A to any non-null subset $B \subset A$ does constitute another conditional decision in the set $\mathscr{D}$. This axiom also asserts that $\mathscr{D}$ is closed under the union of any two disjoint non-null decisions. Specifically, if A, B are disjoint and non-null, and if $f_A, g_B \in \mathscr{D}$, then $f_A \cup g_B$, which is a function defined on $A \cup B$, is also assumed to be in $\mathscr{D}$. Except for the second axiom (which says that $\succsim$ is a weak ordering of $\mathscr{D}$), all remaining axioms are formulated in terms of unions of disjoint conditional decisions and restrictions of decisions to non-null events. It should be emphasized that these closure assumptions do not imply that $\mathscr{D}$ consists of all logically possible decisions. Specifically, in Section 4 we will exhibit an example in which $\mathscr{D}$ does not include any constant decisions.

DEFINITION 1: $\langle X, \mathscr{E}, \mathscr{N}, \mathscr{C}, \mathscr{D}, \succsim\rangle$ is a *conditional decision structure* if for all $A, B \in \mathscr{E} - \mathscr{N}, R, S \in \mathscr{E}$, and all $f_A, f_A^{(i)}, f_{A \cup B}, f_{A \cup R}, g_B, g_B^{(i)}, h_A^{(i)}, k_B^{(i)} \in \mathscr{D}$, the following nine axioms are satisfied.

AXIOM 1: (*i*) *If* $A \cap B = \varnothing$, *then* $f_A \cup g_B \in \mathscr{D}$; (*ii*) *if* $B \subset A$, *then* f_A *is restricted to* $B \in \mathscr{D}$.

AXIOM 2: $\succsim$ *is a weak ordering of* $\mathscr{D}$.

AXIOM 3: *If* $A \cap B = \varnothing$ *and* $f_A \sim g_B$, *then* $f_A \cup g_B \sim f_A$.

AXIOM 4: *If* $A \cap B = \varnothing$, *then* $f_A^{(1)} \succsim f_A^{(2)}$ *if and only if* $f_A^{(1)} \cup g_B \succsim f_A^{(2)} \cup g_B$.

AXIOM 5: *If* $A \cap B = \varnothing$, $f_A^{(i)} \sim g_B^{(i)}$, $i = 1, 2, 3, 4$, $f_A^{(1)} \cup k_B^{(1)} \sim f_A^{(2)} \cup k_B^{(2)}$, *and* $h_A^{(1)} \cup g_B^{(1)} \sim h_A^{(2)} \cup g_B^{(2)}$, *then* $f_A^{(3)} \cup k_B^{(1)} \succsim f_A^{(4)} \cup k_B^{(2)}$ *if and only if* $h_A^{(1)} \cup g_B^{(3)} \succsim h_A^{(2)} \cup g_B^{(4)}$.

AXIOM 6: *If* $A \cap B = \varnothing$, *N is a sequence of consecutive integers, not* $g_B^{(1)} \sim g_B^{(2)}$, *and* $f_A^{(i)} \cup g_B^{(1)} \sim f_A^{(i+1)} \cup g_B^{(2)}$ *for all* $i, i + 1 \in N$, *then either N is finite or* $\{f_A^{(i)} | i \in N\}$ *is unbounded.*

AXIOM 7: (*i*) *If* $R \in \mathscr{N}$ *and* $S \subset R$, *then* $S \in \mathscr{N}$; (*ii*) $R \in \mathscr{N}$ *if and only if, for all* $f_{A \cup R} \in \mathscr{D}$, $f_{A \cup R} \sim f_A$, *where* f_A *is the restriction of* $f_{A \cup R}$ *to* A.

AXIOM 8: (*i*) $\mathscr{E} - \mathscr{N}$ *contains at least three pairwise-disjoint elements*; (*ii*) $\mathscr{D}/\sim$ *contains at least two distinct equivalence classes.*

AXIOM 9: (*i*) *If* A *and* g_B *are given, then there exists* $h_A \in \mathscr{D}$ *for which* $h_A \sim g_B$; (*ii*) *if* $A \cap B = \varnothing$ *and* $h_A^{(1)} \cup g_B \succsim f_{A \cup B} \succsim h_A^{(2)} \cup g_B$, *then there exists* $h_A \in \mathscr{D}$ *such that* $h_A \cup g_B \sim f_{A \cup B}$.

The three structural axioms, 1, 8, and 9, yield rather subtle constraints on the type of situation to which the theory applies. We believe, however, that these constraints are satisfied, at least approximately, in a wide variety of situations.

The simplest of these to discuss is Axiom 8, which assures that the events and decisions are not wholly trivial. It is much weaker than Savage's structural conditions: Axiom 8(ii) is the same as Savage's P5, whereas 8(i) follows from his P6, which, in fact, assures that there are indefinitely fine partitions of his basic set of states. Unlike Savage's axioms, ours have models in which X is finite. Axiom 8(i) can even be weakened to assert that there are at least two (rather than three) disjoint non-null elements, but then Axiom 4 has to be replaced by a stronger, and less natural assertion.

Axiom 1 involves some special difficulties in real situations, although not in laboratory studies of decision making. Suppose, for example, that $C = \{\text{heads, tails}\}$ is the toss of a coin and that $D = \{1, 2, 3, 4, 5, 6\}$ is the roll of a die. Let h_C be some payoff scheme assigning one consequence to heads, another to tails, and k_D a payoff scheme contingent on D. The experimenter can invent a game in which either C or D (but not both) will occur, and where the subject is uncertain which one will occur. Then $h_C \cup k_D$ may be considered the payoff scheme for this game. The quantities $P(C|C \cup D)$ and $P(D|C \cup D)$, which arise in the representation of Theorem 1, represent the subject's subjective probabilities of C and D, given that he chooses to play such a game. Thus, it is perfectly natural to evaluate the utility of the whole game, $u(h_C \cup k_D)$, in terms of the expectation of the utilities $u(h_C)$ and $u(k_D)$ of the subgames h_C and k_D:

$$u(h_C \cup k_D) = u(h_C)P(C|C \cup D) + u(k_D)P(D|C \cup D).$$

Moreover, the experimenter can test this expectation principle by offering the subject various choices between different sorts of games, including $h_C, k_D, h_C \cup k_D$, etc., and from the choices made he can compute utilities and subjective probabilities.

For real-world choices, matters are less simple. Returning to our earlier example, what is $f_A \cup g_B$? It means a trip from H to T, with uncertainty as to mode of transport—airplane or bus. Since there is no experimenter, the uncertainty can only arise from random events in nature. If you are considering whether to go from H to T tomorrow and do not hold a plane reservation, the probability $P(A|A \cup B)$ may represent your subjective probability of obtaining the desired reservation. The decision to travel, if it must be taken without knowing whether or not a reservation is obtainable, is well represented by $f_A \cup g_B$ and is well evaluated by $u(f_A)P(A|A \cup B) + u(g_B)P(B|A \cup B)$. But in this situation, it is difficult to infer $u(f_A)$, since there is no experimenter who, by fiat, can offer you f_A as opposed to other choices. At the moment of decision, $f_A \cup g_B$ is an available decision, but its restriction to A, namely f_A, is not—violating Axiom 1(ii). On the other hand, if you already hold a plane reservation, there may be no uncertainty. You simply evaluate $u(f_A)$, $u(g_B)$, and u (stay at home), and if one of the first two is highest, you make the trip by the appropriate means of transportation. Here, f_A, g_B are in $\mathscr{D}$, but $f_A \cup g_B$ is meaningless—violating Axiom 1(i).

To apply our theory to real-world decisions, we must therefore suppose that "natural" decisions, such as f_A, g_B above, are enriched by certain artificial ones. A game can be constructed in which either A or B will occur, but which one is not certain. The fact that a person would not ordinarily consider such a game, when contemplating travel from H to T, does not prevent us from supposing that, were he to consider it, the game $f_A \cup g_B$ would have a definite place in the ordering $\succsim$

over $\mathscr{D}$. The supposition that $\mathscr{E}$ is an algebra and that Axiom 1 holds is equivalent to the assumption that sufficiently many such games are available for potential consideration.

This gives us a clue as to what is required actually to measure utilities and subjective probabilities of real-world decisions and events. The measurer must be prepared to present for serious consideration by the decision maker some rather artificial alternatives, and the decision maker must be induced to make realistic decisions among them. The usual technique is to pose hypothetical questions. For example, if a decision maker prefers aircraft to bus to car, we might ask him whether he prefers a coin flip to decide between the aircraft and the car or whether he prefers to go by bus for sure. Suitable side payments or a willingness to serve science may induce decision makers to submit to somewhat odd, nonhypothetical choices and to abide by the resulting consequences.

At any rate, our supposition that a decision maker maintains a preference ordering $\succsim$ over a set $\mathscr{D}$, which includes both some natural alternatives and some rather artificial games, is certainly considerably weaker than Savage's assumption that he maintains an ordering over all possible functions relating states to consequences.

The status of the last structural axiom, 9, is similar to that of Axiom 1. In the laboratory, the experimenter can surely devise enough different payoff schemes h_A so that he can match any g_B, etc. A decision maker considering a trip from H to T will not ordinarily think (except, wishfully) of the unlikely payoff scheme h_A specifying (among other things) that if there are fifteen planes ahead of him on the runway for takeoff, he will receive a \$20 refund. But we can still suppose that he has a definite order of preference, or is indifferent, between such an h_A and g_B; and we can either ask him to report this preference or try to induce him to make real choices of this kind. Again, Axiom 9 seems considerably less stringent than Savage's requirement that any possible assignment of consequences to states is potentially available for consideration.

Axioms 2–7 are of a rather different character since they are logically necessary for the representation we are trying to obtain in Theorem 1. (This is not quite true for 7(ii), but is nearly true in a certain sense.) These axioms capture one view of rational decision making.

The weak ordering axiom (2) although questionable on empirical grounds (Tversky [**12**]), is unchallenged as a principle of rationality.

Axiom 3 asserts that if a decision maker is equally happy with f_A or g_B, then he is just as happy in the more uncertain condition $f_A \cup g_B$. If an individual who cannot decide between two lottery tickets is handed one at random, he should have no motivation to see which it is before deciding on its resale price.

Axiom 4 is a version of Savage's "sure-thing principle." If $f_A^{(1)}, f_A^{(2)}$ are constant decisions, there hardly can be any dispute about the principle: if $f_A^{(1)}$ is the more valuable, then it does more than $f_A^{(2)}$ to enhance g_B when combined by disjoint union. For nonconstant decisions, some rather subtle arguments have been raised against the principle (Allais [**1**], Ellsberg [**4**], MacCrimmon [**8**]). The gist of Allais' argument is that if $f_A^{(1)}$ has high variance and $f_A^{(2)}$ has low variance, then $f_A^{(1)}$ may be preferred in isolation, or in combination with a nearly valueless g_B, but that $f_A^{(2)}$ is preferred in combination with a g_B that is approximately equivalent to $f_A^{(2)}$. Although this is a fairly good descriptive statement, Savage's argument [**11**, p. 103] that this behavior is irrational seems quite convincing. It seems to us that violations of Axiom 4 are something like optical illusions: compelling, but wrong. A rational decision maker should analyze choices just as Savage recommends, thereby avoiding the illusion.

Axiom 5 is less complex conceptually than it first seems; nonetheless, it is desirable to replace it by simpler axioms. Let $\mathscr{D}_A$ denote the set of all decisions that are conditional on A. The two equations $f_A^{(1)} \cup k_B^{(1)} \sim f_A^{(2)} \cup k_B^{(2)}$ and $f_A^{(3)} \cup k_B^{(1)} \succsim f_A^{(4)} \cup k_B^{(2)}$ define an order of "differences" in $\mathscr{D}_A$: we infer that the "difference" between $f_A^{(3)}$ and $f_A^{(4)}$ is at least as large as that between $f_A^{(1)}$ and $f_A^{(2)}$. The matching relations $f_A^{(i)} \sim g_B^{(i)}$ show that the $g_B^{(i)}$ "differences" ought to coincide with corresponding $f_A^{(i)}$ "differences," and that is precisely what the axiom asserts.

Axiom 6 is an Archimedean axiom of the type used in additive measurement. It asserts that if the "difference" between $g_B^{(1)}$ and $g_B^{(2)}$ is nonzero, then an infinite chain with that spacing cannot be bounded in the ordering $\succsim$.

Axioms 5 and 6 may be easier to understand if we introduce the concept of a standard sequence used in measurement theory. If N is a sequence of consecutive integers and $A \in \mathscr{E} - \mathscr{N}$, then $\{f_A^{(i)} | i \in N, f_A^{(i)} \in \mathscr{D}_A\}$ is called a *standard sequence* if there exist $B \in \mathscr{E} - \mathscr{N}$, g_B^*, and $g_B^{**} \in \mathscr{D}_B$, such that for all i, $i + 1 \in N, f_A^{(i)} \cup g_B^{**} \sim f_A^{(i+1)} \cup g_B^*$. Axiom 5 is now equivalent to: *If* $\{f_A^{(i)}|N\}$ *and* $\{g_B^{(i)}|N\}$ *are standard sequences with* $A \cap B = \varnothing$ *and if, for some* k, $k + 1 \in N, f_A^{(k)} \sim g_B^{(k)}$ *and* $f_A^{(k+1)} \sim g_B^{(k+1)}$, *then for all* $i \in N$, $f_A^{(i)} \sim g_B^{(i)}$. By its definition, a standard sequence has successive "differences" equal. Axiom 5 asserts that if the size of the "difference" and the "phase" are the same at one point in two standard sequences, then they are equivalent everywhere. The Archimedean Axiom 6 becomes simply: *Any bounded standard sequence is finite.*

Finally Axiom 7 asserts that a subset of an impossible event is itself impossible and that impossible events are completely characterized by the property that consequences contingent on them do not affect the decisions in which they are embedded. For practical purposes, we could dispense entirely with null events, simplifying both the axiom system and the proofs, but we retain them simply because they are traditional in probability theory and we want to show how to deal with them when they exist.

THEOREM 1: *Suppose that* $\langle X, \mathscr{E}, \mathscr{N}, \mathscr{C}, \mathscr{D}, \succsim \rangle$ *is a conditional decision structure in the sense of Definition* 1. *Then there exist real-valued functions* u *on* $\mathscr{D}$ *and* P *on* $\mathscr{E}$ *such that* $\langle X, \mathscr{E}, P \rangle$ *is a finitely additive probability space and, for all* $A, B \in \mathscr{E} - \mathscr{N}$, $R \in \mathscr{E}$, *and* $f_A, g_B \in \mathscr{D}$, (*i*) $R \in \mathscr{N}$ *if and only if* $P(R) = 0$; (*ii*) $f_A \succsim g_B$ *if and only if* $u(f_A) \geqslant u(g_B)$; (*iii*) *if* $A \cap B = \varnothing$, *then*

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

Moreover, P *is unique and* u *is unique up to a positive linear transformation.*

Observe that the theorem asserts the existence of a utility function over $\mathscr{D}$, not over $\mathscr{C}$ as is usual in most theories of utility, and according to part (iii) this function acts like the expectation of a random variable relative to P. We return in Section 4 to the question of utility functions over $\mathscr{C}$.

Theorem 1 is not proved here; the main steps of the proof are indicated in Section 6. Instead, we will prove a modified result.

DEFINITION 1*: In Definition 1 *retain* Axioms 1, 2, 3, 4, 7, and 8, and *replace* 5, 6, and 9 by:

AXIOM 5*: *If* $A, B, C, D \in \mathscr{E} - \mathscr{N}$, $A \cap B = C \cap D = \varnothing$, *and* $f_A^{(i)} \sim h_C^{(i)} \succ g_B^{(i)} \sim k_D^{(i)}$, $i = 1, 2$, *then* $f_A^{(1)} \cup g_B^{(1)} \succsim h_C^{(1)} \cup k_D^{(1)}$ *if and only if* $f_A^{(2)} \cup g_B^{(2)} \succsim h_C^{(2)} \cup k_D^{(2)}$.

AXIOM 6*: $\langle \mathscr{D}, \succsim \rangle$ *is order complete, i.e., every subset of* $\mathscr{D}$ *that has an upper bound has a least upper bound.*

AXIOM 9*: (*i*) *is unchanged*; (*ii*) *if* $A \cap B = \varnothing$ *and* $f_{A \cup B}$ *and* g_B *are given, then there exists* $h_A \in \mathscr{D}$ *such that* $h_A \cup g_B \sim f_{A \cup B}$.

Axiom 5* is about as complex as Axiom 5 (eight decisions on four events instead of twelve on two events), and it serves a somewhat similar role. In particular, it permits us to define qualitative conditional probability on $\mathscr{E} \times (\mathscr{E} - \mathscr{N})$ in Definition 2. Axiom 6* substitutes for and is stronger than the Archimedean property. Part (ii) of Axiom 9* is the unrestricted, and so much stronger, version of the solvability condition invoked in Definition 1. As in Pfanzagl's theory, one further property will be assumed, but it is best stated as a separate hypothesis in Theorem 1*.

DEFINITION 2: Suppose that $\langle X, \mathscr{E}, \mathscr{N}, \mathscr{C}, \mathscr{D}, \succsim \rangle$ satisfies Definition 1*, then $\succsim$ on $\mathscr{E} \times (\mathscr{E} - \mathscr{N})$ is defined by: *if* $A, C \in \mathscr{E}$, *and* $B, D \in \mathscr{E} - \mathscr{N}$, *then* $A|B \succsim C|D$ *if and only if either* (*i*) $C \cap D \in \mathscr{N}$; (*ii*) $\bar{A} \cap B \in \mathscr{N}$; *or* (*iii*) $A \cap B, \bar{A} \cap B, C \cap D, \bar{C} \cap D \in \mathscr{E} - \mathscr{N}$ *and there exist* $f_{A \cap B}, g_{\bar{A} \cap B}, h_{C \cap D}, k_{\bar{C} \cap D} \in \mathscr{D}$ *such that* $f_{A \cap B} \sim h_{C \cap D} \succ g_{\bar{A} \cap B} \sim k_{\bar{C} \cap D}$ *and* $f_{A \cap B} \cup g_{\bar{A} \cap B} \succsim h_{C \cap D} \cup k_{\bar{C} \cap D}$.

Since we must frequently refer to events A for which both A and $\bar{A} \in \mathscr{E} - \mathscr{N}$, it is convenient to have a term for them; we call them *proper events*.

DEFINITION 3: Suppose that $\langle X, \mathscr{E}, \mathscr{N}, \mathscr{C}, \mathscr{D}, \succsim \rangle$ satisfies Definition 1*, that $A \in \mathscr{E}$, and that B is a proper event; then A is *independent* of B if and only if $A|B \sim A|\bar{B}$.

THEOREM 1*: *Suppose that* $\langle X, \mathscr{E}, \mathscr{N}, \mathscr{C}, \mathscr{D}, \succsim \rangle$ *is a conditional decision structure in the sense of Definition* 1*. *If to every proper event there is a proper event that is independent of it, then the assertions of Theorem* 1 *hold and, in addition, if* $A, C \in \mathscr{E}$ *and* $B, D \in \mathscr{E} - \mathscr{N}$, *then* $A|B \succsim C|D$ *if and only if* $P(A \cap B)/P(B) \geqslant P(C \cap D)/P(D)$.

A proof is given in Section 7.

The assumption that to each proper event there is an independent proper event, which is not needed for Theorem 1, is strong: it implies that $\langle X, \mathscr{E}, P\rangle$ is atomless. Suppose the contrary; then there is a proper atom (Axiom 8(i)) x. By assumption, there is an event B that is proper and independent of $A = X - \{x\}$, and so $P(A \cap B) = P(A)P(B)$ and $0 < P(A)$, $P(B) < 1$. Clearly, either $B - A = \varnothing$ or $\{x\}$. In the former case, $P(B) = P(A \cap B) + P(B - A) = P(A)P(B)$, and so either $P(B) = 0$ or $P(A) = 1$, both of which are impossible. In the latter case,

$$\begin{aligned} P(A \cap B) &= P(A)P(B) \\ &= P(A)[P(A \cap B) + P(B - A)] \\ &= [1 - P(x)][P(A \cap B) + P(x)], \end{aligned}$$

whence $P(A \cap B) = 1 - P(x)$, from which $P(B) = 1$, which is impossible. So the space is atomless.

4. CONSTANT DECISIONS AND THE UTILITY OF CONSEQUENCES

As we noted earlier, Theorems 1 and 1* do not assign a utility function to consequences; therefore, they do not seem to fulfill our original goal and certainly they are different in this respect from all other theories of expected utility. We plan now to examine carefully the possibility of defining utility over consequences; however, our discussion by no means exhausts the matter—there are a number of important open problems.

Crucial in all other theories of expected utility are the constant decisions. A conditional decision $f_A \in \mathscr{D}$ is called *constant* if for some $c \in \mathscr{C}$ and all $x \in A$, $f_A(x) = c$. It is frequently convenient to denote by c_A the constant decision on A having as its single value $c \in \mathscr{C}$. A conditional decision is called *finite* if its range is a finite subset of $\mathscr{C}$. Clearly a constant decision is finite. Observe that if c is in the range of f_A, the set of points mapping into c, i.e.,

$$f_A^{-1}(c) = \{x | x \in A \quad \text{and} \quad f_A(x) = c\},$$

may or may not be an event of $\mathscr{E} - \mathscr{N}$. We call f_A a *gamble* if f_A is finite and if for every $c \in f_A(A)$, $f_A^{-1}(c) \in \mathscr{E} - \mathscr{N}$. It is not difficult to see that f_A is a gamble if and only if there is a partition of A into a finite number of non-null events and a constant function exists on each of these events such that f_A is the union of these mutually disjoint constant functions. This definition of a gamble seems consistent with the usage of other authors.

Since all other theories of expected utility, in which the utility function is defined over the consequences, have rested heavily upon arguments involving gambles and constant decisions, an example of a conditional decision structure that includes no constant decisions whatsoever should make us skeptical about constructing a utility function over consequences in the general case. The following is such an example. Let $\mathscr{E}$ be the algebra that includes all open intervals and finite unions of open intervals from [0, 1]. Let P be the usual (probability) measure of length of intervals and finite unions of intervals. Let $\mathscr{N}$ consist solely of the empty set. And let $\mathscr{R}$ be the set of real numbers. If A is an open interval, then for every real α let the function $f_A(x) = \alpha + x$, $x \in A$ be in $\mathscr{D}$, and let $\mathscr{D}$ be closed under finite (disjoint) unions of such functions. Obviously, if $A \in \mathscr{E}$ and $f_A \in \mathscr{D}$, we can define the conditional expectation, $E[f_A|A]$, to be the ordinary integral $\int_A f_A(x)dx/P(A)$. Let $\succsim$ be the ordering on $\mathscr{D}$ induced by $E[f_A|A]$. It is not difficult to see that $\langle [0, 1], \mathscr{E}, \mathscr{N}, \mathscr{R}, \mathscr{D}, \succsim \rangle$ is a conditional decision structure (the last two structural axioms hold since for each non-null A and each real number β, we can choose values for the α's such that $E[f_A|A] = \beta$). Even more obviously, none of the conditional decisions in $\mathscr{D}$ is a constant.

Evidently, then, if we need constant decisions, we must build them into the structure. But even that is not enough to establish the usual representation, as we now show. Let $\langle X, \mathscr{E}, P \rangle$ be a finitely additive probability space, let $\mathscr{R}$ be the real numbers, and let $\mathscr{D}$ be the set of all gambles that can be constructed from $\mathscr{E}$ and $\mathscr{R}$. Obviously, $\mathscr{D}$ satisfies Axiom 1. Let w be any function on $\mathscr{E}$ with the property[2] that for $A, B \in \mathscr{E}$ and $A \cap B = \varnothing$,

$$(1) \qquad w(A \cup B) = w(A)P(A|A \cup B) + w(B)P(B|A \cup B).$$

Observe that since $f_A \in \mathscr{D}$ is a real-valued gamble, its expectation with respect to P exists, that for every non-null A and every real number β there is a gamble $f_A \in \mathscr{D}$ such that $E[f_A|A] = \beta$, and that the function

$$u(f_A) = E[f_A|A] + w(A)$$

has the conditional expected utility property of part (iii) of Theorem 1. Thus, if we define $\succsim$ on $\mathscr{D}$ by part (ii) of the Theorem, we have a conditional decision structure.

Except for the additive term w, this structure has the usual representation (in which the utility over the real-valued consequences is trivially the identity function). The point, however, is that the axioms, even when restricted to gambles, do not exclude the function w (and, quite possibly, more bizarre representations). The function w has a natural interpretation as the utility of the events independent of what consequences are associated with them. For dice, coins, and the like, it is reasonable to suppose that the events have no utility in and of themselves; however, were we to gamble on future weather, with one consequence associated with clear and cold weather and another with clear and warm weather, then it is less evident that the conditional event—being clear—lacks utility in and of itself. In any case, what we have shown is that the assumptions made so far still do not exclude this possibility. Specifically, if A and B are non-null, $c \in \mathscr{C}$, and the constant decison c_A and c_B are in $\mathscr{D}$, there is nothing in the theory that requires

[2] If ϕ is any finitely additive measure on $\mathscr{E} - \mathscr{N}$, then it is easy to see that for any real number α and for all $A \in \mathscr{E} - \mathscr{N}$,

$$W(A) = \frac{\phi(A)}{P(A)} + \alpha$$

has this property, and every W satisfying Equation (1) is of this form. Jeffrey [6] studied Equation (1) as a possible theory of decisions.

$c_A \sim c_B$; whereas, in the ordinary utility representation, $u(c_A) = u(c_B) = u(c)$, which implies $c_A \sim c_B$.

The above remarks suggest the following result.

THEOREM 2: *In addition to the assumptions of Theorem* 1 *or* 1*, *assume that* (*i*) *for every* $c \in \mathscr{C}$, *there exists some* $A(c) \in \mathscr{E} - \mathscr{N}$ *such that* $c_{A(c)} \in \mathscr{D}$, *and* (*ii*) *if* $c \in \mathscr{C}$, A, $B \in \mathscr{E} - \mathscr{N}$, *and* $c_A, c_B \in \mathscr{D}$, *then* $c_A \sim c_B$. *Then there exists a unique real-valued function* v *on* $\mathscr{C}$ *such that for every gamble* $f_A \in \mathscr{D}$ *with* $A \in \mathscr{E} - \mathscr{N}$, $u(f_A) = E[v(f_A)|A]$, *where* u *and* P *are the functions constructed in Theorem* 1 *and* E *is the conditional expectation with respect to* P.

Keeping in mind that when f_A is a gamble, there exists a finite partition of A, $\{A_1, A_2, \ldots, A_n\}$, where $A_i \in \mathscr{E} - \mathscr{N}$ and consequences $c^i \in \mathscr{C}$ such that $f_A = \bigcup_{i=1}^{n} c^i_{A_i}$, then

$$\begin{aligned} u(f_A) &= E[v(f_A)|A] \\ &= \sum_{i=1}^{n} v(c^i)P(A_i|A), \end{aligned}$$

which is the usual form of the expected utility property.

The proof simply involves setting $v(c) = u[c_{A(c)}]$ and calculating the expectation of a gamble using part (iii) of Theorem 1.

5. COMPARISON WITH STATISTICAL DECISION THEORY

The only other axiomatization at a comparable level of generality to this one is Savage's [**11**]. We wish, first, to compare them as axiom systems, and second, to show how in the finite case one can pass from one to the other.

As an axiom system, per se, Savage's has the two interrelated failings that not all of the axioms are stated in terms of the primitives of the system (and to do so seems to lead to a very messy system, indeed) and that several relatively complex definitions and a number of intermediate theorems must be proved before the system can be stated fully.

Aside from the niceties of axiomatization, at least three conceptual criticisms can be leveled at Savage's theory. First, the axioms include one to the effect that, for every positive integer n, the set $\mathscr{S}$ of states of nature can be partitioned into n equally probable events. Thus, the infiniteness of the system is forced onto the states of nature, with no opportunity to place it on the consequences if that seems more convenient or realistic. As we show later, however, once the initial formulation of the problem in unconditional terms has been accepted, there is probably little option about the set of states being (for all practical purposes) infinite. Second, the proof of the representation theorem makes essential use of constant decisions, i.e., those with the same fixed consequence no matter which state of nature obtains. For a complex and realistic set of states, it may not be plausible to suppose that such constant acts can be generally realized. And third, the system and the axioms are so formulated that the choice of a decision leaves the probability distribution over the states unaffected. Frequently, this is alleged to be a virtue of the formulation, in spite of the fact that the most natural way to describe most decision problems is in terms of decisions and states for which the choice of a decision makes impossible many chance occurrences. We explore this point with some care in the finite case, showing that we can always reformulate a conditional decision problem as a statistical one, but that it is usually impractical to do so.

Before turning to that discussion, let us compare the conditional decision model with the statistical one on the first three points that we have made. First, all assumptions (axioms) of the conditional theory are stated simply in terms of the primitive notions (see Definition 1). (This is not quite true of Theorem 1* which needs the defined notion of independent events.) Second, the set of events need not be infinite. And third, as was shown by an example in Section 4, no constant decision need be included among the conditional decisions. So on these three points, the conditional theory seems the more satisfactory.

When the consequences, the states of nature, and the events all form finite sets, it is easy to display (in matrix form) the differences in conceptualization between the two theories. Suppose that the χ elements of $\mathscr{C}$ are enumerated in some way as $\{c_1, \ldots, c_k, \ldots, c_\chi\}$, and that the σ elements of $\mathscr{S}$ are enumerated as $\{S_1, \ldots, S_j, \ldots, S_\sigma\}$. If a typical decision is denoted f_i, then the statistical theory takes the form

$$(2) \qquad \overset{S_j}{f_i[c_{ij}]},$$

and decisions are ordered by

$$(3) \qquad u(f_i) = \sum_{j=1}^{\sigma} v(c_{ij})Q(S_j),$$

where Q is a probability distribution over $\mathscr{S}$ and v is a real-valued function on $\mathscr{C}$. In the conditional theory, let the $\varepsilon + 1$ elements of $\mathscr{E}$ be enumerated in some fashion as $\{\varnothing, A_1, \ldots, A_l, \ldots, A_\varepsilon\}$; then the format is

$$(4) \qquad \overset{c_k}{f_i[A_{ik}]}, \qquad A_{ik} \cap A_{il} = \varnothing \qquad \text{for} \qquad k \neq l,$$

where $A_{ik} \in \mathscr{E}$ is the event under which f_i leads to c_k, and $\bigcup_{k=1}^{\chi} A_{ik}$ is the event on which f_i is conditional. Of course, some of the A_{ik} may be $\varnothing$. Decisions are ordered by

$$(5) \qquad u(f_i) = \sum_{k=1}^{\chi} v(c_k)P(A_{ik}|\textstyle\bigcup_{l=1}^{\chi} A_{il}).$$

The question is: can we pass freely from the model of Equations 2 and 3 into an equivalent one satisfying Equations 4 and 5, and vice-versa. Assuming Equations 2 and 3 first, we simply let $X = \mathscr{S}$, $\mathscr{E}$ be the power set of X, and $P = Q$. For each $c_k \in \mathscr{C}$, define $A_{ik} = \{S_j | c_{ij} = c_k\}$, and the following simple calculation shows that

Equation 5 holds:

$$\begin{aligned} u(f_i) &= \sum_{j=1}^{\sigma} v(c_{ij})Q(S_j) \\ &= \sum_{k=1}^{\chi} v(c_k) \sum_{\substack{j \\ S_j \in A_{ik}}} Q(S_j) \\ &= \sum_{l=1}^{\chi} v(c_k)Q(A_{ik}) \\ &= \sum_{k=1}^{\chi} v(c_k)P(A_{ik}|\mathscr{S}), \end{aligned}$$

which is Equation 5 with every decision conditional on $\mathscr{S}$.

To go from the conditional model to the statistical one is a bit more complicated. Assume that Equations 4 and 5 hold and, with no loss of generality, that for all $A \in \mathscr{E}$, save $A = \varnothing$, $P(A) > 0$ (if this is not true, simply drop all null atoms from X and let $\mathscr{E}$ be the resulting algebra of subsets). Define $\mathscr{S} = \Pi_{l=1}^{\varepsilon} A_l$, and enumerate the elements of $\mathscr{S}$ as $\{S_1, \ldots, S_\sigma\}$. For $1 \leqslant l \leqslant \varepsilon$ and $S_j \in \mathscr{S}$, let S_j^l denote the lth component of S_j; hence $S_j^l \in A_l$. Define

$$Q(S_j) = \sum_{l=1}^{\varepsilon} [P(S_j^l)/P(A_l)],$$

which is easily seen to be a probability distribution over $\mathscr{S}$. Note that if $B \subset A_l$, then

$$\sum_{\substack{j \\ S_j^l \in B}} Q(S_j) = P(B|A_l). \tag{6}$$

For any f_i, $\bigcup_{k=1}^{\chi} A_{ik}$ is some element of $\mathscr{E}$, say A_l. And for any S_j, its lth component, S_j^l, is in one and only one of the $A_{ik} \subset A_l$. If it is in the kth, define $c_{ij} = c_k$. Thus, we have formulated Equation 2, and assuming Equation 5, we prove Equation 3:

$$\begin{aligned} u(f_i) &= \sum_{k=1}^{\chi} v(c_k)P(A_{ik}|A_l) \\ &= \sum_{k=1}^{\chi} v(c_k)\left[\sum_{\substack{j \\ S_j^l \in A_{ik}}} Q(S_j)\right] && \text{(by Equation 6)} \\ &= \sum_{k=1}^{\chi} \sum_{\substack{j \\ S_j^l \in A_{ik}}} v(c_k)Q(S_j) \\ &= \sum_{k=1}^{\chi} \sum_{\substack{j \\ S_j^l \in A_{ik}}} v(c_{ij})Q(S_j) && \text{(definition of } c_{ij}) \\ &= \sum_{\substack{j \\ S_j^l \in A_l}} v(c_{ij})Q(S_j) \\ &= \sum_{j=1}^{\sigma} v(c_{ij})Q(S_j) && \text{(since for every } j,\ S_j^l \in A_l). \end{aligned}$$

Next, we show that $\mathscr{S}$ cannot be collapsed into a smaller set of states if any decision is a possibility. Consider treating $\{S_j, S_k\}$, where $S_j, S_k \in \mathscr{S}$, $j \neq k$, as a single state; then since $S_j \neq S_k$, for some l, $S_j^l \neq S_k^l$. The decision f_i conditional on A_l, defined by

$$f_i(x) = \begin{cases} c & \text{if} \quad x = S_j^l, \\ d & \text{if} \quad x \in A_l - \{S_j^l\}, \end{cases}$$

where $c, d \in \mathscr{C}$ and $c \neq d$, is a possibility. But in the ith row of the statistical model, c is the entry of the S_j column and d is the entry of the S_k column. Hence, the pair cannot be treated as a single state since a row (decision) and a column (state) are supposed to specify a unique consequence.

Although, in principle, the above definitions provide a way to pass from a finite conditional model to a finite statistical model, in practice it is wholly unrealistic because the number of events in $\mathscr{S}$ is, to put it mildly, excessive. We can see this as follows. Suppose that X has n elements, then since there are $\binom{n}{i}$ subsets with i elements, $\mathscr{S}$ has $\Pi_{i=1}^n i^{\binom{n}{i}}$ elements. The values for the first few integers are startling:

n:	2	3	4	5
number in $\mathscr{S}$:	2	24	20,736	5×12^{10}

6. OUTLINE OF A PROOF OF THEOREM 1

A result from the theory of additive, n-dimensional conjoint measurement is used to prove Theorem 1; it is similar to, but more general than, a theorem of Debreu [3]. Suppose that $\succsim$ is a binary relation on $\Pi_{i=1}^n \mathscr{A}_i$, where $n \geqslant 3$ and $\mathscr{A}_i$ are non-empty sets. We say that the relation is *independent* if, for every $M \subset N = \{1, 2, \ldots, n\}$, the ordering induced by $\succsim$ on $\Pi_{i \in M} \mathscr{A}_i$ for a fixed element in $\Pi_{i \in N-M} \mathscr{A}_i$ is independent of that choice. The relation satisfies *restricted solvability* if whenever

$$(b_1, \ldots, b_i^{**}, \ldots, b_n) \succsim (a_1, \ldots, a_i, \ldots, a_n) \succsim (b_1, \ldots, b_i^*, \ldots, b_n),$$

then there exists $b_i \in \mathscr{A}_i$ such that

$$(b_1, \ldots, b_i, \ldots, b_n) \sim (a_1, \ldots, a_i, \ldots, a_n).$$

Let I be a sequence of integers (positive or negative or both, finite or infinite); then a set $\{a_i^j \in \mathscr{A}_i | j \in I\}$ is a *standard sequence* if there exist $b_k^0, b_k^1 \in \mathscr{A}_k$, $k \in N - \{i\}$, such that for all $j, j+1 \in I$,

$$(b_1^1, \ldots, a_i^j, \ldots, b_n^1) \sim (b_1^0, \ldots, a_i^{j+1}, \ldots, b_n^0).$$

And, finally, when $\succsim$ is a weak order and independent, we say that component $\mathscr{A}_i$ is *essential* if $\mathscr{A}_i/\sim$ has at least two equivalence classes.

Then the following result can be proved (Krantz, et al. [7]).

THEOREM 3: *If $\succsim$ on $\Pi_{i=1}\,\mathscr{A}_i$ is an independent weak ordering for which restricted solvability holds, every bounded standard sequence is finite, and at least three components are essential, then there exist real-valued functions φ_i on $\mathscr{A}_i$, $i \in N$, such that for all $a_i, b_i \in \mathscr{A}_i$, $(a_1, \ldots, a_n) \succsim (b_1, \ldots, b_n)$ if and only if*

$$\sum_{i=1}^{n} \varphi_i(a_i) \geqslant \sum_{i=1}^{n} \varphi_i(b_i).$$

Moreover, if φ_i^* is another family of functions having the same property, then there exist numbers $\alpha > 0$ and β_i such that for all $i \in N$, $\varphi_i^* = \alpha\varphi_i + \beta_i$.

We now turn to Theorem 1. Let A_i, $i \in N = \{1, 2, \ldots, n\}$, $n \geqslant 2$, be any set of pairwise disjoint events from $\mathscr{E} - \mathscr{N}$. Let $\mathscr{A}_{A_i}$ denote the set of decisions conditional on A_i and define $\succsim$ on $\Pi_{i\in N}\,\mathscr{A}_{A_i}$ by $f_1 \cup \ldots \cup f_n \succsim g_1 \cup \ldots \cup g_n$. For $n \geqslant 3$, one proves that the hypotheses of Theorem 3 are fulfilled, and for $n = 2$, one uses Axiom 8(i) to convert the problem into a three-component case. The conclusion is that functions φ_i exist that are additive over the components. A variety of functions on $\mathscr{A}_A$ will arise from various partitions of X. One next shows that these are all related to one another by positive linear transformations.

Select any two decisions with $f^1 \succ f^0$ to serve as unit and zero (they exist by Axiom 8(ii)). Choose that normalization u_A of the "additive" functions defined over $\mathscr{A}_A$ such that, for $f_A^1 \sim f^1$ and $f_A^0 \sim f^0$ (which exist by Axiom 9(i)), $u_A(f_A^1) = 1$ and $u_A(f_A^0) = 0$. For any A, $B \in \mathscr{E} - \mathscr{N}$ with $A \cap B = \varnothing$, we know that the additive decomposition

$$u_{A\cup B} = \varphi_{A,B} + \varphi_{B,A}$$

exists. But we also know that $\varphi_{A,B}$ is linearly related to u_A and $\varphi_{B,A}$ to u_B, i.e.,

$$u_{A\cup B} = P(A|A \cup B)u_A + P(B|A \cup B)u_B + \beta_{A,B} + \beta_{B,A}.$$

The proof is completed by showing that the β terms sum to 0, that $P(A|A \cup B)$ can be written as $P(A)/P(A \cup B)$, where the unconditional P is a finitely additive probability measure, and that the functions u_A are order preserving. The last part is by no means obvious and requires an extended argument. Axiom 5 is used to show that standard sequences behave appropriately.

7. PROOF OF THEOREM 1*

LEMMA 1: *If R, $S \in \mathscr{N}$, then $R \cup S \in \mathscr{N}$.*

PROOF: For any $A \in \mathscr{E} - \mathscr{N}$, by Axiom 7(i), $A \cup R$ and $A \cup R \cup S \in \mathscr{E} - \mathscr{N}$. For any $f_{A\cup R\cup S} \in \mathscr{D}$, let $f_{A\cup R}$ and f_A be its restrictions to $A \cup R$ and A, respectively.

By Axiom 7(ii), $f_{A \cup R \cup S} \sim f_{A \cup R} \sim f_A$, and so by Axiom 2, $f_{A \cup R \cup S} \sim f_A$. Thus $R \cup S \in \mathcal{N}$ by Axiom 7(ii). *Q.E.D.*

LEMMA 2: *Let $\succsim$ be the relation on $\mathscr{E} \times (\mathscr{E} - \mathcal{N})$ given in Definition 2.* (i) *If $A|B \succsim C|D$ and $A \cap B \in \mathcal{N}$, then $C \cap D \in \mathcal{N}$*; (ii) $A \cap B|B \sim A|B$; (iii) *$\succsim$ is a weak order.*

PROOF: (i) Since $A \cap B \in \mathcal{N}$ and $A|B \succsim C|D$, condition (iii) of Definition 2 does not apply. Suppose (ii) were to hold, i.e., $\bar{A} \cap B \in \mathcal{N}$; then by Lemma 1, $B = (A \cap B) \cup (\bar{A} \cap B) \in \mathcal{N}$, contrary to choice. Thus condition (i), $C \cap D \in \mathcal{N}$, must hold.

Proof of (ii) is trivial.

(iii) $\succsim$ is obviously connected and reflexive. To show transitivity, suppose that $A|B \succsim C|D$ and $C|D \succsim E|F$. There are seven cases:

(1) $C \cap D \in \mathcal{N}$: By part (i) applied to $C|D \succsim E|F$, we have $E \cap F \in \mathcal{N}$, and so $A|B \succsim E|F$.

(2) $\bar{A} \cap B \in \mathcal{N}$: By definition, $A|B \succsim E|F$.

(3) $A \cap B \in \mathcal{N}$: By part (i) applied to $A|B \succsim C|D$, we have $C \cap D \in \mathcal{N}$, which is case (1).

(4) $\bar{C} \cap D \in \mathcal{N}$: Condition (i) does not hold for $A|B \succsim C|D$, since that would imply $D = (C \cap D) \cup (\bar{C} \cap D) \in \mathcal{N}$. Since condition (iii) cannot hold, (ii) must, i.e., $\bar{A} \cap B \in \mathcal{N}$, which is case (2).

(5) $E \cap F \in \mathcal{N}$: By definition, $A|B \succsim E|F$.

(6) $\bar{E} \cap F \in \mathcal{N}$: As in case (4), conditions (i) and (iii) cannot hold for $C|D \succsim E|F$; hence $\bar{C} \cap D \in \mathcal{N}$, which is case (4).

(7) Condition (iii) applies to both inequalities. By Axiom 9*(i), we may select $f_{A \cap B} \sim f'_{C \cap D} \sim f''_{E \cap F} \succ g_{\bar{A} \cap B} \sim g'_{\bar{C} \cap D} \sim g''_{\bar{E} \cap F}$ and by Axiom 5*, $f_{A \cap B} \cup g_{\bar{A} \cap B} \succsim f'_{C \cap D} \cup g'_{\bar{C} \cap D} \succsim f''_{E \cap F} \cup g''_{\bar{E} \cap F}$, whence $A|B \succsim E|F$ by Axiom 2.

Recall that $A \in \mathscr{E}$ is *proper* if $A, \bar{A} \in \mathscr{E} - \mathcal{N}$.

LEMMA 3: *If A and B are proper events and A is independent of B, then $A \cap B$, $A \cap \bar{B}, \bar{A} \cap B, \bar{A} \cap \bar{B} \in \mathscr{E} - \mathcal{N}$.*

PROOF: Suppose that $A \cap B \in \mathcal{N}$, then Lemma 2(i) applied to $A|B \sim A|\bar{B}$ yields $A \cap \bar{B} \in \mathcal{N}$. By Lemma 1, $A = (A \cap B) \cup (A \cap \bar{B}) \in \mathcal{N}$, contrary to assumption. Similarly, $A \cap \bar{B} \in \mathscr{E} - \mathcal{N}$. Suppose that $\bar{A} \cap B \in \mathcal{N}$; then since $A|\bar{B} \succsim A|B$, either $A \cap B$ or $\bar{A} \cap \bar{B} \in \mathcal{N}$. Lemma 1 yields $B \in \mathcal{N}$ or $\bar{A} \in \mathcal{N}$, both of which are contrary to assumption. The proof that $\bar{A} \cap \bar{B} \in \mathscr{E} - \mathcal{N}$ is similar. *Q.E.D.*

For $A \in \mathscr{E} - \mathcal{N}$, let $\mathscr{D}_A = \{f_A | f_A \in \mathscr{D}\}$. By Axioms 8(ii) and 9*(i), $\mathscr{D}_A \neq \varnothing$.

LEMMA 4: *If $A, B \in \mathscr{E} - \mathcal{N}$, $A \cap B = \varnothing$, and $g \in \mathscr{D}_B$, then the mapping θ_g from $\mathscr{D}_A$ into $\mathscr{D}_{A \cup B}$ defined by $\theta_g(f) = f \cup g$ for $f \in \mathscr{D}_A$ is continuous in the order topology of $\succsim$.*

PROOF: Since a mapping from an ordered set into an ordered set is continuous if and only if the inverse of each open ray is open, it is sufficient to show that for each $h \in \mathscr{D}_{A \cup B}$ there exists $f \in \mathscr{D}_A$ such that

$$\{f' | \theta_g(f') \prec h\} = \{f' | f' \prec f\}.$$

By Axiom 9*(ii), if $h \in \mathscr{D}_{A \cup B}$ there exists $f \in \mathscr{D}_A$ such that $f \cup g \sim h$. If $\theta_g(f') = f' \cup g \prec h \sim f \cup g$, then by Axiom 4, $f' \prec f$. Conversely, if $f' \prec f$ then by Axiom 4, $\theta_g(f') = f' \cup g \prec f \cup g \sim h$. *Q.E.D.*

DEFINITION 4: Suppose that $\mathscr{S}$ is a non-empty set, $\succsim$ is a weak ordering of $\mathscr{S}$, and $\circ$ is a binary operation on $\mathscr{S}$. For $x, y, x', y', \in \mathscr{S}$, (i) $\circ$ is *intern* if for

$$\begin{Bmatrix} x \succ y \\ x \sim y \\ x \prec y \end{Bmatrix}, \quad \text{then} \quad \begin{Bmatrix} x \succ x \circ y \succ y \\ x \sim x \circ y \sim y \\ x \prec x \circ y \prec y \end{Bmatrix};$$

(ii) $\circ$ is *strictly increasing* (in each variable) if for $x \succ x'$, then $x \circ y \succ x' \circ y$, and for $y \succ y'$, then $x \circ y \succ x \circ y'$; (iii) $\circ$ is *continuous* (in each variable) if the mappings $x \to x \circ y$ and $y \to x \circ y$ are continuous in the order topology of $\succsim$; (iv) $\langle \mathscr{S}, \sim, \circ \rangle$ is a *quasi-group* if (1) $x \circ y \sim x' \circ y$ implies $x \sim x'$ and $x \circ y \sim x \circ y'$ implies $y \sim y'$, and (2) for every x, y there exist z and $z' \in \mathscr{S}$ such that $x \circ z \sim y$ and $z' \circ x \sim y$.

To simplify notation, let φ, ψ, etc. denote typical decisions in $\mathscr{D}_X = \{f_X | f_X \in \mathscr{D}\}$. Since Axioms 7(i) and 8(i) imply $X \in \mathscr{E} - \mathscr{N}$, then for each $f_A \in \mathscr{D}$ Axiom 9*(i) assures us that there is a $\varphi \in \mathscr{D}_X$ such that $\varphi \sim f_A$. Moreover, if $\varphi \in \mathscr{D}_X$ and $A \in \mathscr{E} - \mathscr{N}$, there exists $f_A \in \mathscr{D}$ such that $\varphi \sim f_A$. We now define operations on $\mathscr{D}_X$ that reflect the conditional probability structure of the union of disjoint decisions.

DEFINITION 5: For $A \in \mathscr{E}$ and $B \in \mathscr{E} - \mathscr{N}$, define the *binary operation* $\circ_{A|B}$ *on* $\mathscr{D}_X$ as follows. For $\varphi, \psi \in \mathscr{D}_X$, (i) if $A \cap B \in \mathscr{N}$, then $\varphi \circ_{A|B} \psi = \psi$; (ii) if $\bar{A} \cap B \in \mathscr{N}$, then $\varphi \circ_{A|B} \psi = \varphi$; (iii) if $A \cap B$, $\bar{A} \cap B \in \mathscr{E} - \mathscr{N}$, $f_{A \cap B} \sim \varphi$, and $g_{\bar{A} \cap B} \sim \psi$, then $\varphi \circ_{A|B} \psi \sim f_{A \cap B} \cup g_{\bar{A} \cap B}$.

LEMMA 5: *Suppose that* $\langle X, \mathscr{E}, \mathscr{N}, \mathscr{C}, \mathscr{D}, \succsim \rangle$ *satisfies Definition* 1*, *that* $A \in \mathscr{E}$, $B \in \mathscr{E} - \mathscr{N}$, *and that* $\varphi, \varphi', \psi, \psi' \in \mathscr{D}_X$. (i) *If* $A \cap B$, $\bar{A} \cap B \in \mathscr{E} - \mathscr{N}$, *then* $\circ_{A|B}$ *is intern, strictly increasing, and continuous, and* $\langle \mathscr{D}_X, \sim, \circ_{A|B} \rangle$ *is a quasi-group*; (ii) $\varphi \circ_{A|B} \psi \sim \psi \circ_{\bar{A}|B} \varphi$; (iii) *if* $A \cap B$, $A \cap \bar{B}$, $\bar{A} \cap B$, $\bar{A} \cap \bar{B} \in \mathscr{E} - \mathscr{N}$, *then*

$$(\varphi \circ_{A|B} \psi) \circ_{B|X} (\varphi' \circ_{A|\bar{B}} \psi') \sim (\varphi \circ_{B|A} \varphi') \circ_{A|X} (\psi \circ_{B|\bar{A}} \psi');$$

(iv) *if* $A \cap B$, $\bar{A} \cap B$, $\bar{B} \in \mathscr{E} - \mathscr{N}$, *then*

$$(\varphi \circ_{A|B} \psi) \circ_{B|X} \psi \sim \varphi \circ_{A \cap B|X} \psi.$$

PROOF: (i) $\circ_{A|B}$ is intern. Suppose that $\varphi \succ \psi$. Since $A \cap B$, $\bar{A} \cap B \in \mathscr{E} - \mathscr{N}$, by Axiom 9*(i) there exist $f_{A \cap B} \sim g_{\bar{A} \cap B} \sim \varphi \succ \psi \sim h_{\bar{A} \cap B}$. By Axiom 4,

$$f_{A \cap B} \cup g_{\bar{A} \cap B} \succ f_{A \cap B} \cup h_{\bar{A} \cap B} \sim \varphi \circ_{A|B} \psi.$$

By Axiom 3,

$$\varphi \sim f_{A \cap B} \sim f_{A \cap B} \cup g_{\bar{A} \cap B},$$

hence $\varphi \succ \varphi \circ_{A|B} \psi$. The other cases are handled similarly.

$\circ_{A|B}$ is strictly increasing. Suppose that $\varphi \succ \varphi'$. By Axiom 9*(i), there exist $f_{A \cap B} \sim \varphi \succ \varphi' \sim f'_{A \cap B}$ and $g_{\bar{A} \cap B} \sim \psi$, whence, by Axiom 5,

$$\varphi \circ_{A|B} \psi \sim f_{A \cap B} \cup g_{\bar{A} \cap B} \succ f'_{A \cap B} \cup g_{\bar{A} \cap B} \sim \varphi' \circ_{A|B} \psi.$$

The other case is similar.

$\circ_{A|B}$ is continuous. Lemma 4.

$\langle \mathscr{D}_X, \sim, \circ_{A|B} \rangle$ is a quasi-group. Suppose that $\varphi \circ_{A|B} \psi \sim \varphi \circ_{A|B} \psi'$ and let $f_{A \cap B} \sim \varphi, g_{\bar{A} \cap B} \sim \psi, g'_{\bar{A} \cap B} \sim \psi'$. Thus, $f_{A \cap B} \cup g_{\bar{A} \cap B} \sim f_{A \cap B} \cup g'_{\bar{A} \cap B}$, so by Axiom 4, $\psi \sim g_{\bar{A} \cap B} \sim g'_{\bar{A} \cap B} \sim \psi'$. The second cancellation property is similar. Let φ and ψ be given and let $g_{A \cap B} \sim \varphi$ and $h_B \sim \psi$. Since $\bar{A} \cap B \in \mathscr{E} - \mathscr{N}$, Axiom 9*(ii) asserts that there is $f_{\bar{A} \cap B} \in \mathscr{D}$ such that $h_B \sim g_{A \cap B} \cup f_{\bar{A} \cap B}$. Let $\theta \sim f_{\bar{A} \cap B}$; thus $\varphi \circ_{A|B} \theta \sim \psi$. The second case is similar.

Part (ii) is proved by Definition 5.

(iii) By Axiom 9*(i), let

$$f_{A \cap B} \sim \varphi, f'_{A \cap \bar{B}} \sim \varphi', g_{\bar{A} \cap B} \sim \psi, g'_{\bar{A} \cap \bar{B}} \sim \psi',$$

$$h_B \sim \varphi \circ_{A|B} \psi \sim f_{A \cap B} \cup g_{\bar{A} \cap B}, h'_{\bar{B}} \sim \varphi' \circ_{A|\bar{B}} \psi' \sim f'_{A \cap \bar{B}} \cup g'_{\bar{A} \cap \bar{B}},$$

$$k_A \sim \varphi \circ_{B|A} \varphi' \sim f_{A \cap B} \cup f'_{A \cap \bar{B}}, k'_{\bar{A}} \sim \psi \circ_{B|\bar{A}} \psi' \sim g_{\bar{A} \cap B} \cup g'_{\bar{A} \cap \bar{B}}.$$

Then by Definition 5, Axiom 4, the commutativity and associativity of $\cup$,

$$\begin{aligned}(\varphi \circ_{A|B} \psi) \circ_{B|X} (\varphi' \circ_{A|\bar{B}} \psi') &\sim h_B \cup h'_{\bar{B}} \\ &\sim f_{A \cap B} \cup g_{\bar{A} \cap B} \cup f'_{A \cap \bar{B}} \cup g'_{\bar{A} \cap \bar{B}} \\ &\sim f_{A \cap B} \cup f'_{A \cap \bar{B}} \cup g_{\bar{A} \cap B} \cup g'_{\bar{A} \cap \bar{B}} \\ &\sim k_A \cup k'_{\bar{A}} \\ &\sim (\varphi \circ_{B|A} \varphi') \circ_{A|X} (\psi \circ_{B|\bar{A}} \psi').\end{aligned}$$

(iv) $\overline{A \cap B} \in \mathscr{E} - \mathscr{N}$ since if $\bar{A} \cup \bar{B} = \overline{A \cap B} \in \mathscr{N}$, then by Axiom 7(i), $\bar{B} \in \mathscr{N}$, contrary to assumption. Let $f_{A \cap B} \sim \varphi$, $g_{\overline{A \cap B}} \sim g'_{\bar{A} \cap B} \sim g''_{\bar{B}} \sim \psi$, and $h_B \sim f_{A \cap B} \cup g'_{\bar{A} \cap B} \sim \varphi \circ_{A|B} \psi$. Note that $(\bar{A} \cap B) \cup \bar{B} = \bar{A} \cup \bar{B} = \overline{A \cap B}$. By Axiom 4, $g'_{\bar{A} \cap B} \cup g''_{\bar{B}} \sim g'_{\bar{A} \cap B} \sim g_{\overline{A \cap B}}$. So

$$\begin{aligned}\varphi \circ_{A \cap B|X} \psi &\sim f_{A \cap B} \cup g_{\overline{A \cap B}} \\ &\sim f_{A \cap B} \cup g'_{\bar{A} \cap B} \cup g''_{\bar{B}} \\ &\sim h_B \cup g''_{\bar{B}} \\ &\sim (\varphi \circ_{A|B} \psi) \circ_{B|X} \psi.\end{aligned}$$

Q.E.D.

From this point on, the proof is nearly identical to that of Pfanzagl [**9**], although the interpretation is considerably different. Within his system, φ and ψ are consequences, i.e., elements of $\mathscr{C}$, and $\varphi \circ_{A|B} \psi$ is a simple gamble in which the event B is assumed to occur and φ is the outcome if $A \cap B$ occurs and ψ if $\bar{A} \cap B$ occurs. To be a bit more specific, since $\langle \mathscr{D}_X, \succsim, \circ_{A|B} \rangle$ is a weakly ordered quasi-group, $\langle \mathscr{D}_X, \succsim \rangle$ is order complete (Axiom 6*), and $\circ_{A|B}$ is intern, strictly increasing, and continuous, then there exists a real-valued, continuous, and strictly increasing function on $\mathscr{D}_X$. Using Lemma 5 and the Axioms, it is easy to see that this function exhibits properties P1–P9 of Pfanzagl's $m_{A|B}(\varphi, \psi)$, and so his proof, which depends upon the existence of proper events independent of proper events and the validity of Lemma 3, yields our conclusions. *Q.E.D.*

The Institute for Advanced Study
and
University of Michigan

Manuscript received April, 1968; revision received October, 1969.

REFERENCES

[**1**] Allais, M.: "Le Comportement de l'Homme Rationnel Devant le Risque: Critique des Postulats et Axiomes de l'École Americaine," *Econometrica*, 21 (1953), 503–546.
[**2**] Davidson, D., and P. Suppes: "A Finitistic Axiomatization of Subjective Probability and Utility," *Econometrica*, 24 (1956), 264–275.
[**3**] Debreu, G.: "Topological Methods in Cardinal Utility Theory," in K. J. Arrow, S. Karlin, and P. Suppes (Editors), *Mathematical Methods in the Social Sciences*. Stanford: Stanford Univ. Press, 1960, pp. 16–26.
[**4**] Ellsberg, D.: "Risk, Ambiguity, and the Savage Axioms," *Quarterly Journal of Economics*, 75 (1961), 643–669.
[**5**] Fishburn, P. C.: *Decision and Value Theory*. New York: Wiley, 1964.
[**6**] Jeffrey, R. C.: *The Logic of Decision*. New York: McGraw-Hill, 1965.
[**7**] Krantz, D. H., R. D. Luce, P. Suppes, and A. Tversky: *Foundations of Measurement*. Vol. I. New York: Academic Press, 1971 (in press).
[**8**] MacCrimmon, K. R.: "Descriptive and Normative Implications of the Decision Theory Postulates," in K. H. Borch (Editor), *Risk and Uncertainty*. New York: MacMillan, 1967.
[**9**] Pfanzagl, J.: "Subjective Probability Derived from the Morgenstern-von Neumann Utility Concept," in M. Shubik (Editor), *Essays in Mathematical Economics in Honor of Oskar Morgenstern*. Princeton: Princeton Univ. Press, 1967, pp. 237–251. (Or see Chapter 12 of J. Pfanzagl, *Theory of Measurement*. New York: Wiley, 1968.)
[**10**] Ramsey, F. P.: "Truth and Probability," in F. P. Ramsey, *The Foundations of Mathematics and Other Logical Essays*. New York: Harcourt, Brace, 1931, pp. 156–198. (Also in H. E. Kyburg and H. E. Smokler (Editors), *Studies in Subjective Probability*. New York: Wiley, 1964, pp. 61–92.)
[**11**] Savage, L. J.: *The Foundations of Statistics*. New York: John Wiley, 1954.
[**12**] Tversky, A.: "Intransitivity of Preferences," *Psychological Review*, 76 (1969), 31–48.