A CONFLICT BETWEEN FINITE ADDITIVITY AND AVOIDING DUTCH BOOK*

TEDDY SEIDENFELD

Department of Philosophy
Washington University, St. Louis

MARK J. SCHERVISH†

Department of Statistics
Carnegie Mellon University, Pittsburgh

For Savage (1954) as for de Finetti (1974), the existence of subjective (personal) probability is a consequence of the normative theory of preference. (De Finetti achieves the reduction of belief to desire with his generalized Dutch-Book argument for *previsions*.) Both Savage and de Finetti rebel against legislating countable additivity for subjective probability. They require merely that probability be finitely additive. Simultaneously, they insist that their theories of preference are weak, accommodating all but self-defeating desires. In this paper we dispute these claims by showing that the following three cannot simultaneously hold:

(i) Coherent belief is reducible to rational preference, i.e. the generalized Dutch-Book argument fixes standards of coherence.
(ii) Finitely additive probability is coherent.
(iii) Admissible preference structures may be free of *consequences*, i.e. they may lack prizes whose values are robust against all contingencies.

1. Introduction. One of the most important results of the subjectivist theories of Savage and de Finetti is the thesis that, normatively, preference circumscribes belief. Specifically, these authors argue that the theory of subjective probability is reducible to the theory of reasonable preference, i.e. coherent belief is a consequence of rational desire. In Savage's (1954) axiomatic treatment of preference, the existence of a quantitative subjective probability is assured once the postulates governing preference are granted. In de Finetti's (1974) discussion of *prevision*, avoidance of a (uniform) loss for certain is thought to guarantee agreement with the requirements of subjective probability (sometimes called the avoidance of "Dutch Book").

Obviously, the significance of these results depends upon the avowed

*Received December 1982; revised March 1983.

†We would like to thank P. C. Fishburn, Jay Kadane, Isaac Levi, Patrick Maher and a referee for their helpful comments. Also, we have benefited from discussions about consequences with E. F. McClennen.

Philosophy of Science, 50 (1983) pp. 398–412.

liberalism regarding the range of preferences and beliefs the theories are said to tolerate. Both Savage and de Finetti are explicit in their opposition to the stipulation of countable additivity for probability, and, of course, each insists that the constraints imposed on reasonable preference are weak, permitting all but self-defeating desires. It is our purpose in this paper to challenge these claims. We aim to show that the reduction of belief to preference cannot be carried off as Savage and de Finetti suggest without contracting the range of admissible states of preference and belief. In particular, we argue that the purported reduction fails unless

(i) subjective probability is countably additive, or
(ii) each agent is required to acknowledge the existence of a rich supply of *consequences*, i.e. prizes whose values are robust against the contingencies of nature.

As we see in section 3, Savage recognized that consequences serve as expedients in his theory for constructing "constant acts" and should not be essential to subjectivism. We find no convenient stockpile of consequences. In fact, it seems reasonable to deny that there are consequences in practical decisions. Thus, our position is that, lacking consequences, expected utility theory must treat subjective probability distributions as *extraneous* (Fishburn 1970, §12.2 and chapter 13). Otherwise, probabilities which are not countably additive cannot be sanctioned. In light of our findings in section 4, the problem is deeply rooted indeed. The expected utility hypothesis fails for acts with denumerably many outcomes, when probability (extraneous or otherwise) is merely finitely additive and consequences are absent.

Fortunately, Savage's theory is axiomatized so that the first six of his seven postulates deal with the structure of preference for *gambles*, i.e. acts which produce only finitely many different outcomes almost surely. It is the seventh postulate, P7, which carries the extension of expected utility theory to acts in general. Thus, our focus in section 2 is on the final axiom. We examine several conjectures about the conditions under which P7 remains independent of P1–P6 and demonstrate that the independence is *not* a matter of the additivity of the probability. Hence, one may satisfy P1–P6 with a countably additive probability but violate the expected utility hypothesis for acts that are not gambles. In other words, compliance with P1–P6 fails to guarantee the expected utility hypothesis for random variables in general, even on the condition that probability (based on P1–P6) is countably additive. Readers unfamiliar with Savage's theory may wish to skip section 2 on a first reading.

In section 3, we analyze P7 and show that its role in extending utility to acts in general trades on an undesirable feature consequences are conceded to require. For instance, when P7 is reformulated to avoid this feature of consequences, the resulting theory precludes all but countably additive subjective probability. In our discussion of de Finetti's argument against "Dutch Book" in section 4, we grant his working hypothesis that there is a linear utility function for outcomes (as when he assumes dollars linear in utility), then show that his standards for coherence of previsions prohibit merely finitely additive probability. In parallel with Savage's theory, if consequences are introduced and coherence confined to previsions involving consequences exclusively, then, as desired, finite additivity is all that follows from avoidance of "Dutch Book". Thus, the dilemma is between mandating consequences and denying the admissibility of merely finitely additive distributions.

2. On the Independence of P7. In his classic *The Foundations of Statistics* (1954), L. J. Savage constructs a theory of utility, axiomatized in seven postulates. The first six of Savage's axioms yield a theory of expected utility for *gambles*, i.e. acts which produce at most finitely many consequences almost surely. The seventh postulate (P7) extends the theory to acts in general. Immediately following the introduction of P7, Savage demonstrates its independence from the first six with the aid of a finitely, but not countably additive probability. He concludes the demonstration with this terse remark:

> Finite, as opposed to countable, additivity seems to be essential to this example; perhaps, if the theory were worked out in a countably additive spirit from the start, little or no counterpart of P7 would be necessary (1954, p. 78).[1]

Our purpose in this section is to demonstrate that the conjecture implicit in the above remark is not accurate. That is, we will produce examples involving only countably additive probabilities for which P1–P6 are satisfied but P7 is not. This means, on the condition that the expected utility hypothesis is valid for acts in general, some replacement for P7 is necessary even if the theory is worked out in a countably additive spirit.

We will assume that the reader either is familiar with Savage's postulate system or else has a copy of Savage (1954) readily available. Ad-

[1]Savage had misgivings about this comment. In a letter to P. C. Fishburn (dated 30 June 1965) he wrote:

> You suggest that I review the last sentence on page 78 of F. of S. [*Foundations of Statistics*] It is hard for me now to feel sure what I meant by that sentence, and I have serious doubts that it is defensible. But what it seems to say is not that something stronger than P7 would be needed in a countably additive context, but rather something weaker might suffice.

And in another letter to Fishburn (dated 9 September 1966):

> Once you convince yourself, with Zorn's lemma, that the Blackwell-Girshick theorem cannot be had without some counterpart of P7, you will have shown that the conjecture at the bottom of page 78 of F. of S. is more or less incorrect.

We thank Professor Fishburn for bringing these to our attention.

ditionally, we recommend Fishburn (1970, Ch. 14) and Fishburn (1981) for helpful discussions of Savage's theory. In any event, P1–P4 are stated in the proof of Lemma 1 below, a lemma useful for the investigation of conditions under which P7 remains independent of P1–P6. The remaining three postulates, P5–P7, are stated following Lemma 1.

Savage's postulates concern *states* (elements of a set $\boldsymbol{S}$), *events* (subsets of $\boldsymbol{S}$), *consequences* (elements of a set $\boldsymbol{F}$), *acts* (functions from $\boldsymbol{S}$ to $\boldsymbol{F}$), and a relation between acts $\leq$ (read "is not preferred to"). If $\boldsymbol{f} \leq \boldsymbol{g}$ and $\boldsymbol{g} \leq \boldsymbol{f}$, we say $\boldsymbol{g}$ and $\boldsymbol{f}$ are *equivalent*. If $\boldsymbol{f}$ is an act, the consequence of $\boldsymbol{f}$ occurring in the state s is denoted $f(s)$. To avoid additional notation and with only slight encumbrance on the reader, we often identify a consequence with the act which produces that consequence in all states, that is, the constant act. If B is an event, we will use the notation I_B to denote the indicator of B, that is, the function which is 1 if B occurs and 0 if not. Following Savage, we will denote the complement of B, $\sim B$.

Lemma 1: Let S be a measurable set, and let F be a subset of the real numbers containing zero and closed under division (by non-zero elements) and multiplication. Assume all acts are measurable functions from $\boldsymbol{S}$ to $\boldsymbol{F}$, and assume that whenever $\boldsymbol{f}$ and $\boldsymbol{g}$ are acts, and B is an event, $\boldsymbol{f}I_B + \boldsymbol{g}I_{\sim B}$ is an act. Assume all constant functions are acts, and denote the act which is constantly 0 as $\mathbf{0}$. Let W be a mapping from the set of all acts to the finite real numbers which statisfies $W(\mathbf{0}) = 0$,

$$W(\boldsymbol{f} + \boldsymbol{g}) = W(\boldsymbol{f}) + W(\boldsymbol{g}), \tag{2.1}$$

whenever $\boldsymbol{fg} = \mathbf{0}$, and

$$cW(fI_B) = fW(cI_B), \tag{2.2}$$

for all events B and all consequences $f, c \in \boldsymbol{F}$. Define $\boldsymbol{f} \leq \boldsymbol{g}$ if and only if $W(\boldsymbol{f}) \leq W(\boldsymbol{g})$. Then P1–P4 are satisfied.

Proof:

P1: The relation $\leq$ is a simple ordering.
This is trivial and needs no proof.

P2: If $\boldsymbol{f}$, $\boldsymbol{g}$, and $\boldsymbol{f}'$, $\boldsymbol{g}'$ are acts and B is an event such that:

1. for $s \in \sim B$, $f(s) = g(s)$, and $f'(s) = g'(s)$,
2. for $s \in B$, $f(s) = f'(s)$, and $g(s) = g'(s)$,
3. $\boldsymbol{f} \leq \boldsymbol{g}$;

then $\boldsymbol{f}' \leq \boldsymbol{g}'$.

The conditions of P2 say that $\boldsymbol{f}I_{\sim B} = \boldsymbol{g}I_{\sim B}$, $\boldsymbol{f}'I_{\sim B} = \boldsymbol{g}'I_{\sim B}$, $\boldsymbol{f}I_B = \boldsymbol{f}'I_B$, and $\boldsymbol{g}I_B = \boldsymbol{g}'I_B$. Since $\boldsymbol{h} = \boldsymbol{h}I_B + \boldsymbol{h}I_{\sim B}$, for every act $\boldsymbol{h}$, and $I_BI_{\sim B} = \mathbf{0}$, it follows from (2.1) that $\boldsymbol{f}' \leq \boldsymbol{g}'$ if $\boldsymbol{f} \leq \boldsymbol{g}$.

P3: If $f \equiv g, f' \equiv g'$, and B is not null; then $\boldsymbol{f} \leq \boldsymbol{f}'$ given B, if and only if $g \leq g'$ (as constant acts).

Savage defines "$\boldsymbol{f} \leq \boldsymbol{f}'$ given B" to mean $\boldsymbol{g} \leq \boldsymbol{g}'$ for every pair of acts $\boldsymbol{g}$, $\boldsymbol{g}'$ satisfying $\boldsymbol{g}I_B = \boldsymbol{f}I_B$, $\boldsymbol{g}'I_B = \boldsymbol{f}'I_B$, and $\boldsymbol{g}I_{\sim B} = \boldsymbol{g}'I_B$. Under (2.1) and the conditions of P3, this can only happen if $g \leq g'$ (as constant acts).

P4: If f, f', g, g' are consequences, A, B are events, and $\boldsymbol{f}_A, \boldsymbol{f}_B, \boldsymbol{g}_A, \boldsymbol{g}_B$ are acts such that

1. $f' < f$, $g' < g$ (as constant acts),
2. $f_A(s) = f$, $g_A(s) = g$ for $s \in A$,
3. $f_A(s) = f'$, $g_A(s) = g'$ for $s \in \sim A$,
4. $f_B(s) = f$, $g_B(s) = g$ for $s \in B$,
5. $f_B(s) = f'$, $g_B(s) = g'$ for $s \in \sim B$,
6. $\boldsymbol{f}_A \leq \boldsymbol{f}_B$,

then $\boldsymbol{g}_A \leq \boldsymbol{g}_B$.

The conditions of P4 say that $\boldsymbol{f}_A = fI_A + f'I_{\sim A}$, $\boldsymbol{f}_B = fI_B + f'I_{\sim B}$, $\boldsymbol{g}_A = gI_A + g'I_{\sim A}$, and $\boldsymbol{g}_B = gI_B + g'I_{\sim B}$. Condition 6 together with (2.1) and (2.2) yields

$$(f - f')\{W(cI_A) - W(cI_B)\} \leq 0,$$

for some positive $c \in \boldsymbol{F}$. Condition 1 implies that $(f - f')$ and $(g - g')$ have the same sign. Hence

$$(g - g')\{W(cI_A) - W(cI_B)\} \leq 0. \tag{2.3}$$

It follows from (2.2) and (2.3) that $\boldsymbol{g}_A \leq \boldsymbol{g}_B$. □

The final three postulates are:

P5: There is at least one pair of consequences f, f' such that $f' < f$ (as constant acts).

P6: If $\boldsymbol{g} < \boldsymbol{h}$, and f is any consequence; then there exists a (finite) partition of $\boldsymbol{S}$ such that, if $\boldsymbol{g}$ or $\boldsymbol{h}$ is so modified on any one element of the partition as to take the value f at every s there, other values being undisturbed; then the modified $\boldsymbol{g}$ remains strictly not preferred to $\boldsymbol{h}$, or $\boldsymbol{g}$ remains strictly not preferred to the modified $\boldsymbol{h}$, as the case may require.

P7: If $\boldsymbol{f} \leq (\geq)\ g(s)$ given B for every $s \in B$, then $\boldsymbol{f} \leq (\geq)\ \boldsymbol{g}$ given B.[2]

[2]Fishburn (1970, Theorem 14.1) offers a weakened version of P7 which suffices to extend expected utility theory to acts in general. The slight weakening of P7 is achieved by insisting on a *strict* inequality, $f < (>)\ g(s)$, in the antecedent. Our examples 2.2 and 2.3 apply to that form of P7 as well.

The following lemma is stated without proof because it is so straightforward. We then proceed to Savage's example.

Lemma 2: P5 will be satisfied if W assigns different values to at least two different constant acts. Under the conditions of Lemma 1, P6 will be satisfied if for each act $\mathbf{g}$, each consequence f, and each $\epsilon > 0$ there exists a finite partition $B_1, \ldots, B_n$ such that

$$|W(\mathbf{g}I_{B_i}) - W(fI_{B_i})| < \epsilon,$$

for all i.

Example 2.1: (Savage 1954) Let $\mathbf{S}$ be the set of positive integers and $\mathbf{F}$ the interval [0.0, 1.0). Let P be any finitely additive probability on $\mathbf{S}$ which assigns probability 0 to each integer, assigns probability $^1/_2$ to the even integers, and admits a (finite) partition of $\mathbf{S}$ into events of arbitrarily small probability. Any limit point (as $n \to \infty$) of the sequence of discrete uniform distributions over the first n integers will do. Define

$$W(\mathbf{f}) = \int_S f(s)dP(s) + \lim_{\epsilon\to 0} P\{f(s) \geq 1 - \epsilon\}.$$

It is easy to see that W satisfies the conditions of Lemmas 1 and 2. Note that if $\mathbf{f}$ is a gamble, i.e. having finitely many consequences almost surely,

$$\lim_{\epsilon\to 0} P\{f(s) \geq 1 - \epsilon\} = 0.$$

Thus, for gamble $\mathbf{f}$, $W(\mathbf{f})$ is a utility, with $W(\mathbf{f}) = f$ for a constant act $\mathbf{f} \equiv f$. To see that P7 is violated, let $\mathbf{f}$ equal $1 - 1/n$ for even n and 0 for odd n, and let $g(s)$ equal the larger of $^3/_4$ and $f(s)$. We now have $W(\mathbf{f}) = 1$, $W(\mathbf{g}) = {}^{11}/_8$, and $W(g(s)) < 1$, for all $s \in \mathbf{S}$. So $\mathbf{f} \geq g(s)$ given $\mathbf{S}$ for all s, but $\mathbf{f} < \mathbf{g}$.

The following is a similar example which uses countably additive probabilities.

Example 2.2: Let $\mathbf{S} = \mathbf{F}$ be the interval of real numbers [0.0, 1.0], and let P be uniform probability on Lebesgue measurable subsets of $\mathbf{S}$. Let all measurable functions from $\mathbf{S}$ to $\mathbf{F}$ be acts, and define

$$a(\mathbf{f}) = \inf\{P(E) : f(s) \text{ assumes only finitely many values on } {\sim}E\}.$$

Note that $a(\mathbf{f} + \mathbf{g}) = a(\mathbf{f}) + a(\mathbf{g})$ whenever $\mathbf{fg} = \mathbf{0}$. For each act $\mathbf{f}$ define

$$W(\mathbf{f}) = \int_{[0,1]} f(s)dP(s) + a(\mathbf{f}).$$

It is easy to see that W satisfies the conditions of Lemmas 1 and 2. If f is a gamble, $a(f) = 0$, hence $W(f)$ is a utility. To see that P7 is violated, let $g(s) = s$ for all s except $s = 1$, and let $g(1) = 0$. Let $f(s) = 1$, for all s. Then $W(\mathbf{g}) = 1.5$, $W(\mathbf{f}) = 1$, $W(g(s)) = s$ for all s except $s = 1$, and $W(g(1)) = 0$. So $\mathbf{f} > g(s)$ for all s, while $\mathbf{f} < \mathbf{g}$.

The feature that drives Example 2.2 is the fact that the "worth" W of an act is increased from its expected value by the extent to which the act produces uncountably many consequences. Savage proves that (given P1–P6) P7 holds for gambles (effectively Theorem 2.7.3 of Savage 1954). His example (2.1, above) shows that P7 need not hold for acts that assume countably many consequences. The following example shows that this remains the case even when the probability is countably additive (unlike example 2.1).

Example 2.3: Let $\mathbf{S}$ be the half-open interval [0.0, 1.0), and $\mathbf{F}$ the rational numbers in $\mathbf{S}$. Let P be uniform probability on Lebesgue measurable subsets of $\mathbf{S}$. Let all measurable functions $\mathbf{f}$ from $\mathbf{S}$ to $\mathbf{F}$ satisfying

$$a(f) = \lim_{i\to\infty} P\{f(s) \geq 1 - 2^{-i}\}2^{i} < \infty$$

be acts (all subsets of $\mathbf{F}$ being measurable). Define

$$W(f) = \int_{S} f(s)dP(s) + a(f).$$

Once again, W satisfies the conditions of Lemmas 1 and 2. If $\mathbf{f}$ is a gamble, $a(f) = 0$, hence $W(f)$ is a utility. To see that P7 is violated, let $f(s) = 1 - 1/2^{k}$ and let $g(s)$ equal $1 - 1/2^{k+1}$ for $1 - 1/2^{k+1} > s \geq 1 - 1/2^{k}$. Then $W(\mathbf{f}) = 1/3 + 1 = 4/3 > W(g(s))$, for each $s \in \mathbf{S}$. Yet $W(\mathbf{g}) = 2/3 + 2 = 8/3$, so that $\mathbf{g} > \mathbf{f}$.

What examples 2.2 and 2.3 illustrate is the independence of the relationship between P1–P6 and P7 from the degree of additivity which personal probability possesses. What we hope to show in the next section is that this independence follows from the special role that consequences play in P7. If we deny the existence of consequences, attempts to reformulate P7 lead to the exclusion of merely finitely additive probabilities.[3]

3. Dominance and Conglomerability of Probability. Savage's seventh postulate contrasts acts in general through a comparison of one with the

[3]Fishburn (1970) has also studied the relationship between P7 and countable additivity. See the Appendix to this paper for a discussion of the connection between his results and those of the present paper.

consequences of the other (on some non-null event). This is possible because for each consequence, say $g(s)$, there is an act, the constant act $g^* \equiv g(s)$, which serves naturally as the counterpart for the consequence. However, the reader is reminded that constant acts (consequences) have, in virtue of P3, rather distinguished properties. To wit: as stipulated by P3, the relative values of consequences are unaffected given non-null events, i.e. their values are invariant under different states. Thus, a consequence must behave like a prize whose value is robust against whatever (non-null) information we might acquire.[4]

In practical terms, P3 prohibits approximating a consequence by an award of, e.g. stock options where the relative attractiveness of two stocks may be a function of the state of the economy. Are there good candidates for consequences? As Savage argues in his typically evenhanded style,

> . . . what are often thought of as consequences (that is, sure experiences of the deciding person) in isolated decision situations typically are in reality highly uncertain. Indeed, in the final analysis, a consequence is an idealization that can perhaps never be well approximated. I therefore suggest that we must expect acts with actually uncertain consequences to play the role of sure consequences in typical isolated decision situations (1954, p. 84).

If we concede that consequences, in the sense required by P3, are not the entities typically viewed as outcomes in familiar decisions, what can we offer in place of P7 if we ignore the relativization to contrast by consequences? We can recast the question this way. Let $\pi^B = \{h_l, \ldots\}$ be a, possibly infinite, partition of the event B by non-null elements h_i. Can we make sense of a comparison, given B, between act f and each outcome $\boldsymbol{g}I_{h_i}$ even though $\boldsymbol{g}I_{h_i}$ is not a consequence in the fashion of P3? (Unfortunately, we are forced to consider non-null h_i exclusively because in Savage's theory all acts are equivalent given a null event. That is, Savage's program cannot generate probability conditional upon an event of 0 probability.)

Suppose we attempt to avoid consequences entirely. Instead of comparing acts f and g through consequences, we might contrast the outcomes of f and g given h_i $(i = 1, \ldots)$ directly. Thus, we have

P8: If $f \leq (\geq)\ g$ given h_i for all i, then $f \leq (\geq)\ g$ given B.

[4]This feature of consequences is separate from the requirement discussed by Fishburn (1970, p. 166 and 1981, p. 162) that each consequence be "relevant", i.e. an outcome for some act, for each state. In other words, the entire class of Savage-type consequences $\boldsymbol{F}$ is needed to exhaust the range of outcomes for each state. This restriction prohibits the strategy of adopting outcomes under fine descriptions for consequences, since for different states the descriptions are contraries, in violation of the clause that each consequence be "relevant" to each state. We do not know of an axiomatic approach that avoids completely the existence of (at least a pair of) Savage-type consequences "relevant" for each state.

Of course, P8 does not suffice as a replacement for P7 if the goal is to extend expected utility theory to acts in general. To wit: the preference structures $W(\cdot)$ of examples 2.2 and 2.3 satisfy P1–P6 and P8, yet $W(\cdot)$ does not admit a ranking of acts by their expected utility of consequences. Thus, a substitute for P7 must do more (or other) than P8 to reach the goal of the expected utility hypothesis.

Ideally, we would formulate a rule like P8 without the restriction that h_i be non-null, i.e. to permit strict preferences conditional upon an event of zero probability. To repeat, this move is not available to us within Savage's theory because of his definition of null events. Nonetheless, we find it productive to analyze the relation between P8 and P1–P7. Our investigation uncovers a hidden tie to countable additivity, a tie that, perhaps, underpins Savage's (1954, p.78) statement about P7.

P8 trades for its plausibility on a dominance principle, extended to infinite partitions by non-null events. That is, the tacit assumption behind P8 is this: since $\boldsymbol{f}$ is not preferred to $\boldsymbol{g}$ on each element of π^B, $\boldsymbol{f}$ should not be preferred to $\boldsymbol{g}$ given B.

Surprisingly, P8 is inconsistent with P1–P7 unless finitely, but not countably additive probability is precluded. That is, P8 fails, though P1–P7 do not, whenever the agent's subjective probability $P(\cdot)$ is not conglomerable in π^B, or equivalently (Dubins 1975) whenever $P(\cdot)$ is not disintegrable in π^B. Without loss of generality, hereafter we assume $B = \boldsymbol{S}$, the sure event.[5]

Non-conglomerability (de Finetti 1972, §5.30) of a probability $P(\cdot)$ occurs in a partition $\pi = \{h_1, \ldots\}$ if, for some event E and constants k_1 and k_2, $k_1 \leq P(E|h_i) \leq k_2$ for each $h_i \in \pi$, yet $P(E) < k_1$ or $P(E) > k_2$. For example (due to Dubins, see de Finetti 1972, p. 205), let the sure-event be the union of the events (i,j) where $i = 1, 2, \ldots$ is an integer and $j = 0$ or 1. Let I stand for the first coordinate, and let J stand for the second coordinate. Let $\pi = \{h_i | h_i = \{(i,0),(i,1)\}\}$ be a partition. Consider a finitely, but not countably additive probability $P(\cdot)$ such that $P(J = 0) = P(J = 1) = 1/2$, and $P(I = i|J = j) = 2^{-(i+j)}$. There are many such $P(\cdot)$. Each has probability "adherent" along the sequence of $(i,1)$ events. That is, $P(J = 1) = P(\cup_i\{(i,1)\}) = 1/2$, but $\Sigma_i P(i,1) = 1/4$. By finite additivity, $P(h_i) = 3\times 2^{-(i+2)}$, so each element of π is non-null. Bayes' Theorem entails that $P(J = j|h_i) = (2 - j)/3$. Hence, for each element of π, $P(J = 1|h_i) = 1/3$, however $P(J = 1) = 1/2$. $P(\cdot)$ is not conglomerable in π.

Failure of P8 follows directly from non-conglomerablity of $P(\cdot)$ in π.

[5]The convenience, $B = S$, is justified by the result that conglomerability of $P(\cdot)$ fails in a partition π of the sure-event just in case it fails for $P(\cdot|B)$, given some non-null B, in the restricted partition π^B (Kadane, Schervish, and Seidenfeld 1981, §4).

To show that $P(\cdot)$ does not satisfy P8, let the dollar symbol \$ denote utiles, and assume that prizes worth any desired number of utiles exist. (These prizes need not be consequences in Savage's sense.) Consider the acts

f: an even odds bet on E, the event that $J = 1$, with stake of \$2,

that is, the agent places \$1 on E against an opponent who places \$1 on $\sim E$ and

g: the bet on $\sim E$ $\{J = 0\}$ at odds of 3:2 with a stake of \$2,

that is, the agent places \$1.2 on $\sim E$ against an opponent who places \$0.80 on E. Since $P(E) = 1/2$, $g < f$. However, since $P(E|h_i) = 1/3$, for each h_i, $i = 1, \ldots, f < g$ given h_i for all i. This contradicts P8.

Hence, P1–P7 do not entail a dominance rule P8 unless conglomerability in denumerable partitions is mandated, in which case countable additivity holds (see Schervish, Seidenfeld, and Kadane 1981). Moreover, P1–P6 and P8 are insufficient to extend utility theory from gambles to acts, even when probability is countably additive (example 2.2) and the space of consequences is denumerable (example 2.3). In short, the relation between Savage's "extended sure-thing" principle and dominance does not parallel the relation between the basic "sure-thing" principle (P2) and the dominance-in-finite-partition rule which follows from it (Savage 1954, Theorm 2.7.2).[6]

Our failure to duplicate the force of P7 with a principle of dominance raises the following problem, which we discuss at length in the next section. Savage's theory, like de Finetti's which antedates it, extracts probability from preference. There is no extraneous probability in either program. However, unlike Savage's approach, de Finetti's theory begins with coherence of previsions, and establishes the calculus of probability from that. But coherence of previsions is a requirement of dominance; specifically, de Finetti's concern is that one's previsions be undominated. De Finetti's definition of coherence requires that dominance hold in finite partitions regardless of how we view the status of outcomes, i.e. for coherent previsions dominance in finite partitions does not demand a contrast by consequences. Our question, then, is this. How does de Finetti's approach apply for previsions not limited to finitely many outcomes, without legislating countable additivity while retaining a principle of coherence?

[6]Theorem 2.7.2 of Savage (1954) reads as follows:

If $\pi^B = \{B_1, \ldots, B_n\}$ is a finite partition of B, and $f \leq g$ given B_i for each i, then $f \leq g$ given B.

The proof of this theorem depends only on P1 and P2.

4. On the Avoidance of "Dutch Book". De Finetti (1974, chapter 3) presents his concept of a prevision and offers an "operational definition" in terms of lotteries.[7] An agent's *prevision* is a function P from real valued random quantities X to real numbers $x^+ = P(X)$, where x^+ is the agent's "fair price equivalent" for the lottery that yields a prize worth x-units when $X = x$. If it can be shown that a person's prevision ought to satisfy the constraints:

(a) $P(X + Y) = P(X) + P(Y)$ (additive previsions) and
(b) $\inf X \leq P(X) \leq \sup S$ (previsions lie within the range of values for X),

then (normatively) previsions are finitely additive expectations, and entail finitely additive probability for events as a special case. Specifically, if events E are identified with their indicator functions I_E, then previsions for events are subjective probabilities.

De Finetti defines a set of previsions to be *coherent* if no finite selection of "fair" lotteries yields a uniformly negative return for each possible outcome of the random variables involved in the lotteries. (See, e.g., Shimony 1955 for a discussion of coherence and "Dutch Book" regarding events.) Specifically, for each random variable X, the agent is required to hold a prevision $P(X) = x^+$ that he feels makes "fair" a lottery yielding a prize worth $c(X - x^+)$ units, where c is some real-valued constant selected by an "opponent". Again, the agent's previsions are coherent if among such "fair" lotteries there is no finite selection of non-zero c's that guarantee a uniform loss regardless of which values the random variables assume.[8]

De Finetti operationalizes all this by supposing that, for modest quantities of money, people are prepared to use dollar(lire)-unit prizes. That is, with the c's confined to some small interval about 0, de Finetti posits the existence of previsions when lotteries are given in small dollar amounts. The effect of this working hypothesis is to make utility linear in dollars

[7] De Finetti offers two criteria of coherence which, of course, he shows to be equivalent. We make use of his first criterion in the text, but our objection applies also to the second criterion. The second criterion is that the agent is attempting to minimize the value of a proper scoring rule (c.f. Savage 1971).

[8] The two clauses: (i) that coherence involves *finitely* many lotteries, and (ii) that merely *uniformly* negative losses be avoided, are each necessary to avoid restricting coherence to countably additive previsions. To see that (i) is necessary, consider a finitely additive probability which assigns 0 probability to each of a denumerable exhaustive collection of pair-wise disjoint events. By accepting a \$1 wager at odds of 1:0 on each of the denumerably many indicator variables (corresponding to these events), the agent permits an "opponent" a \$1 win for sure. To see that (ii) is necessary, consider the random variable $X = x_i = 1/i$ for $i = 1, 2, \ldots$ with a finitely additive probability that assigns $P(x_i) = 0$ for all i. Then the prevision for X, is $P(X) = 0$. But with $x^+ = 0$, $-(X - x^+) < 0$, but not uniformly less than 0.

in the region near \$0. For example, a prevision of 0.5 for the event E means that for small dollar stakes the agent is prepared to offer even odds, regardless of whether (by choosing $c < 0$) he bets on E, or (by choosing $c > 0$) he bets against E.

We separate de Finetti's thesis, that there exist previsions when prizes are measured in the appropriate scale, i.e. utiles, from his working hypothesis, that small dollar amounts are utiles. Is the thesis neutral regarding disputes in inductive inference? Kyburg (1978) argues in the negative. For our purposes, fortunately, we do not need to enter this debate. We are prepared to grant both the thesis and working hypothesis, since our goal is to show that the reduction of belief to preference does *not* follow from the standard of coherence alone and our criticism is compatible with both assumptions.

However, to grant the working hypothesis is not to concede that dollar prizes (or whatever) are consequences (in Savage's sense). For constructing a lottery for a prevision of X, we need only have available prizes that can be awarded in cx units when $X = x$, subject to the thesis that the prevision is independent of the sign and magnitude of c. The unreasonableness of viewing prizes as Savage-type consequences is as apparent to de Finetti as to Savage. When discussing the dispensibility of his working hypothesis (to cover the familiar problem of "risk aversion"), he says

> It would be more appropriate, instead of considering the variable x representing the gain, to take f + x, where f is the individual's 'fortune' (in order to avoid splitting hairs, inappropriate in this context, one could think of the value of his estate). Anyway, it would be convenient to choose a less arbitrary origin to take into account the possibility that judgments may alter because in the meantime variations have occurred in one's fortune, or risks have been taken, and in order not to preclude for oneself the possibility of taking these things into account, should the need arise. Indeed, as a recognition of the fact that the situation will always involve risks, it would be more appropriate to denote the fortune itself by F (considering it as a random quantity), instead of with f (a definite value) (de Finetti 1974, p.79).

What de Finetti says of the agent's 'fortune' applies *mutatis mutandis* to the payoff of the lottery. That is, given $X = x$, one augments the agent's capital reserve by an amount, cx units. But the prize itself is no different in kind from the fortune (F), each of which the agent owns outright and neither of which (normatively) need be a full-blooded consequence.

In light of our discussion in section 3, it should come as no surprise to learn that unless lottery prizes are consequences, de Finetti's criterion of coherence precludes all but countably additive distributions. As before, failure of conglomerability entails a (uniform) failure of dominance. Coherence, after all, requires avoidance of previsions that induce a failure of dominance with respect to the alternative: no bet. But conglomerability in denumerable partitions is equivalent to countable additivity. Thus, for previsions of random variables assuming more than finitely many outcomes, coherence entails countable additivity. The following (typical) construction illustrates the incoherence of merely finitely additive distributions.

For simplicity, let $P(\cdot)$ be a finitely additive probability assuming infinitely many different values. It follows from Theorem 3.1 of Schervish, Seidenfeld, and Kadane (1981) that there exist an event E, a positive number d, and a partition $\pi = \{h_1 \ldots\}$ such both that $P(h_i) > 0$ and $P(E) - P(E|h_i) > d$ for all i. That is, conglomerability fails in π with regard to the event E. Let $x_i = P(E|h_i)$ and $k = \sup_i x_i$ so that $P(E) \geq k + d$. Consider, next, a wager W that yields a \$1 (one utile) prize in case E occurs and \$0 otherwise. W is worth at least $k + d$. However, given h_i, W is worth at most k. So, the prevision of W, $P(W) \geq k + d$, and $P(W|h_i) < k$ for all i. Define the random variable X so that $X = x_i$ just in case h_i obtains. That is

$$X = \sum_i x_i I_{h_i}.$$

Then X is the (conditional expected) utility of any lottery whose prizes have (conditional expected) utility given h_i equal to x_i for all i. W is such a lottery. Define the lottery Y by saying that Y awards prize WI_{h_i} if h_i obtains for $i = 1, 2, \ldots$. It is clear that $Y = W$; however, note that the prizes WI_{h_i} awarded by Y are not consequences. Since $0 \leq x_i \leq k$ for all i, the prevision for any lottery whose prizes are worth x_i under h_i should be between 0 and k. Y is such a lottery, hence y^+ should be between 0 and k. But, $W = Y$ so $-(W - y^+)$ should be considered fair. However, $w^+ \geq k + d$, so $-(W - y^+)$ is worth no more than $-d$, and hence is unfair. On the other hand, if y^+ is chosen equal to w^+, then y^+ is not the expected utility over the partition π.

In summary, we see that de Finetti's criterion of coherence for prevision rules out merely finitely additive distributions unless lotteries are restricted to prizes which are consequences in the sense of Savage. In the construction of the previous paragraph, WI_{h_i} serves as a prize whose value, given h_i, equals x_i; however WI_{h_i} is not a consequence, since its value is not independent of, e.g., the event E. The dilemma is, of course, that consequences are hard to come by and, it would seem, beyond what is required by consideration of rational preference. Can we not argue, like Savage and de Finetti, that there always are risks? The lesson is clear. We will not have all three of the following:

(i) reduction of coherent belief to rational preference;
(ii) coherence of finitely additive probability;
(iii) admissible preference structures free of consequences.

In light of the highly questionable character of consequences, it seems best to us to dispose of either (i) or (ii). Since many classical statistical procedures require finitely additive "priors" in order to be Bayesian (see e.g. Heath and Sudderth 1978), there is good reason to resist abandoning (ii). However, given (ii) and (iii), the expected utility hypothesis fails (as shown above). Without consequences to fall back upon, non-conglomerability of finitely additive distributions cannot be squared with a requirement of (uniform) dominance for acts. This leaves statistical decision theory devoid of a formidable criterion: *admissibility* (see Savage 1954, p.114). We do not find this an easy choice to make. Perhaps further discussion will suggest a solution.

APPENDIX

In this appendix we examine several results and suggestions of Fishburn (1970) which are related to our discussion in section 2. First, Fishburn (1970, Ch. 10) considers the relationship between countable additivity and a set of postulates not equivalent to Savage's P1–P6. For example, Fishburn's (1970, p.137) postulate *S*4, that the set of probability measures be closed under countable convex combination, is not implied by Savage's system. In fact, this postulate rules out example 2.3 but not example 2.2.

Second, Savage (1954 [second edition 1972], p. 78*n*.) references a suggestion by Fishburn (1970, p.213, ex. 21) for weakening P7 to accommodate acts in general, subject to the constraint that probabilities are countably additive. Fishburn's suggested version, called P7b, requires that

if $f \leq (\geq)\ g(s)$ given A, for each $s \in A$, then $f \leq (\geq)\ \boldsymbol{g}$ given A, for constant act $\boldsymbol{f} = f$

P7b is less demanding than P7 in that it contrasts acts in general with constant acts solely, and not with other acts in general. However, P7b is *not* sufficient for extending utility theory to the class of acts in general even when probability is countably additive. In example 2.3, $W(\cdot)$ satisfies P7b though acts are not ranked by W in accord with their expected utilities (as fixed by P1–P6). To see that W of example 2.3 satisfies P7b it is sufficient to verify that:

(case 1) if $f \leq g(s)$ given A, for each $s \in A$, then $fI_A \leq \int g(s)I_A dP(s)$. Since $a(\cdot)$ is non-negative, it follows that $W(fI_A) \leq W(gI_A)$; hence $\boldsymbol{f} \leq \boldsymbol{g}$ given A as required.

(case 2) if $f \geq g(s)$ given A, for all $s \in A$, then as $f < 1$, $a(gI_A) = 0$. Thus, $W(fI_A) \geq \int g(s)I_A dP(s) = W(gI_A)$, and $\boldsymbol{f} \geq \boldsymbol{g}$ given A as required.

It is to be observed that there can be no counterexample to the conjecture that P7b suffices for extending utility theory to acts in general if $\boldsymbol{F}$ is closed under limits of utility (as fixed by P1–P6). Since P1–P6 entail that consequences have finite utility (Savage 1954, p.81), the assumption of "closure" yields bounded utilities for gambles. However, subject to "closure" of $\boldsymbol{F}$ under preference, P7b entails P7, given P1–P6, regardless of the additivity of probability. This is seen as follows:

Assume the antecedent of P7, that is $\boldsymbol{f} \leq (\geq)\ g(s)$ given A, for each $s \in A$. By hypothesis of "closure", there exist constant acts $\boldsymbol{g}_*$ and $\boldsymbol{g}^*$ which are the infemum and supremum of the consequences of g for $s \in A$. Clearly $\boldsymbol{f} \leq \boldsymbol{g}_*$ $(\geq \boldsymbol{g}^*)$ given A. Then by P7b, $\boldsymbol{g}_* \leq (\boldsymbol{g}^* \geq)\ \boldsymbol{g}$. By transitivity, $\boldsymbol{f} \leq (\geq)\ \boldsymbol{g}$, as required by P7.

REFERENCES

de Finetti, B. (1972), *Probability, Induction and Statistics*. New York: Wiley.
de Finetti, B. (1974), *Theory of Probability* (2 vols.). New York: Wiley.
Dubins, L. (1975), "Finitely additive conditional probabilities, conglomerability and disintegrations", *Annals of Probability 3*: 89–99.
Fishburn, P. C. (1970), *Utility Theory for Decision Making*. New York: Wiley.
Fishburn, P. C. (1981), "Subjective Expected Utility: A Review of Normative Theories", *Theory and Decision 13*: 139–199.
Heath, D. and Sudderth, W. (1978), "On finitely additive priors, coherence, and extended admissibility", *Annals of Statistics 6*: 333–345
Kadane, J. B., Schervish, M. J., and Seidenfeld, T. (1981), *Statistical Implications of Finitely Additive Probability*. Technical Report 206. Department of Statistics, Pittsburgh: Carnegie-Mellon University, July.
Kyburg, H. E. (1978), "Subjective probability: considerations, reflections, and problems", *Journal of Philosophical Logic 7*: 157–180.
Savage, L. J. (1954), *Foundations of Statistics*. New York: Wiley.
Savage, L. J. (1971), "Elicitation of Personal Probabilities and Expectations", *Journal of the American Statistical Association 66*: 783–801.
Schervish, M. J., Seidenfeld, T., and Kadane, J. B. (1981), *The Extent of Non-conglomerability of Finitely Additive Probabilities*. Technical Report 198, Department of Statistics, Pittsburgh: Carnegie-Mellon University, March.
Shimony, A. (1955), "Coherence and the axioms of confirmation", *Journal of Symbolic Logic 20*: 1–28.

LINKED CITATIONS

A Conflict between Finite Additivity and Avoiding Dutch Book
Teddy Seidenfeld; Mark J. Schervish
Philosophy of Science, Vol. 50, No. 3. (Sep., 1983), pp. 398-412.
Stable URL:
http://links.jstor.org/sici?sici=0031-8248%28198309%2950%3A3%3C398%3AACBFAA%3E2.0.CO%3B2-3

[Footnotes]

[7] **Elicitation of Personal Probabilities and Expectations**
Leonard J. Savage
Journal of the American Statistical Association, Vol. 66, No. 336. (Dec., 1971), pp. 783-801.
Stable URL:
http://links.jstor.org/sici?sici=0162-1459%28197112%2966%3A336%3C783%3AEOPPAE%3E2.0.CO%3B2-V

References

Finitely Additive Conditional Probabilities, Conglomerability and Disintegrations
Lester E. Dubins
The Annals of Probability, Vol. 3, No. 1. (Feb., 1975), pp. 89-99.
Stable URL:
http://links.jstor.org/sici?sici=0091-1798%28197502%293%3A1%3C89%3AFACPCA%3E2.0.CO%3B2-Q

On Finitely Additive Priors, Coherence, and Extended Admissibility
David Heath; William Sudderth
The Annals of Statistics, Vol. 6, No. 2. (Mar., 1978), pp. 333-345.
Stable URL:
http://links.jstor.org/sici?sici=0090-5364%28197803%296%3A2%3C333%3AOFAPCA%3E2.0.CO%3B2-5

LINKED CITATIONS

Elicitation of Personal Probabilities and Expectations
Leonard J. Savage
Journal of the American Statistical Association, Vol. 66, No. 336. (Dec., 1971), pp. 783-801.
Stable URL:
http://links.jstor.org/sici?sici=0162-1459%28197112%2966%3A336%3C783%3AEOPPAE%3E2.0.CO%3B2-V

Coherence and the Axioms of Confirmation
Abner Shimony
The Journal of Symbolic Logic, Vol. 20, No. 1. (Mar., 1955), pp. 1-28.
Stable URL:
http://links.jstor.org/sici?sici=0022-4812%28195503%2920%3A1%3C1%3ACATAOC%3E2.0.CO%3B2-T

NOTE: *The reference numbering from the original has been maintained in this citation list.*