ECONOMETRICA
JOURNAL OF THE ECONOMETRIC SOCIETY

A More Robust Definition of Subjective Probability

Author(s): Mark J. Machina and David Schmeidler

Source: *Econometrica*, Jul., 1992, Vol. 60, No. 4 (Jul., 1992), pp. 745-780

Published by: The Econometric Society

Stable URL: https://www.jstor.org/stable/2951565

JSTOR is a not-for-profit service that helps scholars, researchers, and students discover, use, and build upon a wide range of content in a trusted digital archive. We use information technology and tools to increase productivity and facilitate new forms of scholarship. For more information about JSTOR, please contact support@jstor.org.

Your use of the JSTOR archive indicates your acceptance of the Terms & Conditions of Use, available at https://about.jstor.org/terms

The Econometric Society is collaborating with JSTOR to digitize, preserve and extend access to *Econometrica*

JSTOR

This content downloaded from
154.59.124.38 on Thu, 28 Aug 2025 13:46:36 UTC
All use subject to https://about.jstor.org/terms

Econometrica, Vol. 60, No. 4 (July, 1992), 745–780

A MORE ROBUST DEFINITION OF SUBJECTIVE PROBABILITY

BY MARK J. MACHINA AND DAVID SCHMEIDLER[1]

The goal of choice-theoretic derivations of subjective probability is to separate a decision maker's underlying *beliefs* (subjective probabilities of events) from their *preferences* (attitudes toward risk). Classical derivations have all relied upon some form of the Marschak-Samuelson "Independence Axiom" or the Savage "Sure-Thing Principle," which imply that preferences over lotteries conform to the expected utility hypothesis. This paper presents a choice-theoretic derivation of subjective probability, in a Savage-type setting of purely subjective uncertainty, which neither assumes nor implies that the decision maker's preferences over lotteries necessarily conform to the expected utility hypothesis.

KEYWORDS: Subjective probability, uncertainty, subjective uncertainty, non-expected utility theory, expected utility theory, Leonard J. Savage.

1. INTRODUCTION

THE MODERN OR "CHOICE-THEORETIC" theory of subjective probability, as developed in the seminal works of Ramsey (1931), Anscombe and Aumann (1963), and Savage (1954), can be viewed as the culmination of two separate lines of inquiry. The first of these came out of the mathematics and mathematical statistics literature, and addressed the question:

> "When can an individual's beliefs over the relative likelihoods of events be said to be consistent with classical[2] probability theory?"

In other words, given a binary relation $\succeq_l$ over events (including their unions, intersections, and complements) where $A \succeq_l B$ denotes that A is believed to be at least as likely as B, when can $\succeq_l$ be represented by a classical probability measure $\mu(\cdot)$ in the sense that $A \succeq_l B$ if and only if $\mu(A) \geqslant \mu(B)$? This concept, known as *intuitive* or *qualitative probability*, was explored by de Finetti (1937, 1949), Koopman (1940a, 1940b, 1941), Kraft, Pratt, and Seidenberg (1959), Chateauneuf (1985), and others,[3] who obtained necessary and sufficient conditions for such a relation to be representable by a classical probability measure.

[1] We are grateful to Vince Crawford, Valentino Dardanoni, Eddie Dekel, Mark Durst, Simon Grant, Birgit Grodal, Faruk Gul, Edi Karni, David Kreps, Irving LaValle, Ehud Lehrer, Duncan Luce, Salvatore Modica, Michael Rothschild, Joel Sobel, Max Stinchcombe, Menahem Yaari, the Editor, anonymous referees, and especially Peter Fishburn and Peter Wakker for helpful discussions on this material. Support from the Deutsche Forschungsgemeinschaft, Gottfried-Wilhelm-Leibniz Förderpreis during the 1989 Bonn Workshop in Mathematical Economics is gratefully acknowledged. Responsibility for errors and opinions is our own.

[2] By "classical" probability we mean a measure that satisfies the axioms of finitely additive probability theory. We consider the countably additive case in Section 6.2.

[3] E.g., Villegas (1964, 1967), Fishburn (1969b, 1983a, 1983b, 1986), Luce (1967, 1968), Wakker (1981), and Chateauneuf and Jaffray (1984).

This content downloaded from
154.59.124.38 on Thu, 28 Aug 2025 13:46:36 UTC
All use subject to https://about.jstor.org/terms

While this approach gave a characterization of what may be termed "probabilistically sophisticated" *beliefs*, it could not fully serve the goals of economists, who were interested in the issue of *choice* under uncertainty, or in other words, in the question:

> "When can an individual's *choices over uncertain economic prospects* be said to be consistent with classical probability theory?"

This was the question addressed by Ramsey, Anscombe-Aumann, and Savage.

The second line of inquiry that led to this question, and to its resolution by the above researchers, came out of the literature on the foundations of the expected utility model of decision making. Most researchers in this area, from Bernoulli (1738) through von Neumann and Morgenstern (1944, 1947, 1953) and Herstein and Milnor (1953), represented uncertainty by means of explicit probabilities, so that the objects of choice consisted of well-defined probability distributions over outcomes. (This form of representation is known as *objective uncertainty*.) These researchers obtained necessary and sufficient conditions for preferences over probability distributions to be representable by the expectation of a von Neumann-Morgenstern utility function over outcomes.

The representation of uncertainty by formal probabilities has allowed for the application of a tremendous number of results from probability theory, and it is hard to imagine where the theory of games, the theory of search, or the theory of auctions would be without it.[4] However, real-world uncertainty seldom presents itself in terms of exogenously specified probabilities, but rather, as alternative "events" or "states of nature," so that instead of well-defined objective probability distributions, the objects of choice are typically "bets" or "acts" which assign outcomes to the alternative possible events or states. (This form of representation is known as *subjective uncertainty*.) Given the unrealism of using probabilities as a *primitive* notion, but given the value of being able to *derive* them from preferences or choice, it is no surprise that this line of research also led back to the previously displayed question, which can be phrased more formally as:

> "When can choices over subjectively uncertain acts be said to be consistent with probabilistically sophisticated beliefs over event likelihoods?"

In other words, whereas the theory of qualitative probability had developed a method of deriving probabilities from *beliefs*, it had not yet derived them from *choice*, and whereas the original axiomatic work on expected utility had derived *utilities* from choice, it had not yet derived *probabilities* in a similar manner.

The choice-theoretic approach to subjective probability, as pioneered by Ramsey, Anscombe-Aumann, and Savage, and further explored by Arrow (1965, Lect. 1; 1970, Ch. 3), Fishburn (1970), and others, answered this question by obtaining conditions on preferences over subjectively (or mixed subjectively/ob-

[4] Although researchers such as Dempster (1968), Shafer (1976), and Zadeh (1978) have proposed alternative approaches for the numerical representation of uncertain beliefs, classical probability theory remains the predominant normative and analytical framework for the quantitative representation of uncertainty (e.g., Lindley (1982)).

This content downloaded from
154.59.124.38 on Thu, 28 Aug 2025 13:46:36 UTC
All use subject to https://about.jstor.org/terms

jectively) uncertain acts which imply that the individual: (i) possesses a unique, well-defined classical probability distribution over events, (ii) possesses a von Neumann–Morgenstern utility function over outcomes, and (iii) ranks subjectively uncertain acts according to the expected utility of their induced probability distributions over outcomes.

The earliest of these approaches, that of Ramsey (1931), worked in terms of subjective uncertainty, including an event with a known subjective probability of 1/2, from which the utility function over outcomes and hence the subjective probabilities of all other events could be derived. Anscombe and Aumann (1963) used "horse race-roulette lotteries" involving both subjective and objective uncertainty, from which both a utility function and subjective event probabilities were derived.[5] In what has been termed the "crowning glory of choice theory,"[6] Savage used purely subjective acts over an infinitely divisible state space, deriving a utility function over outcomes and a classical subjective probability measure over states. Although their settings differ, each provides a characterization of what can be termed a *probabilistically sophisticated*[7] *expected utility maximizer*. Since the derived probability of each event as well as the derived utility of each outcome are independent of the particular assignment of outcomes to events, these representations are often referred to as yielding a *separation of preferences from beliefs*.

The purpose of this paper is to investigate the robustness of these characterizations of probabilistically sophisticated choice behavior. Specifically, we want to determine how much of the choice-theoretic development of subjective probability depends upon the hypothesis of expected utility maximization; that is,

> "Do departures from the expected utility property of *preferences* affect our ability to characterize the concept of probabilistically sophisticated *beliefs?*"

Or put in another manner,

> "What does it take for choice behavior that does not necessarily conform to the expected utility hypothesis to nonetheless be based on probabilistic beliefs?"

We will call such an agent a *probabilistically sophisticated* ***non***-*expected utility maximizer*.

The rationale for this undertaking comes from three sources. The first is the experimental work of Allais (1953, 1979), Kahneman and Tversky (1979), and others, who have uncovered systematic violations of the expected utility hypothesis in experimentally observed preferences over gambles involving explicit numerical probabilities. If agents do not maximize expected utility in such well-defined settings of objective uncertainty, it is hard to believe that they will

[5] Similar derivations have been developed by Davidson and Suppes (1956), Pratt, Raiffa, and Schlaifer (1964), Ferreira (1972), and Fishburn (1967, 1969a).

[6] Kreps (1988, p. 120).

[7] Fishburn (1987, p. 828; 1988, p. 27) has used the term "reduction principle" to denote the property we are terming "probabilistic sophistication."

This content downloaded from
154.59.124.38 on Thu, 28 Aug 2025 13:46:36 UTC
All use subject to https://about.jstor.org/terms

do so in real world settings of subjective uncertainty. Not surprisingly, these findings have led to the development of numerous non-expected utility models of decision making.[8] However, with the exceptions of Yaari (1969), Schmeidler (1982, 1989), Gilboa (1987), Hazen (1987), Fishburn (1989), Wakker (1989a, Ch. VI; 1989b), and Sugden (1992), these new theories apply to preferences over objective probability distributions, so that they too are subject to the criticism that real world uncertainty is subjective, and stand in need of an operational foundation in terms of preferences over subjectively uncertain acts.

The second justification for our work stems from the fact that probabilistically sophisticated non-expected utility preferences arise naturally in situations of *delayed resolution* of uncertainty. Extending the well-known argument of Markowitz (1959, Ch. 11), Mossin (1969), and Spence and Zeckhauser (1972) to the case of subjective uncertainty, we show that while an expected utility maximizer's preferences over delayed-resolution prospects typically *violate* the expected utility hypothesis, they *retain* the property of probabilistic sophistication. Thus, situations involving delayed-resolution subjective uncertainty induce probabilistically sophisticated non-expected utility behavior.

A final reason for "liberating" the theory of subjective probability from the expected utility hypothesis stems from one of the original goals of the Ramsey/Anscombe-Aumann/Savage approach, namely the separation of an individual's *preferences* from their *beliefs*. If such a separation is analytically or normatively desirable (and we agree that it is), it makes sense to separate the characterization of probabilistic beliefs from as many *restrictions* on risk preferences as possible. In other words, just as a characterization of subjective probability that applied to all expected utility maximizers would be more desirable than one that applied just to risk-neutral agents, a characterization that applied to both expected utility *and* non-expected utility maximizers would be more desirable still.

The following section describes the expected utility-based characterization of Savage (1954), which provides the starting point for our own contribution. Section 3 gives a formal description of probabilistically sophisticated non-expected utility preferences, and shows how such preferences naturally arise in situations of delayed resolution of uncertainty. It also shows that neither of the two most obvious approaches to our topic—(i) postulating all of the Savage axioms except his expected utility-based "Sure-Thing Principle," or (ii) applying the qualitative probability approach to preferences over "bets"—are sufficient to even imply probabilistic sophistication, much less characterize it. Section 4 shows how dropping the Sure-Thing Principle and strengthening one of the remaining Savage axioms *does* provide a characterization of such agents—that is to say, a choice-theoretic axiomatization of classical subjective probability which neither assumes nor implies the expected utility hypothesis. Section 5

[8] See Machina (1983, 1987), Sugden (1986), Weber and Camerer (1987), Fishburn (1988), and Karni and Schmeidler (1991) for surveys of the experimental evidence as well as these alternative models.

This content downloaded from
154.59.124.38 on Thu, 28 Aug 2025 13:46:36 UTC
All use subject to https://about.jstor.org/terms

extends the analysis to the case of conditional preferences and beliefs, and Section 6 gives a discussion of related work and an extension to countably additive subjective probability. Proofs appear in an Appendix.

2. SAVAGE'S CHARACTERIZATION OF PROBABILISTICALLY SOPHISTICATED EXPECTED UTILITY MAXIMIZERS

2.1. *Savage's Framework, Axioms, and Theorem*

Since the Savage approach does not assume the existence of any extraneous randomization device, the objects of choice consist of "acts" $f(\cdot), g(\cdot)$, etc. which assign an outcome to each state of nature. The price of being able to work with purely subjective uncertainty is that the state space must be infinitely divisible. Formally, Savage's setting consists of:

$\mathscr{S} = \{\ldots, s, \ldots\}$ a set of *states*;

$\mathscr{E} = 2^{\mathscr{S}} = \{\ldots, A, B, E, \ldots\}$ the set of all *events* (that is, all subsets of $\mathscr{S}$);

$\mathscr{X} = \{\ldots, x, \ldots\}$ a set of *outcomes* or *consequences*; and

$\mathscr{A} = \{\ldots, f(\cdot), g(\cdot), \ldots\}$ the set of *finite-outcome*[9] *acts* on $\mathscr{S}$.

In addition, an event E is said to be *null* if any pair of acts which differ only on E are indifferent. Finally, we write $y \succeq z$ whenever the constant act yielding y for all $s \in \mathscr{S}$ is weakly preferred to the constant act yielding z (we will sometimes refer to this induced relation on $\mathscr{X}$ by the symbol $\succeq_x$). Given this, Savage's axioms are as follows:[10]

AXIOM P1 (Ordering): *The relation $\succeq$ is complete, reflexive, and transitive.*

AXIOM P2 (Sure-Thing Principle): *For all events E and acts $f(\cdot)$, $f^*(\cdot)$, $g(\cdot)$, and $h(\cdot)$,*

$$\begin{bmatrix} f^*(s) & \text{if } s \in E \\ g(s) & \text{if } s \notin E \end{bmatrix} \succeq \begin{bmatrix} f(s) & \text{if } s \in E \\ g(s) & \text{if } s \notin E \end{bmatrix}$$

$$\Rightarrow \quad \begin{bmatrix} f^*(s) & \text{if } s \in E \\ h(s) & \text{if } s \notin E \end{bmatrix} \succeq \begin{bmatrix} f(s) & \text{if } s \in E \\ h(s) & \text{if } s \notin E \end{bmatrix}.$$

[9] $f(\cdot)$ is said to be a *finite-outcome act* if its outcome set $f(\mathscr{S}) = \{f(s) | s \in \mathscr{S}\}$ is finite. For infinite outcome acts, see our discussion of Axiom P7 as well as footnote 11 below.

[10] Savage did not provide explicit names for his individual axioms. The following names are intended to be suggestive of their respective functions, or in the case of P4, to distinguish it from our own, stronger version of this axiom.

This content downloaded from
154.59.124.38 on Thu, 28 Aug 2025 13:46:36 UTC
All use subject to https://about.jstor.org/terms

AXIOM P3 (Eventwise Monotonicity): *For all outcomes x and y, non-null events E and acts $g(\cdot)$,*

$$\begin{bmatrix} x & \text{if } s \in E \\ g(s) & \text{if } s \notin E \end{bmatrix} \succeq \begin{bmatrix} y & \text{if } s \in E \\ g(s) & \text{if } s \notin E \end{bmatrix} \quad \Leftrightarrow \quad x \succeq y.$$

AXIOM P4 (Weak Comparative Probability): *For all events A, B, and outcomes $x^* \succ x$ and $y^* \succ y$,*

$$\begin{bmatrix} x^* & \text{if } A \\ x & \text{if } \sim A \end{bmatrix} \succeq \begin{bmatrix} x^* & \text{if } B \\ x & \text{if } \sim B \end{bmatrix} \Rightarrow \begin{bmatrix} y^* & \text{if } A \\ y & \text{if } \sim A \end{bmatrix} \succeq \begin{bmatrix} y^* & \text{if } B \\ y & \text{if } \sim B \end{bmatrix}.$$

AXIOM P5 (Nondegeneracy): *There exist outcomes x and y such that $x \succ y$.*

AXIOM P6 (Small Event Continuity): *For any acts $f(\cdot) \succ g(\cdot)$ and outcome x, there exists a finite set of events $\{A_1, \ldots, A_n\}$ forming a partition of $\mathscr{S}$ such that*

$$f(\cdot) \succ \begin{bmatrix} x & \text{if } s \in A_i \\ g(s) & \text{if } s \notin A_i \end{bmatrix} \quad \text{and} \quad \begin{bmatrix} x & \text{if } s \in A_j \\ f(s) & \text{if } s \notin A_j \end{bmatrix} \succ g(\cdot)$$

$$\text{for all } i, j \in \{1, \ldots, n\}.$$

AXIOM P7 (Uniform Monotonicity): *For all events E and all acts $f(\cdot)$ and $f^*(\cdot)$, if*

$$\begin{bmatrix} f^*(s) & \text{if } s \in E \\ g(s) & \text{if } s \notin E \end{bmatrix} \succeq (\preceq) \begin{bmatrix} x & \text{if } s \in E \\ g(s) & \text{if } s \notin E \end{bmatrix}$$

for all $g(\cdot)$ and each $x \in f(E)$, then

$$\begin{bmatrix} f^*(s) & \text{if } s \in E \\ h(s) & \text{if } s \notin E \end{bmatrix} \succeq (\preceq) \begin{bmatrix} f(s) & \text{if } s \in E \\ h(s) & \text{if } s \notin E \end{bmatrix}$$

for all $h(\cdot)$.

In the above,

$$\begin{bmatrix} f(s) & \text{if } s \in E \\ g(s) & \text{if } s \notin E \end{bmatrix}$$

denotes the act that agrees with $f(\cdot)$ over the event E and with $g(\cdot)$ over the event $\sim E$.

Axiom P1 (ordering) is standard. Axiom P2, the Sure-Thing Principle, states that if two acts imply different *subacts* ($f^*(\cdot)$ versus $f(\cdot)$) over an event E, but the *same* subact over the complementary event $\sim E$, the ranking of these acts will not depend on what this common subact is. This axiom implies that preferences are *separable across mutually exclusive events*, which is the key

This content downloaded from
154.59.124.38 on Thu, 28 Aug 2025 13:46:36 UTC
All use subject to https://about.jstor.org/terms

property of expected utility preferences, either over objective probability distributions or over purely subjective acts.

Axiom P3 (Eventwise Monotonicity) states that replacing any outcome y on a non-null event E by a preferred outcome x always leads to a preferred act. Axiom P4 (Weak Comparative Probability) is crucial for the existence of subjective probabilities. When $x^* \succ x$, we can say that the ranking $[x^* \text{ if } A; x \text{ if } \sim A] \succeq [x^* \text{ if } B; x \text{ if } \sim B]$ "reveals" that the individual believes the event A to be at least as likely as the event B. Axiom P3 states that this revealed likelihood ranking is independent of the specific outcomes used.

Axiom P5 (Nondegeneracy) states that the relation $\succeq$ is not trivial. Axiom P6 (Small Event Continuity) states that for any pair of non-indifferent acts and any outcome x, the set $\mathscr{S}$ can be partitioned into small enough events so that altering either act to equal *x on just one of these events* is not enough to reverse their original ranking. For further discussion of the role of this axiom in the theory of subjective probability, the reader is referred to Savage (1954, pp. 27–43) and Kreps (1988, pp. 122–125). Finally, Axiom P7 (Uniform Monotonicity) states that if, for all $g(\cdot)$, $[f^*(\cdot) \text{ if } E; g(\cdot) \text{ if } \sim E]$ is weakly preferred to $[x \text{ if } E; g(\cdot) \text{ if } \sim E]$ *for each outcome* x of a subact $f(\cdot)$ over E, then the individual will weakly prefer $[f^*(\cdot) \text{ if } E; h(\cdot) \text{ if } \sim E]$ to $[f(\cdot) \text{ if } E; h(\cdot) \text{ if } \sim E]$ for all $h(\cdot)$ (and similarly for weakly not preferred). Since the purpose of this axiom is to extend the theory to the case of infinite-outcome acts,[11] we shall not invoke it in the sequel.

Given the above axioms, Savage's result is as follows.

THEOREM (Savage): *Axioms P1 through P6 imply that there exists a unique, finitely additive, non-atomic*[12] *probability measure $\mu(\cdot)$ on $\mathscr{E}$, and a state-independent utility function $U(\cdot)$ on $\mathscr{X}$, such that the individual ranks finite-outcome acts $f(\cdot)$ on the basis of*

$$\hat{\mathscr{V}}(f(\cdot)) \equiv \int U(f(s)) \cdot d\mu(s) \equiv \sum_{i=1}^{n} U(x_i) \cdot \mu(f^{-1}(x_i)),$$

where $\{x_1, \ldots, x_n\}$ is the outcome set of the act $f(\cdot)$.

As seen in the above equation, there are two ways of calculating the value of the expected utility preference functional $\hat{\mathscr{V}}(f(\cdot))$. It can be determined by integrating over states, weighting the utility of the outcome in each state, $U(f(s))$, with respect to the subjective probability measure $\mu(\cdot)$. Or, it can be determined by summing over the finite set of outcomes $\{x_1, \ldots, x_n\}$ implied by the act $f(\cdot)$, weighting the utility of each outcome, $U(x_i)$, by the subjective probability that it will occur. This subjective probability is the measure that $\mu(\cdot)$ assigns to the set $f^{-1}(x_i) = \{s \mid f(s) = x_i\}$, or in other words, $\mu(f^{-1}(x_i))$.

[11] See Savage (1954, pp. 76–82).
[12] We define a probability measure $\mu(\cdot)$ to be *non-atomic* if, for any event E with $\mu(E) > 0$ and any $\alpha \in (0, 1)$, there exists some event $E^* \subset E$ such that $\mu(E^*) = \alpha \cdot \mu(E)$.

This content downloaded from
154.59.124.38 on Thu, 28 Aug 2025 13:46:36 UTC
All use subject to https://about.jstor.org/terms

This is an extremely important result in the theory of subjective probability.[13] Unlike the approach of Ramsey, it does not posit the principle of expected utility maximization from the outset, but rather derives it from the axioms. And unlike Anscombe-Aumann, it does not require any extraneous objective randomization device, but instead derives the individual's subjective probabilities in a framework of purely subjective uncertainty. As mentioned above, it will form the basis for our own extension.

2.2. *What Would it Mean Not to be Probabilistically Sophisticated?*

Each of the classic choice-theoretic approaches—Ramsey, Anscombe-Aumann, and Savage—characterizes an individual who assigns well-behaved subjective probabilities to events, and who makes use of them in ranking subjectively uncertain acts. In order to appreciate the refutable implications of this hypothesis of "probabilistic sophistication," it is worth considering what it would mean to *violate* this property.

The most well-known examples of such violations are a class of problems due to Ellsberg (1961). Consider, for example, an urn containing ninety balls, identical except for color. You know that exactly thirty of the balls are red. Each of the remaining sixty balls is either black or yellow, but you do not know the relative numbers of black and yellow balls (it could be anywhere from 0:60 to 60:0). You are allowed to draw one ball from the urn. Consider the following four acts (where the act f_1 yields \$100 if you draw a red ball and \$0 if you draw a black or a yellow ball, etc.).

	30	60	
	red	black	yellow
f_1	\$100	\$0	\$0
f_2	\$0	\$100	\$0
f_3	\$100	\$0	\$100
f_4	\$0	\$100	\$100

The typical preferences in this example are $f_1 \succ f_2$ and $f_4 \succ f_3$. However, these preferences are inconsistent with *any* set of subjective probabilities $\{p_r, p_b, p_y\}$ over the events {red, black, yellow}. To see this, note that the ranking $f_1 \succ f_2$ implies $p_r > p_b$, but the ranking $f_4 \succ f_3$ implies $p_b + p_y > p_r + p_y$, which leads to a contradiction. The reader is referred to Ellsberg (1961, pp. 651–653), Chipman (1960, pp. 79–88), Raiffa (1961), Becker and Brownson (1964), Slovic and Tversky (1974), and MacCrimmon and Larsson (1979, §7, 8) for similar observed violations of probabilistic sophistication.

[13] See Savage (1951, 1961, 1967), Hacking (1967), and Shimony (1967) for additional early discussions of this approach; Kreps (1988, Chs. 8–10) and Fishburn (1982, Chs. 9–12) for modern expositions of it; and Luce and Suppes (1965) and Fishburn (1981) for surveys of the tremendous body of research it has inspired.

This content downloaded from
154.59.124.38 on Thu, 28 Aug 2025 13:46:36 UTC
All use subject to https://about.jstor.org/terms

The preferences $f_1 \succ f_2$ and $f_4 \succ f_3$ in this example also violate the Sure-Thing Principle, and hence are not compatible with the expected utility hypothesis. Does this mean that *all* violations of the Sure-Thing Principle (or in other words, all violations of expected utility theory) constitute violations of probabilistic sophistication? We shall see that the answer is "no."[14]

3. PROBABILISTICALLY SOPHISTICATED NON-EXPECTED UTILITY MAXIMIZERS

3.1. *Description and Properties*

Our own setting of purely subjective uncertainty (states, events, outcomes, and acts) is identical to that of Savage as described above. However, since it is our intention to *drop* the property of expected utility maximization but *retain* the property of probabilistic sophistication, it is useful to offer an equivalent description of Savage-type preferences in a manner which clearly separates these two properties:

DEFINITION: Let $\mathscr{P}_0(\mathscr{X}) = \{(x_1, p_1; \ldots; x_m, p_m) | m \geq 1,\ \sum_{i=1}^m p_i = 1,\ x_i \in \mathscr{X},\ p_i \geq 0\}$ denote the set of finite-outcome probability distributions over $\mathscr{X}$. An individual is said to be a *probabilistically sophisticated expected utility maximizer* if there exists a probability measure $\mu(\cdot)$ on $\mathscr{E}$ and an expected utility preference functional $\hat{V}(x_1, p_1; \ldots; x_m, p_m) \equiv \sum_{i=1}^m U(x_i) \cdot p_i$ on $\mathscr{P}_0(\mathscr{X})$, such that their preference relation $\succeq$ over acts can be represented by the preference functional

$$\hat{\mathscr{V}}(f(\cdot)) \equiv \hat{V}\big(x_1, \mu\big(f^{-1}(x_1)\big); \ldots; x_n, \mu\big(f^{-1}(x_n)\big)\big),$$

where $\{x_1, \ldots, x_n\}$ is the outcome set of the act $f(\cdot)$.

Such an individual accordingly uses (or acts as if using) the subjective probability measure $\mu(\cdot)$ to determine the probability distribution over consequences implied by any act $f(\cdot)$, and compares alternative acts solely on the basis of their induced probability distributions over consequences, using an expected utility preference functional $\hat{V}(\cdot)$.

Our notion of a probabilistically sophisticated *non*-expected utility maximizer differs only in that the preference functional $V(x_1, p_1; \ldots; x_m, p_m)$ over probability distributions is not (necessarily) expected utility. Accordingly, it is useful to identify the senses in which non-expected utility preference functionals over probability distributions differ from expected utility preference functionals, and

[14] A final issue, to which we shall only allude, is that of *state-independence* in the theory of subjective probability under purely subjective uncertainty. As is well known, unless preferences are state-independent, subjective probabilities cannot be uniquely defined (see Arrow (1974) and Kadane and Winkler (1988) on the implications of this problem for applied decision analysis). In the present paper we follow Savage by adopting axioms sufficient to imply the existence of state-independent preferences, and hence a unique subjective probability measure. See Karni, Schmeidler, and Vind (1983) and Karni (1991) for treatments of this problem in the case of mixed subjective-objective uncertainty, and Karni and Schmeidler (1992) for a representation of preferences in Savage's framework with non-unique subjective probability.

This content downloaded from
154.59.124.38 on Thu, 28 Aug 2025 13:46:36 UTC
All use subject to https://about.jstor.org/terms

the senses in which they are similar. The key *difference* is that any expected utility preference functional $\hat{V}(x_1, p_1; \ldots; x_m, p_m) \equiv \Sigma U(x_i) \cdot p_i$ will be "linear in the probabilities," whereas non-expected utility preference functionals needn't be. In other words, non-expected utility preference functionals will typically violate the Marschak (1950)/Samuelson (1952) "Independence Axiom."[15] The key *similarities* are that we shall retain the same monotonicity and continuity properties as are generically exhibited by expected utility preference functionals. Since these two concepts appear in our main result, we review them here.

The standard notion of monotonicity for preference functionals over probability distributions is monotonicity with respect to first order stochastic dominance. Although first order stochastic dominance is well-known for the case of univariate distributions, its extension to distributions over arbitrary outcome sets $\mathscr{X}$ is less so:[16]

DEFINITION: A distribution $P = (x_1, p_1; \ldots; x_m, p_m)$ is said to *first order stochastically dominate* $Q = (y_1, q_1; \ldots; y_n, q_n)$ with respect to the order $\succeq_x$ if

$$\sum_{\{i \mid x_i \preceq_x x\}} p_i \leqslant \sum_{\{j \mid y_j \preceq_x x\}} q_j \qquad \text{for all } x \in \mathscr{X},$$

and P is said to *strictly first order stochastically dominate* Q if the above holds with strict inequality for some $x^* \in \mathscr{X}$.

When there is no ambiguity regarding the order $\succeq_x$, we shall say simply that *P stochastically dominates Q*. Given this, our definition of monotonicity for preference functionals over probability distributions is as follows:

DEFINITION: Given an order $\succeq_x$ over the outcome set $\mathscr{X}$, $V(\cdot)$ is said to be *monotonic with respect to stochastic dominance* if $V(P)\,(>)\geqslant V(Q)$ whenever P (strictly) stochastically dominates Q.[17]

The standard notion of continuity for expected utility preferences over probability distributions is mixture continuity. The $\lambda : (1-\lambda)$ probability mixture of the probability distributions $P = (x_1, p_1; \ldots; x_m, p_m)$ and $Q = (y_1, q_1; \ldots; y_n, q_n)$, written $\lambda \cdot P + (1-\lambda) \cdot Q$, is defined as the distribution

[15] Formally, this axiom states that a probability distribution P^* is weakly preferred to P if and only if the mixture $\lambda \cdot P^* + (1-\lambda) \cdot Q$ is weakly preferred to $\lambda \cdot P + (1-\lambda) \cdot Q$ for all P^*, P, Q and $\lambda \in (0,1]$, where the mixture $\lambda \cdot P^* + (1-\lambda) \cdot Q$ is as defined below.

[16] See Fishburn and Vickson (1978, §2.21) for a discussion of this more general concept of stochastic dominance.

[17] Similarly, we say that a preference relation $\succeq_{\not{p}}$ over probability distributions is *monotonic with respect to stochastic dominance* if $P\,(\succ_{\not{p}})\succeq_{\not{p}} Q$ whenever P (strictly) stochastically dominates Q. It is important to note that the property of monotonicity over general probability distributions does not necessarily possess the same normative strength as does monotonicity over univariate distributions (e.g., wealth lotteries). See Grant (1991) for an extension of our approach to preferences which are not necessarily monotonic in this sense.

This content downloaded from
154.59.124.38 on Thu, 28 Aug 2025 13:46:36 UTC
All use subject to https://about.jstor.org/terms

$(x_1, \lambda p_1; \ldots; x_m, \lambda p_m; y_1, (1-\lambda)q_1; \ldots; y_n, (1-\lambda)q_n)$. Given this, we have the following:

DEFINITION: $V(\cdot)$ is said to be *mixture continuous* if, for any distributions P, Q, and R in $\mathscr{P}_0(\mathscr{X})$, the sets $\{\lambda \in [0,1] | V(\lambda \cdot P + (1-\lambda) \cdot Q) \geqslant V(R)\}$ and $\{\lambda \in [0,1] | V(\lambda \cdot P + (1-\lambda) \cdot Q) \leqslant V(R)\}$ are both closed.[18]

Given this, we define a probabilistically sophisticated non-expected utility maximizer as follows:

DEFINITION: An individual is said to be a *probabilistically sophisticated non-expected utility maximizer* if there exists a probability measure $\mu(\cdot)$ on $\mathscr{E}$ and a non-expected utility preference functional $V(x_1, p_1; \ldots; x_m, p_m)$ on $\mathscr{P}_0(\mathscr{X})$ satisfying mixture continuity and monotonicity with respect to stochastic dominance, such that their preference relation $\succeq$ over acts can be represented by the preference functional

$$\mathscr{V}(f(\cdot)) \equiv V\big(x_1, \mu\big(f^{-1}(x_1)\big); \ldots; x_n, \mu\big(f^{-1}(x_n)\big)\big),$$

where $\{x_1, \ldots, x_n\}$ is the outcome set of the act $f(\cdot)$.

Note that this definition (as well as its expected utility counterpart) implies that preferences are "state-independent," in that the effect of assigning an outcome x to an event E depends solely on the resulting contribution $\mu(E)$ to the overall probability $\mu(f^{-1}(x))$ of obtaining x, rather than on any specific "state-dependence" between x and E.

For a simple example of probabilistically sophisticated non-expected utility preferences, let the outcome set be $\mathscr{X} = \{1, 2, 3\}$ and the state space be $\mathscr{S} = [0, 1]$, and consider an individual whose subjective probability measure $\mathscr{L}(\cdot)$ is Lebesgue measure[19] on $[0, 1]$, and whose preference functional over probability distributions on $\mathscr{X}$ is $V(p_1, p_2, p_3) \equiv [1 + p_1]^2 + [3 + p_2]^2 + [5 + p_3]^2$, where $p_i = \text{prob}\,(i)$. This preference functional is clearly mixture continuous, and since it satisfies

$$\partial V(p_1, p_2, p_3)/\partial p_1 < \partial V(p_1, p_2, p_3)/\partial p_2 < \partial V(p_1, p_2, p_3)/\partial p_3$$

for all (p_1, p_2, p_3), all first order stochastically dominating transfers of probability mass from a less preferred outcome to a more preferred outcome will be strictly preferred. Such an individual will evaluate acts on the basis of the

[18] Similarly, we say that a preference relation $\succeq_{\mathscr{P}}$ over probability distributions is *mixture continuous* if, for any P, Q, and R, the sets $\{\lambda \in [0,1] | \lambda \cdot P + (1-\lambda) \cdot Q \succeq_{\mathscr{P}} R\}$ and $\{\lambda \in [0,1] | \lambda \cdot P + (1-\lambda) \cdot Q \preceq_{\mathscr{P}} R\}$ are both closed. Note that if $V(\cdot)$ is mixture continuous in this sense, then $V(P) > V(R) > V(Q)$ implies that there exists some $\lambda \in (0,1)$ such that $V(R) = V(\lambda \cdot P + (1-\lambda) \cdot Q)$, and similarly for the preference relation $\succeq_{\mathscr{P}}$.

[19] Formally, since standard Lebesgue measure is not defined over all subsets of $[0,1]$, we must either restrict our class of events $\mathscr{E}$ in this example to the set of Lebesgue measurable sets, or else use some finitely-additive extension of Lebesgue measure to the class of all subsets of $[0,1]$ (e.g., Savage (1954, p. 41)). This also applies to the measure $\mathscr{L}(\cdot)$ in our examples of Sections 3.2 and 3.3.

This content downloaded from
154.59.124.38 on Thu, 28 Aug 2025 13:46:36 UTC
All use subject to https://about.jstor.org/terms

preference functional

$$\mathscr{V}(f(\cdot)) \equiv \left[1+\mathscr{L}(f^{-1}(x_1))\right]^2 + \left[3+\mathscr{L}(f^{-1}(x_2))\right]^2 + \left[5+\mathscr{L}(f^{-1}(x_3))\right]^2.$$

Note that, as with any probabilistically sophisticated preference functional, the act $f(\cdot)$ enters only through its implied probabilities $\mathscr{L}(f^{-1}(x_i))$ of the respective outcomes.

What are the key similarities and differences between probabilistically sophisticated expected utility and probabilistically sophisticated non-expected utility preferences over acts? The fundamental difference is that expected utility preferences satisfy the Sure-Thing Principle (Axiom P2), whereas non-expected utility preferences need not.[20] However, just as with non-expected utility preferences over probability distributions, probabilistically sophisticated non-expected utility preferences over acts exhibit the same monotonicity and continuity properties as their expected utility counterpart. The appropriate notion of monotonicity for preferences over finite-outcome acts is Eventwise Monotonicity (Axiom P3). Since an eventwise dominating act will induce a stochastically dominating probability distribution over outcomes, monotonicity with respect to stochastic dominance ensures that probabilistically sophisticated non-expected utility preferences over acts will be eventwise monotonic. The appropriate notion of continuity, Savage's Axiom P6, will also be satisfied by probabilistically sophisticated non-expected utility preferences over acts, as long as the subjective probability measure $\mu(\cdot)$ is non-atomic and the preference functional $V(\cdot)$ over probability distributions is mixture continuous.

What about the other Savage axioms? Axiom P1 (Ordering) follows from the fact that probabilistically sophisticated non-expected utility preferences are derived from a real-valued preference functional $\mathscr{V}(\cdot)$ over acts. Axiom P5 (Nondegeneracy) follows in any nontrivial instance. For finite-outcome acts, Axiom P7 (Uniform Monotonicity) also follows from monotonicity with respect to stochastically dominating probability distributions. To verify Axiom P4 (Weak Comparative Probability), note that $[x^*$ if A; x if $\sim A] \succeq [x^*$ if B; x if $\sim B]$ implies $V(x^*, \mu(A); x, \mu(\sim A)) \geqslant V(x^*, \mu(B); x, \mu(\sim B))$ which, given $x^* \succ x$ and monotonicity with respect to stochastic dominance, implies $\mu(A) \geqslant \mu(B)$. Given $y^* \succ y$, monotonicity accordingly implies that $V(y^*, \mu(A); y, \mu(\sim A)) \geqslant V(y^*, \mu(B); y, \mu(\sim B))$, which in turn implies that $[y^*$ if A; y if $\sim A] \succeq [y^*$ if B; y if $\sim B]$.

On the other hand, the hypothesis of probabilistically sophisticated non-expected utility preferences *is* contradicted by the typical behavior in the Ellsberg Paradox. This is to be expected: as noted above, such behavior is simply inconsistent with the existence of subjective probabilities, whether or not risk preferences are expected utility.

[20] For example, defining the events $A = [0, .8)$, $B = [.8, .9)$, and $C = [.9, 1]$, the individual in the previous paragraph prefers the act {\$2 if A; \$2 if B; \$2 if C} to {\$2 if A; \$1 if B; \$3 if C}, but prefers {\$3 if A; \$1 if B; \$3 if C} to {\$3 if A; \$2 if B; \$2 if C}.

This content downloaded from
154.59.124.38 on Thu, 28 Aug 2025 13:46:36 UTC
All use subject to https://about.jstor.org/terms

To sum up: As long as the underlying preference functional over probability distributions satisfies monotonicity with respect to stochastic dominance and mixture continuity, probabilistically sophisticated non-expected utility preferences over acts satisfy eventwise monotonicity, continuity, and all of the Savage axioms, except the expected utility based Sure-Thing Principle.

3.2. *Natural Instances of Probabilistically Sophisticated Non-Expected Utility Preferences*

Although the Savage framework is usually thought of as "timeless," real-world situations of decision-making under uncertainty invariably involve at least some delay between the time a choice must be made and the time the uncertainty is actually resolved. If a Savage-type individual should happen to have any *other* decisions to make during this interim period (even if they are only consumption/savings decisions), he or she will end up exhibiting probabilistically sophisticated *non*-expected utility preferences over acts.

To see this, take an individual with von Neumann-Morgenstern utility function $U(x,\alpha)$ and subjective probability measure $\mu(\cdot)$, who must choose from a set of delayed-resolution acts $\{f(\cdot)\}$, and whose interim decisions are represented by the choice of some element α from a set A. His or her expected utility for a given choice of $f(\cdot)$ and α is

$$\hat{\mathscr{V}}(f(\cdot),\alpha)=\int U(f(s),\alpha)\cdot d\mu(s)=\sum_i U(x_i,\alpha)\cdot\mu\big(f^{-1}(x_i)\big).$$

How will such an individual rank delayed-resolution acts? Since the choice of α must be made *before* the state is known, he or she will rank them on the basis of the induced preference functional

$$\begin{aligned}\mathscr{W}(f(\cdot))&\equiv\max_{\alpha\in A}\big[\hat{\mathscr{V}}(f(\cdot),\alpha)\big]=\max_{\alpha\in A}\Big[\sum_i U(x_i,\alpha)\cdot\mu\big(f^{-1}(x_i)\big)\Big]\\&\equiv\sum_i U\big(x_i,\hat{\alpha}_f(f(\cdot))\big)\cdot\mu\big(f^{-1}(x_i)\big),\end{aligned}$$

where

$$\begin{aligned}\hat{\alpha}_f(f(\cdot))&\equiv\hat{\alpha}_P\big(x_1,\mu\big(f^{-1}(x_1)\big);\ldots;x_n,\mu\big(f^{-1}(x_n)\big)\big)\\&\equiv\operatorname*{argmax}_{\alpha\in A}\Big[\sum_i U(x_i,\alpha)\cdot\mu\big(f^{-1}(x_i)\big)\Big].\end{aligned}$$

Note that the optimal interim decision $\hat{\alpha}_f(f(\cdot))$ depends upon $f(\cdot)$ only through its outcomes $\{x_1,\ldots,x_n\}$ and their respective probabilities $\{\mu(f^{-1}(x_1)),\ldots,\mu(f^{-1}(x_n))\}$. The preference functional $\mathscr{W}(\cdot)$ accordingly induces probabilistically sophisticated non-expected utility preferences over acts, with the subjective probability measure $\mu(\cdot)$ and non-expected utility preference func-

This content downloaded from
154.59.124.38 on Thu, 28 Aug 2025 13:46:36 UTC
All use subject to https://about.jstor.org/terms

tion over probability distributions given by

$$V(x_1, p_1; \ldots; x_n, p_n) \equiv \sum_{i=1}^{n} U(x_i, \hat{\alpha}_P(x_1, p_1; \ldots; x_n, p_n)) \cdot p_i.$$

Thus, to the extent that real-world decisions involve delayed-resolution uncertainty, even Savage-type individuals will exhibit probabilistically sophisticated non-expected utility preferences over acts.[21]

3.3. *Do the Savage Axioms Minus the Sure-Thing Principle Imply Probabilistic Sophistication*?

Since probabilistically sophisticated non-expected utility preferences satisfy each of the Savage axioms except the Sure-Thing Principle, it is natural to ask whether these remaining axioms are sufficient to *imply* probabilistically sophistication. To show that this is not true, we offer a preference functional over acts which is *not* probabilistically sophisticated, but which nevertheless satisfies all of the Savage axioms except the Sure-Thing Principle.

Let the outcome space be the interval $[0,100]$, let the state space be the interval $[0,1]$ with Lebesgue measure $\mathscr{L}(\cdot)$, and define the *non-additive measure* $\mu(\cdot)$ on $[0,1]$ by

$$\mu(A) \equiv 3 \cdot \left[\mathscr{L}\left(A \cap \left[0, \tfrac{1}{3}\right]\right)\right]^2 + \tfrac{3}{2} \cdot \left[\mathscr{L}\left(A \cap \left(\tfrac{1}{3}, 1\right]\right)\right]^2.$$

Note that $\mu(\cdot)$ satisfies monotonicity with respect to set inclusion, and that

$$\mu(\phi) = 0, \qquad \mu\left(\left[0, \tfrac{1}{3}\right]\right) = \tfrac{1}{3}, \qquad \mu\left(\left(\tfrac{1}{3}, 1\right]\right) = \tfrac{2}{3}, \quad \text{and} \quad \mu([0,1]) = 1.$$

But since

$$\mu\left(\left(\tfrac{1}{3}, \tfrac{2}{3}\right]\right) = \tfrac{1}{6} \qquad \text{and} \qquad \mu\left(\left[0, \tfrac{1}{3}\right] \cup \left(\tfrac{2}{3}, 1\right]\right) = \tfrac{1}{2},$$

it follows that $\mu(\cdot)$ is not additive, and hence not a probability measure.

Given the non-additive measure $\mu(\cdot)$, define the preference functional $\mathscr{V}(\cdot)$ over finite-outcome acts by the Choquet (1953–54) integral (e.g., Schmeidler (1982, 1989)):

$$\mathscr{V}(f(\cdot)) \equiv \int_0^{100} \mu(f^{-1}([x, 100]))\, dx.$$

Since $f(s) \geqslant g(s)$ for all $s \in [0,1]$ implies $f^{-1}([x,100]) \supseteq g^{-1}([x,100])$ and hence $\mu(f^{-1}([x,100])) \geqslant \mu(g^{-1}([x,100]))$ for all $x \in [0,100]$, $\mathscr{V}(\cdot)$ exhibits eventwise monotonicity. If $f(\cdot)$ is a two-outcome act with outcomes $x_1 < x_2$, the formula

[21] One might argue that in this setting the "outcomes" are really the (x, α) pairs and the "acts" are really the $(f(\cdot), \alpha)$ pairs, so that the preference functional $\hat{\mathscr{V}}(f(\cdot), \alpha) \equiv \int U(f(x(s)), \alpha) \cdot d\mu(s)$ over such "extended acts" remains expected utility. However this approach requires that the outcomes and acts be defined to include *every interim choice* between the choice of $f(\cdot)$ and the revelation of the state s. As Kreps and Porteus (1979, p. 83) have noted, "the obvious difficulty with this approach is that such complete models may become overburdened with detail and analytically intractable."

This content downloaded from
154.59.124.38 on Thu, 28 Aug 2025 13:46:36 UTC
All use subject to https://about.jstor.org/terms

for $\mathscr{V}(\cdot)$ reduces to

$$\mathscr{V}(f(\cdot)) \equiv x_1 \cdot \mu(f^{-1}([x_1, 100])) + (x_2 - x_1) \cdot \mu(f^{-1}([x_2, 100]))$$
$$\equiv x_1 + (x_2 - x_1) \cdot \mu(f^{-1}(x_2)).$$

Consider the following Ellsberg Paradox-type acts:

	$[0, \frac{1}{3}]$	$(\frac{1}{3}, \frac{2}{3}]$	$(\frac{2}{3}, 1]$
f_1	100	0	0
f_2	0	100	0
f_3	100	0	100
f_4	0	100	100

Applying the above formula for two-outcome acts, we obtain

$$\mathscr{V}(f_1) = 100 \cdot \mu([0, \tfrac{1}{3}]) = 33\tfrac{1}{3},$$
$$\mathscr{V}(f_2) = 100 \cdot \mu((\tfrac{1}{3}, \tfrac{2}{3}]) = 16\tfrac{2}{3},$$
$$\mathscr{V}(f_3) = 100 \cdot \mu([0, \tfrac{1}{3}] \cup (\tfrac{2}{3}, 1]) = 50,$$
$$\mathscr{V}(f_4) = 100 \cdot \mu((\tfrac{1}{3}, 1]) = 66\tfrac{2}{3}.$$

We thus have $f_1 \succ f_2$ but $f_4 \succ f_3$, which violates probabilistic sophistication, since (under the assumption of monotonicity with respect to stochastic dominance) the ranking $f_1 \succ f_2$ implies that the event $[0, \frac{1}{3}]$ must have a strictly greater probability than $(\frac{1}{3}, \frac{2}{3}]$, whereas the ranking $f_4 \succ f_3$ implies the opposite. Nevertheless, the reader may verify that $\mathscr{V}(\cdot)$ satisfies all of the Savage axioms other than the Sure-Thing Principle.

3.4. *Do Probabilistically Sophisticated Betting Preferences Imply Probabilistic Sophistication?*

As mentioned in the Introduction, the so-called "intuitive" or "qualitative" probability axioms on a comparative likelihood relation $\succeq_l$ offer a characterization of probabilistically sophisticated *beliefs* which neither invokes nor implies the expected utility hypothesis. If we were to derive the relation $\succeq_l$ from preferences over subjectively uncertain acts, and assume that it satisfied all of the qualitative probability axioms, wouldn't this lead to a characterization of probabilistically sophisticated *preferences* which did not rely upon the expected utility hypothesis?

More specifically, define "$A \succeq_l B$" if the individual would always weakly prefer to bet on the event A rather than on the event B, that is, weakly prefers the act $[x \text{ if } A;\ y \text{ if } \sim A]$ to the act $[x \text{ if } B;\ y \text{ if } \sim B]$ for all $x \succ y$. Assume that $\succeq_l$ is complete (so that such likelihood rankings are always well-defined) and that $\succeq_l$ satisfies all of the qualitative probability axioms, so that it can be represented by a probability measure $\mu(\cdot)$ on events. Will this imply that preferences over general, many-outcome acts are probabilistically sophisticated?

This content downloaded from
154.59.124.38 on Thu, 28 Aug 2025 13:46:36 UTC
All use subject to https://about.jstor.org/terms

Again, the answer is no. To show this, we construct a preference functional $\mathscr{V}(f(\cdot))$ over acts whose ranking of two-outcome "bets" is consistent with a well-defined subjective probability distribution over events, but which can exhibit Ellsberg Paradox-type preferences over more general, many-outcome acts. Let the state space be the interval $[0,1]$ with Lebesgue measure $\mathscr{L}(\cdot)$, let the outcome space be $\{0,1,2\}$, and define

$$\mathscr{V}(f(\cdot)) \equiv \int_0^1 f(s)\,ds + \Big[\mathscr{L}(f^{-1}(0))\cdot\mathscr{L}(f^{-1}(1)) \cdot\mathscr{L}(f^{-1}(2))\cdot\int_0^1 s\cdot f(s)\,ds\Big].$$

This preference function exhibits eventwise monotonicity.[22] When restricted to the class of *two-outcome* acts (in which case $\mathscr{L}(f^{-1}(0))\cdot\mathscr{L}(f^{-1}(1))\cdot\mathscr{L}(f^{-1}(2)) = 0$), it reduces to the formula $\int_0^1 f(s)\,ds$, that is, to the expected value of $f(\cdot)$ with respect to the uniform measure on $[0,1]$. This implies that for any pair of outcomes $x > y$ from $\{0,1,2\}$, the relation $\succeq_l$ derived from preferences over bets involving these outcomes coincides with the uniform measure on $[0,1]$. However, given the set of acts

	$[0,.305)$	$[.305,.609)$	$[.609,1]$
f_1	2	1	1
f_2	1	2	1
f_3	2	1	0
f_4	1	2	0

we have $\mathscr{V}(f_1) = 1.3050$, $\mathscr{V}(f_2) = 1.3040$, $\mathscr{V}(f_3) = .9224$, and $\mathscr{V}(f_4) = .9248$, so that $f_1 \succ f_2$ yet $f_4 \succ f_3$, which violates probabilistic sophistication, since monotonicity with respect to stochastic dominance and $f_1 \succ f_2$ imply that the event $[0,.305)$ must have a strictly greater probability than $[.305,.609)$, which (again by monotonicity) would imply that $f_3 \succ f_4$. The intuition is that no matter how many restrictions we place on preferences over "bets," these restrictions only apply to preferences over the set of two-outcome acts, which is not enough to imply probabilistic sophistication over general, many-outcome acts.

4. A MORE ROBUST CHARACTERIZATION OF SUBJECTIVE PROBABILITY

4.1. *The Strong Comparative Probability Axiom*

The counterexample of Section 3.3 demonstrated that in the absence of the Sure-Thing Principle, the Weak Comparative Probability Axiom (Axiom P4) is *not* strong enough to ensure probabilistic sophistication, even in the presence of

[22] To see this, take any act $f(\cdot)$ and raise its payoff by one unit over some event $A \subseteq f^{-1}(0)$ or $A \subseteq f^{-1}(1)$. The integral $\int f(s)\,ds$ will rise by $\mathscr{L}(A)$, and if the term in square brackets drops, it will do so by strictly less than $\mathscr{L}(A)$, so $\mathscr{V}(\cdot)$ will strictly increase.

This content downloaded from
154.59.124.38 on Thu, 28 Aug 2025 13:46:36 UTC
All use subject to https://about.jstor.org/terms

the remaining Savage axioms. In order to characterize probabilistically sophisticated non-expected utility preferences, it is necessary to strengthen the Weak Comparative Probability Axiom. We thus offer the following axiom:

AXIOM P4* (Strong Comparative Probability): *For all pairs of disjoint events A and B, outcomes $x^* \succ x$ and $y^* \succ y$, and acts $g(\cdot)$ and $h(\cdot)$,*

$$\begin{bmatrix} x^* & \text{if } s \in A \\ x & \text{if } s \in B \\ g(s) & \text{if } s \notin A \cup B \end{bmatrix} \succeq \begin{bmatrix} x & \text{if } s \in A \\ x^* & \text{if } s \in B \\ g(s) & \text{if } s \notin A \cup B \end{bmatrix}$$

$$\Rightarrow \begin{bmatrix} y^* & \text{if } s \in A \\ y & \text{if } s \in B \\ h(s) & \text{if } s \notin A \cup B \end{bmatrix} \succeq \begin{bmatrix} y & \text{if } s \in A \\ y^* & \text{if } s \in B \\ h(s) & \text{if } s \notin A \cup B \end{bmatrix}.$$

The intuition behind this axiom is similar to that of the Weak Comparative Probability Axiom: any individual exhibiting the upper preference ranking is revealing A to be at least as likely as B in the case when the relevant outcomes are $x^* \succ x$ and the complementary event $\sim(A \cup B)$ yields the subact $g(\cdot)$. If he or she possesses a unique subjective probability distribution over events *and uses it* in the evaluation of acts (in other words, if they are probabilistically sophisticated), he or she must continue to reveal A as being at least as likely as B for any other outcomes $y^* \succ y$ and subact $h(\cdot)$.

Note that none of the events A, B or $\sim(A \cup B)$ in this axiom are required to be nonempty. When $\sim(A \cup B)$ is empty, the intuition is as given in the above paragraph. When B is empty, the axiom follows from Eventwise Monotonicity (Axiom P3). When A is empty, it follows from Axiom P3 that any B satisfying the upper preference ranking must be null, which implies the lower preference ranking. When any *two* of the events A, B, or $\sim(A \cup B)$ are empty, the axiom is trivial.

For the remainder of this paper, we shall use the symbol $\succeq_l$ to denote the revealed likelihood relation implied by the Strong Comparative Probability Axiom—in other words, we write "$A \succeq_l B$" when the pair of disjoint events A and B satisfy

$$\begin{bmatrix} x^* & \text{if } s \in A \\ x & \text{if } s \in B \\ g(s) & \text{if } s \notin A \cup B \end{bmatrix} \succeq \begin{bmatrix} x & \text{if } s \in A \\ x^* & \text{if } s \in B \\ g(s) & \text{if } s \notin A \cup B \end{bmatrix}$$

for all outcomes $x^* \succ x$ and acts $g(\cdot)$. When A and B are not disjoint, we write "$A \succeq_l B$" if and only if $(A - B) \succeq_l (B - A)$. The reader may verify that the Strong Comparative Probability Axiom implies that the relation $\succeq_l$ over events will inherit the properties of completeness, reflexivity, and transitivity[23] from the preference relation $\succeq$ over acts.

[23] Below we show that under P4*, $A \succeq_l B$ is equivalent to $[x^* \text{ on } A;\ x \text{ on } \sim A] \succeq [x^* \text{ on } B;\ x \text{ on } \sim B]$ for any A, B and $x^* \succ x$. Thus $A \succeq_l B$ and $B \succeq_l C$ implies $[x^* \text{ on } A;\ x \text{ on } \sim A] \succeq [x^* \text{ on } B;\ x \text{ on } \sim B] \succeq [x^* \text{ on } C;\ x \text{ on } \sim C]$ which by transitivity of $\succeq$ implies $A \succeq_l C$.

This content downloaded from
154.59.124.38 on Thu, 28 Aug 2025 13:46:36 UTC
All use subject to https://about.jstor.org/terms

The typical preferences of $f_1 \succ f_2$ and $f_4 \succ f_3$ in the Ellsberg Paradox violate the Strong Comparative Probability Axiom.[24] This is to be expected—the purpose of P4* is to characterize probabilistically sophisticated preferences, and we have already seen that Ellsberg Paradox-type preferences are *not* probabilistically sophisticated. It is interesting to note that, although Ellsberg Paradox-type preferences violate Savage's Sure-Thing Principle P2, they *do not* violate his Weak Comparative Probability Axiom P4.[25]

To see that P4* implies Savage's Weak Comparative Probability Axiom P4, choose any pair of (not necessarily disjoint) events A and B and outcomes $x^* \succ x$ and $y^* \succ y$. By equivalent description of acts and P4* we have

$$\begin{bmatrix} x^* & \text{if } A \\ x & \text{if } \sim A \end{bmatrix} \succeq \begin{bmatrix} x^* & \text{if } B \\ x & \text{if } \sim B \end{bmatrix}$$

$$\Leftrightarrow \begin{bmatrix} x^* & \text{if } s \in A - B \\ x & \text{if } s \in B - A \\ x^* & \text{if } s \in A \cap B \\ x & \text{if } s \in \sim A \cap \sim B \end{bmatrix} \succeq \begin{bmatrix} x & \text{if } s \in A - B \\ x^* & \text{if } s \in B - A \\ x^* & \text{if } s \in A \cap B \\ x & \text{if } s \in \sim A \cap \sim B \end{bmatrix}$$

$$\Rightarrow \begin{bmatrix} y^* & \text{if } s \in A - B \\ y & \text{if } s \in B - A \\ y^* & \text{if } s \in A \cap B \\ y & \text{if } s \in \sim A \cap \sim B \end{bmatrix} \succeq \begin{bmatrix} y & \text{if } s \in A - B \\ y^* & \text{if } s \in B - A \\ y^* & \text{if } s \in A \cap B \\ y & \text{if } s \in \sim A \cap \sim B \end{bmatrix}$$

$$\Leftrightarrow \begin{bmatrix} y^* & \text{if } A \\ y & \text{if } \sim A \end{bmatrix} \succeq \begin{bmatrix} y^* & \text{if } B \\ y & \text{if } \sim B \end{bmatrix},$$

which implies the Weak Comparative Probability Axiom.

4.2. *Comparison with the Sure-Thing Principle*

How does the Strong Comparative Probability Axiom differ from the Sure-Thing Principle? Structurally, the axioms are very similar. The key distinction between them can be illustrated by comparing the paradoxes of Allais (1953) and Ellsberg (1961):

Allais Paradox

	#1	#2–#11	#12–#100
a_1	\$1M	\$1M	\$1M
a_2	\$0	\$5M	\$1M
a_4	\$1M	\$1M	\$0
a_3	\$0	\$5M	\$0

Ellsberg Paradox

	red	black	yellow
f_1	\$100	\$0	\$0
f_2	\$0	\$100	\$0
f_3	\$100	\$0	\$100
f_4	\$0	\$100	\$100

[24] Let $A = \{\text{red ball}\}$, $B = \{\text{black ball}\}$, $x^* = y^* = \$100$, $x = y = \$0$, $g(s) \equiv \$0$, and $h(s) \equiv \$100$.

[25] Say the individual believed there were 30 balls of each type, let $\mathscr{L}(\cdot)$ be the counting measure over balls, and define the non-additive measure $\mu(\cdot)$ from $\mathscr{L}(\cdot)$ and the preference functional $\mathscr{V}(\cdot)$ from $\mu(\cdot)$ as in Section 3.3. We have seen that this preference functional over acts satisfies the Weak Comparative Probability Axiom, yet it exhibits the Ellsberg Paradox-type rankings $f_1 \succ f_2$ and $f_4 \succ f_3$.

This content downloaded from
154.59.124.38 on Thu, 28 Aug 2025 13:46:36 UTC
All use subject to https://about.jstor.org/terms

The left-hand table illustrates the four choices of the famous Allais Paradox, converted from their usual probability distribution format into Savage-type acts. Here, the individual faces an urn of 100 numbered balls, so that, for example, the act a_4 yields \$1M if ball #1 through #11 is drawn, otherwise \$0, which implies a .11:.89 chance of winning \$1M or \$0 (where \$1M = \$1,000,000).[26] In order to highlight their structural similarity with the Ellsberg acts, these acts are listed in the order a_1, a_2, a_4, and a_3.

The Strong Comparative Probability Axiom implies $f_1 \succeq f_2$ if and only if $f_3 \succeq f_4$, and the Sure-Thing Principle implies $a_1 \succeq a_2$ if and only if $a_4 \succeq a_3$. Both axioms thus impose some form of "consistency" on preferences across different pairs of acts. The key distinction is whether this consistency involves *beliefs* or *preferences*. The ranking of a_1 versus a_2 reflects both the individual's beliefs over the relative likelihoods of the events #1 and #2–#11 *as well as his or her risk preferences*: specifically, whether he or she prefers receiving \$1M across both events to the riskier prospect of obtaining possibly \$5M, but possibly \$0. The Sure–Thing Principle thus imposes consistency on beliefs *and* risk preferences across the pairs $\{a_1, a_2\}$ and $\{a_4, a_3\}$, which is why it serves to jointly characterize probabilistically sophisticated beliefs and expected utility preferences.

On the other hand, the Strong Comparative Probability Axiom places no restrictions on preferences over the acts $\{a_1, a_2\}$ versus $\{a_4, a_3\}$. It only applies to pairs of acts that take the more specialized form of $\{f_1, f_2\}$ and $\{f_3, f_4\}$, which differ by exchanging the outcomes on a pair of events, in this case, red and black. But an individual's ranking of f_1 and f_2 reflects only their *probabilistic beliefs*, and nothing concerning their *risk preferences*. Thus, the Strong Comparative Probability Axiom only imposes consistency of *beliefs* across pairs of acts, which is why it can characterize probabilistically sophisticated beliefs without implying the expected utility hypothesis on preferences.

What is the logical relationship between the Sure-Thing Principle and the Strong Comparative Probability Axiom? Provided the set $\mathscr{X}$ contains at least three outcomes, the two axioms are logically independent. The example of Section 3.1 satisfies the Strong Comparative Probability Axiom but not the Sure-Thing Principle. For an example of the opposite, consider the outcome set $\mathscr{X} = \{x, y, z\}$, the state space $\mathscr{S} = [0,1]$, and define

$$\mathscr{V}(f(\cdot)) \equiv \int U(f(s), s)\, ds \qquad \text{where}$$

$$\{U(x,s), U(y,s), U(z,s)\} \equiv \begin{cases} \{2,1,2\} & \text{if } s \in \left[0, \frac{1}{3}\right), \\ \{1,2,1\} & \text{if } s \in \left[\frac{1}{3}, \frac{2}{3}\right), \\ \{9,8,7\} & \text{if } s \in \left[\frac{2}{3}, 1\right]. \end{cases}$$

[26] We represent the Allais Paradox in this manner solely to illustrate the theoretical restrictions imposed by the Sure-Thing Principle. See MacCrimmon and Larsson (1979), Moskowitz (1974), Slovic and Tversky (1974), and Keller (1985) for experimental studies of the effect of problem representation on the actual incidence of Allais-type choices (see also the remarks of Machina (1982, p. 289 and footnote 30)).

This content downloaded from
154.59.124.38 on Thu, 28 Aug 2025 13:46:36 UTC
All use subject to https://about.jstor.org/terms

Since this preference function is additively separable across states, it satisfies the Sure-Thing Principle. However, since it implies the rankings $x \succ y$ and $y \succ z$, as well as

$$\begin{bmatrix} x & \text{if } s \in [0, \frac{1}{3}) \\ y & \text{if } s \in [\frac{1}{3}, \frac{2}{3}) \\ z & \text{if } s \in [\frac{2}{3}, 1] \end{bmatrix} \succ \begin{bmatrix} y & \text{if } s \in [0, \frac{1}{3}) \\ x & \text{if } s \in [\frac{1}{3}, \frac{2}{3}) \\ z & \text{if } s \in [\frac{2}{3}, 1] \end{bmatrix} \quad \text{and}$$

$$\begin{bmatrix} y & \text{if } s \in [0, \frac{1}{3}) \\ z & \text{if } s \in [\frac{1}{3}, \frac{2}{3}) \\ z & \text{if } s \in [\frac{2}{3}, 1] \end{bmatrix} \prec \begin{bmatrix} z & \text{if } s \in [0, \frac{1}{3}) \\ y & \text{if } s \in [\frac{1}{3}, \frac{2}{3}) \\ z & \text{if } s \in [\frac{2}{3}, 1] \end{bmatrix},$$

it does not satisfy the Strong Comparative Probability Axiom.[27]

Although logically independent in the case of three or more outcomes, the two axioms are in fact equivalent in any world with exactly two outcomes $z^* \succ z$. Since we use this fact in the proof of our main result (Theorem 2), it is useful to demonstrate it here. To see that Strong Comparative Probability implies the Sure-Thing Principle in such a world, pick any event E and acts $f(\cdot)$, $f^*(\cdot)$, $g(\cdot)$, and $h(\cdot)$. Since the subacts $f^*(\cdot)$ and $f(\cdot)$ over E can be written as $f^*(\cdot) = [z^* \text{ on } E_1;\ z \text{ on } E_2;\ z^* \text{ on } E_3;\ z \text{ on } E_4]$ and $f(\cdot) = [z \text{ on } E_1;\ z^* \text{ on } E_2;\ z^* \text{ on } E_3;\ z \text{ on } E_4]$ for some partition $\{E_1, E_2, E_3, E_4\}$ of E, we have

$$\begin{bmatrix} f^*(s) & \text{if } s \in E \\ g(s) & \text{if } s \notin E \end{bmatrix} \succeq \begin{bmatrix} f(s) & \text{if } s \in E \\ g(s) & \text{if } s \notin E \end{bmatrix}$$

$$\Leftrightarrow \begin{bmatrix} z^* & \text{if } s \in E_1 \\ z & \text{if } s \in E_2 \\ z^* & \text{if } s \in E_3 \\ z & \text{if } s \in E_4 \\ g(s) & \text{if } s \notin E \end{bmatrix} \succeq \begin{bmatrix} z & \text{if } s \in E_1 \\ z^* & \text{if } s \in E_2 \\ z^* & \text{if } s \in E_3 \\ z & \text{if } s \in E_4 \\ g(s) & \text{if } s \notin E \end{bmatrix}$$

$$\Rightarrow \begin{bmatrix} z^* & \text{if } s \in E_1 \\ z & \text{if } s \in E_2 \\ z^* & \text{if } s \in E_3 \\ z & \text{if } s \in E_4 \\ h(s) & \text{if } s \notin E \end{bmatrix} \succeq \begin{bmatrix} z & \text{if } s \in E_1 \\ z^* & \text{if } s \in E_2 \\ z^* & \text{if } s \in E_3 \\ z & \text{if } s \in E_4 \\ h(s) & \text{if } s \notin E \end{bmatrix}$$

$$\Leftrightarrow \begin{bmatrix} f^*(s) & \text{if } s \in E \\ h(s) & \text{if } s \notin E \end{bmatrix} \succeq \begin{bmatrix} f(s) & \text{if } s \in E \\ h(s) & \text{if } s \notin E \end{bmatrix},$$

which is the Sure-Thing Principle.

[27] Since it is "state-dependent," this preference function also violates Axioms P3 and P4. As seen from our Theorem 2, the Sure-Thing Principle *does* imply the Strong Comparative Probability Axiom in the presence of all the other Savage axioms.

This content downloaded from
154.59.124.38 on Thu, 28 Aug 2025 13:46:36 UTC
All use subject to https://about.jstor.org/terms

To see that the Sure-Thing Principle implies the Strong Comparative Probability Axiom in such a world, pick any disjoint events A and B, outcomes $x^* \succ x$ and $y^* \succ y$, and acts $g(\cdot)$ and $h(\cdot)$. Since we have $x^* = y^* = z^*$ and $x = y = z$, the Sure-Thing Principle implies

$$\begin{bmatrix} x^* & \text{if } s \in A \\ x & \text{if } s \in B \\ g(s) & \text{if } s \notin A \cup B \end{bmatrix} \succeq \begin{bmatrix} x & \text{if } s \in A \\ x^* & \text{if } s \in B \\ g(s) & \text{if } s \notin A \cup B \end{bmatrix}$$

$$\Rightarrow \begin{bmatrix} y^* & \text{if } s \in A \\ y & \text{if } s \in B \\ h(s) & \text{if } s \notin A \cup B \end{bmatrix} \succeq \begin{bmatrix} y & \text{if } s \in A \\ y^* & \text{if } s \in B \\ h(s) & \text{if } s \notin A \cup B \end{bmatrix}.$$

This equivalence should come as no surprise. In a world with only two outcomes, all acts consist of two-outcome bets, and all probabilistically sophisticated agents will rank acts solely on the basis of their subjective probability of the preferred outcome. Since risk preferences play no role, it is impossible to distinguish expected utility preferences from probabilistically sophisticated *non*-expected utility preferences, and the Sure-Thing Principle (which characterizes the former type of preferences) will be equivalent to the Strong Comparative Probability Axiom (which characterizes the latter).

4.3. *Formal Characterization*

As mentioned above, our characterization of probabilistic sophistication in the absence of the expected utility hypothesis involves taking the Savage Axioms P1–P6, dropping the Sure–Thing Principle P2, and replacing his Weak Comparative Probability Axiom P4 by our Strong Comparative Probability Axiom P4*.

Theorem 1 captures the spirit of this approach by demonstrating that if preferences satisfy this set of axioms, there exists a subjective probability measure $\mu(\cdot)$ over events which represents the revealed likelihood ranking $\succeq_l$, and such that any two acts that imply the same probability distribution over outcomes will be indifferent.

THEOREM 1: *If the preference relation $\succeq$ over $\mathscr{A}$ satisfies the axioms: P1—Ordering, P3—Eventwise Monotonicity, P4*—Strong Comparative Probability, P5—Nondegeneracy, and P6—Small Event Continuity, then there exists a unique finitely additive, non-atomic probability measure $\mu(\cdot)$ on $\mathscr{E}$ such that $\mu(\cdot)$ represents the comparative likelihood relation $\succeq_l$ from P4*, and such that for any pair of acts $f(\cdot), g(\cdot)$ in $\mathscr{A}$ with respective outcome sets*[28] *$\{x_1, \ldots, x_r\}$ and $\{y_1, \ldots, y_s\}$: if $\mu(f^{-1}(z)) = \mu(g^{-1}(z))$ for each $z \in \{x_1, \ldots, x_r\} \cup \{y_1, \ldots, y_s\}$, then $f(\cdot) \sim g(\cdot)$.*

[28] Note that the outcome sets of the acts $f(\cdot)$ and $g(\cdot)$ needn't be strictly identical, since either act could have outcomes which only occur on sets of zero probability.

This content downloaded from
154.59.124.38 on Thu, 28 Aug 2025 13:46:36 UTC
All use subject to https://about.jstor.org/terms

Theorem 2 provides our formal characterization of a probabilistically sophisticated individual, that is, one who possesses a subjective probability measure $\mu(\cdot)$ over events and evaluates acts on the basis of a non-expected utility preference function $V(\cdot)$ over their implied probability distributions over outcomes.

THEOREM 2: *The following conditions on a preference relation* $\succeq$ *over* $\mathscr{A}$ *are equivalent*:

(i): $\succeq$ *satisfies*: *P1—Ordering*, *P3—Eventwise Monotonicity*, *P4*—Strong Comparative Probability*, *P5—Nondegeneracy*, *P6—Small Event Continuity*.

(ii) *There exists a unique finitely additive, non-atomic probability measure* $\mu(\cdot)$ *on* $\mathscr{E}$ *and a non-constant, mixture continuous preference functional* $V(P) \equiv V(x_1, p_1; \ldots; x_m, p_m)$ *on* $\mathscr{P}_0(\mathscr{X})$ *which exhibits monotonicity with respect to stochastic dominance, such that the relation* $\succeq$ *over acts can be represented by the preference functional*

$$\mathscr{V}(f(\cdot)) \equiv V\big(x_1, \mu(f^{-1}(x_1)); \ldots; x_n, \mu(f^{-1}(x_n))\big),$$

where $\{x_1, \ldots, x_n\}$ *is the outcome set of the act* $f(\cdot)$.

5. CONDITIONAL PREFERENCES AND CONDITIONAL PROBABILITY

In addition to allowing for the explicit separation of preferences (as modeled by expected utility theory) from beliefs (as modeled by classical probability theory), Savage's characterization also allows for the separation of *conditional preferences* (modeled by expected utility) from *conditional beliefs* (modeled by conditional probabilities). By "conditional" we mean the individual's preferences and beliefs upon learning that a specific event E has occurred. Such concepts are required if we are to model an agent's *ex ante* plans, or *ex post* actions, contingent upon the arrival of new information.[29] The purpose of this section is to show that, just as in the unconditional case, the derivation and separation of conditional risk preferences and conditional probabilistic beliefs does not require that preferences satisfy the expected utility hypothesis.

5.1. *Conditional Preferences over Acts and over Probability Distributions*

As before, the difference between our characterization and Savage's is not in the modeling of probabilistic beliefs (which is identical in the two characterizations), but rather in our extension to non-expected utility preferences. Since this extension carries particular significance for conditional preferences, we begin by deriving a non-expected utility maximizer's conditional preferences over acts and over probability distributions.

[29] E.g., Kreps (1988, Ch. 10). In this paper, we restrict ourselves to the case of *ex ante* preferences/plans. See Machina (1989) for a summary of the debate regarding non-expected utility maximizers' *ex ante* versus *ex post* preferences.

This content downloaded from
154.59.124.38 on Thu, 28 Aug 2025 13:46:36 UTC
All use subject to https://about.jstor.org/terms

The concept of conditional preferences over acts is straightforward. Given an event $E \in \mathscr{E}$, let $\mathscr{A}_E$ denote the set of finite-outcome subacts over E. Then, for any finite-outcome subact $h(\cdot)$ defined over $\sim E$, define the conditional preference relation $\succeq|_{E,h(\cdot)}$ on $\mathscr{A}_E$ by

$$f(\cdot) \succeq|_{E,h(\cdot)} g(\cdot) \qquad \text{if and only if}$$

$$\begin{bmatrix} f(s) & \text{if } s \in E \\ h(s) & \text{if } s \notin E \end{bmatrix} \succeq \begin{bmatrix} g(s) & \text{if } s \in E \\ h(s) & \text{if } s \notin E \end{bmatrix}.$$

Under the Sure-Thing Principle, the conditional preference relation $\succeq|_{E,h(\cdot)}$ would be independent of the subact $h(\cdot)$ over $\sim E$. However, we have seen that the characteristic feature of non-expected utility preferences over acts is that they are generally *not* separable across mutually exclusive events, so that what might occur in the event $\sim E$ *could* matter. This dependence across mutually exclusive events can be seen in the example of induced non-expected utility preferences over delayed-resolution acts in Section 3.2. Dependence of conditional preferences is, of course, a feature of non-separability over any economic dimension, be it commodities, time periods, or events.

The derivation of a non-expected utility maximizer's conditional preference functional over probability distributions is analogous. Given a probability $\rho \in (0,1]$ and a distribution $Q = (y_1, q_1; \ldots; y_n, q_n)$, define the conditional preference functional $V(\cdot|\rho; Q)$ over $\mathscr{P}_0(\mathscr{X})$ by

$$V(P|\rho; Q) \equiv V(\rho \cdot P + (1-\rho) \cdot Q)$$

or equivalently

$$V((x_1, p_1; \ldots; x_m, p_m)|\rho; (y_1, q_1; \ldots; y_n, q_n))$$
$$\equiv V(x_1, \rho p_1; \ldots; x_m, \rho p_m; y_1, (1-\rho)q_1; \ldots; y_n, (1-\rho)q_n).$$

The expected utility hypothesis—specifically, the Independence Axiom—would imply that this conditional preference functional over distributions is independent of both the probability ρ and the alternative distribution Q. However, the characteristic feature of conditional non-expected utility preferences over distributions is that they *can* depend upon the alternative outcomes $\{y_1, \ldots, y_n\}$ and their probabilities $\{(1-\rho)q_1, \ldots, (1-\rho)q_n\}$. This non-separability is reflected in empirical phenomena such as the Allais Paradox, the "certainty effect," the "common consequence effect," etc., observed by Allais (1953, 1979), Kahneman and Tversky (1979), and others.[30]

5.2. *Characterization of Conditional Preferences and Beliefs*

Theorem 2 showed how a non-expected utility maximizer's unconditional preferences over acts can be represented by means of a subjective probability distribution $\mu(\cdot)$ and a non-expected utility preference functional $V(\cdot)$ over

[30] See footnote 8 for additional references.

This content downloaded from
154.59.124.38 on Thu, 28 Aug 2025 13:46:36 UTC
All use subject to https://about.jstor.org/terms

probability distributions. The following theorem shows that for any non-null event E and any subact $h(\cdot)$ on $\sim E$, such an individual's conditional preferences over subacts on E can be similarly represented by a subjective probability distribution over subsets of E and a preference functional over probability distributions. This new probability distribution and preference functional over distributions are derived from the original $\mu(\cdot)$ and $V(\cdot)$ in the natural manner: the new probability distribution will be the conditional distribution of $\mu(\cdot)$ given E, and the new preference functional will be the conditional preference functional $V(\cdot\,|\rho;Q)$, where ρ is the probability of E, and Q is the conditional distribution implied by the subact $h(\cdot)$ over $\sim E$.

THEOREM 3: *Assume the preference relation $\succeq$ over $\mathscr{A}$ satisfies the conditions of Theorem 2, so that $\succeq$ can be represented by a subjective probability measure $\mu(\cdot)$ on $\mathscr{E}$ and a preference functional $V(\cdot)$ over $\mathscr{P}_0(\mathscr{X})$. Let E be any event satisfying $\mu(E)\in(0,1)$, and let $h(\cdot)$ be any finite-outcome subact defined over $\sim E$ with outcome set $\{z_1,\ldots,z_k\}$. Then:*

(i): *There exists a unique finitely additive, non-atomic probability measure $\mu_E(\cdot)$ over subsets of E and a non-constant, mixture continuous preference functional $V_{E,h(\cdot)}(\cdot)$ over $\mathscr{P}_0(\mathscr{X})$, monotonic with respect to stochastic dominance, such that for any subacts $f(\cdot)$ and $g(\cdot)$ in $\mathscr{A}_E$ with respective outcome sets $\{x_1,\ldots,x_r\}$ and $\{y_1,\ldots,y_s\}$*

$$f(\cdot)\succeq|_{E,h(\cdot)}\, g(\cdot) \qquad \textit{if and only if} \qquad \mathscr{V}_{E,h(\cdot)}(f(\cdot))\geqslant \mathscr{V}_{E,h(\cdot)}(g(\cdot)),$$

where

$$\mathscr{V}_{E,h(\cdot)}(f(\cdot))\equiv V_{E,h(\cdot)}\big(x_1,\mu_E\big(f^{-1}(x_1)\big);\ldots;x_r,\mu_E\big(f^{-1}(x_r)\big)\big).$$

(ii): *The subjective probability measure $\mu_E(\cdot)$ and the preference functional $V_{E,h(\cdot)}(\cdot)$ are given by*

$$\mu_E(A)\equiv\mu(A)/\mu(E) \qquad \textit{for all } A\in\mathscr{E} \textit{ such that } A\subseteq E$$

and

$$V_{E,h(\cdot)}(P)\equiv V(P|\mu(E);Q) \qquad \textit{for all } P\in\mathscr{P}_0(\mathscr{X}),$$

where $Q=(z_1,\mu(h^{-1}(z_1))/\mu(\sim E);\ldots;z_k,\mu(h^{-1}(z_k))/\mu(\sim E))$ is the conditional probability distribution implied by the subact $h(\cdot)$ conditional on the event $\sim E$ occurring.

6. CONCLUDING REMARKS

6.1. *Related Work*

Ours is not the only choice-theoretic derivation of subjective probability outside of the expected utility framework. Gilboa (1985), for example, has derived conditions for a non-additive measure of the type used in Section 3.3 to be representable as a monotonic transformation of an additive probability measure. This result can be applied to Gilboa (1987) to obtain a characteriza-

This content downloaded from
154.59.124.38 on Thu, 28 Aug 2025 13:46:36 UTC
All use subject to https://about.jstor.org/terms

tion of probabilistic sophistication which yields the class of non-expected utility models developed by Quiggin (1982), Segal (1984), and Yaari (1987).

In another line of work, Fishburn (1984; 1988, Ch.9; 1989), Fishburn and LaValle (1987a, 1987b, 1988), and Sugden (1992) have provided axioms on a preference relation $\succeq$ over acts which imply a probability measure $\mu(\cdot)$ on events and a function $\phi(\cdot,\cdot)$ over *pairs* of outcomes, satisfying $\phi(x,y) \equiv -\phi(y,x)$, such that for any acts $f(\cdot)$ and $g(\cdot)$

$$f(\cdot) \succeq g(\cdot) \qquad \text{if and only if}$$

$$\mathscr{V}(f(\cdot),g(\cdot)) \equiv \int \phi(f(s),g(s)) \cdot d\mu(s) \geqslant 0.$$

Since the functions $\mathscr{V}(\cdot,\cdot)$ and $\phi(\cdot,\cdot)$ are defined over pairs of acts or outcomes, such preferences are typically not transitive.[31] Intuitively, this is because the individual's "valuation" $\phi(x,y)$ of an outcome x can depend upon the outcome y with which it is being compared. In the special case where $\phi(x,y) \equiv U(x) - U(y)$ for some $U(\cdot)$, we have

$$f(\cdot) \succeq g(\cdot) \quad \Leftrightarrow \quad \int \phi(f(s),g(s)) \cdot d\mu(s)$$

$$= \int U(f(s)) \cdot d\mu(s) - \int U(g(s)) \cdot d\mu(s) \geqslant 0,$$

which is precisely Savage's model of probabilistically sophisticated expected utility preferences. This, however, is the only instance where this model will be transitive.[32]

Are such preferences probabilistically sophisticated? An argument due to Faruk Gul (1989) demonstrates that in general they are not. To see this, partition $\mathscr{S}$ into events $\{E_1, E_2, E_3\}$ such that $\mu(E_1) = \mu(E_2) = \mu(E_3) = 1/3$. Since for any x, y, and z, the acts

$$[x \text{ if } E_1; y \text{ if } E_2; z \text{ if } E_3] \quad \text{and} \quad [y \text{ if } E_1; z \text{ if } E_2; x \text{ if } E_3]$$

imply the same probability distribution over outcomes, probabilistic sophistication implies they must be indifferent. The definition of $\mathscr{V}(\cdot,\cdot)$ and skew symmetry of $\phi(\cdot,\cdot)$ then imply

$$\tfrac{1}{3} \cdot \phi(x,y) + \tfrac{1}{3} \cdot \phi(y,z) \equiv \tfrac{1}{3} \cdot \phi(x,z)$$

for any x, y, and z. Defining $U(x) \equiv \phi(x, y_0)$ for some fixed y_0, it follows that for any x and z

$$\phi(x,z) \equiv \phi(x,y_0) + \phi(y_0,z) \equiv \phi(x,y_0) - \phi(z,y_0) \equiv U(x) - U(z).$$

Accordingly, the only form of this model that satisfies our own definition of

[31] If $\phi(x,y)$, $\phi(y,z)$, and $\phi(z,x)$ are all positive, for example, we have $x \succ y$, $y \succ z$, and $z \succ x$.
[32] Adding the condition of transitivity to Fishburn's (1989) axiomatization of this model yields the Savage axioms, and hence expected utility.

This content downloaded from
154.59.124.38 on Thu, 28 Aug 2025 13:46:36 UTC
All use subject to https://about.jstor.org/terms

probabilistic sophistication is the expected utility form

$$\mathscr{V}(f(\cdot), g(\cdot)) \equiv \int [U(f(s)) - U(g(s))] \cdot d\mu(s).$$

On the other hand, since for any acts $f(\cdot)$ and $g(\cdot)$ with outcome sets $\{x_1, \ldots, x_m\}$ and $\{y_1, \ldots, y_n\}$ we can write

$$\mathscr{V}(f(\cdot), g(\cdot)) \equiv \int \phi(f(s), g(s)) \cdot d\mu(s)$$
$$\equiv \sum_{i=1}^{m} \sum_{j=1}^{n} \phi(x_i, y_j) \cdot \mu\left(f^{-1}(x_i) \cap g^{-1}(y_j)\right),$$

it follows that $f(\cdot)$ and $g(\cdot)$ enter into $\mathscr{V}(\cdot, \cdot)$ only through their outcomes x_i, y_j and corresponding *joint* probabilities $\mu(f^{-1}(x_i) \cap g^{-1}(y_j))$. This may be thought of as a weaker form of "probabilistic sophistication" in a world where individuals do not value outcomes or acts individually, but rather in comparison with what they would have received had they made a different choice.

6.2. *Extension to Countable Additivity*

One of the features that our characterization inherits from Savage is that, while it assigns a subjective probability to every set of states, the subjective probability measure $\mu(\cdot)$ is in general only finitely rather than countably additive.[33] In a similar vein, Ramsey (1931) only established finite additivity for his "degrees of belief," and Anscombe and Aumann (1963) worked with only a finite number of subjectively uncertain states. But since many of the fundamental results of probability theory require the assumption of *countable* additivity, it is worthwhile sketching how our characterization could be adapted to obtain a countably additive subjective probability measure.

As noted by Savage (1954, pp. 40–43), the key issue is the fact that, even for state spaces as simple as the unit interval, a probability measure cannot be simultaneously nonatomic, countably additive, and defined over every subset of states. Accordingly, to characterize countably additive subjective probability, we must restrict the domain of events to some sufficiently small σ-algebra[34] $\mathscr{E}' \subseteq \mathscr{E}$ of subsets of $\mathscr{S}$, and restrict the domain of acts to the set $\mathscr{A}'$ of finite-outcome, $\mathscr{E}'$-measurable[35] functions from $\mathscr{S}$ to $\mathscr{X}$.

Derivations of countably additive subjective probability measures over σ-algebras of events have been given by Villegas (1964) and Malmnäs (1981, Ch. IV) for qualitative probability and by Fishburn (1982, pp. 126–134) for expected utility preferences on acts involving mixed subjective/objective uncertainty. The

[33] A probability measure $\mu(\cdot)$ is said to be *countably additive* if $\mu(\cup_{i=1}^{\infty} E_i) = \sum_{i=1}^{\infty} \mu(E_i)$ for any countable set of disjoint events $\{E_1, E_2, \ldots\}$.

[34] A collection of sets is said to be a ***σ-algebra*** if it contains the empty set and is closed under complementation and countable unions (and hence also countable intersections).

[35] A finite-outcome act $f(\cdot)$ is said to be ***$\mathscr{E}'$-measurable*** if $f^{-1}(x) \in \mathscr{E}'$ for all $x \in \mathscr{X}$.

This content downloaded from
154.59.124.38 on Thu, 28 Aug 2025 13:46:36 UTC
All use subject to https://about.jstor.org/terms

derivation closest to Savage's is that of Arrow (1965, Lect. 1; 1970, Ch. 2), who obtains countable additivity by means of the following axiom (1970, p. 48):

AXIOM (Monotone Continuity): *Given any acts* $f(\cdot) \succ g(\cdot)$, *outcome* x, *and sequence of events* $\{E_1, E_2, \ldots\}$ *such that* $E_{i+1} \subset E_i$ *for all* i *and* $\bigcap_{i=1}^{\infty} E_i$ *is empty, there exists an* i^* *such that for all* i *greater than* i^*

$$\begin{bmatrix} x & \text{if } s \in E_i \\ f(s) & \text{if } s \in \sim E_i \end{bmatrix} \succ g(\cdot) \qquad \textit{and} \qquad f(\cdot) \succ \begin{bmatrix} x & \text{if } s \in E_i \\ g(s) & \text{if } s \in \sim E_i \end{bmatrix}.$$

By adapting our framework and axioms to those of Arrow, dropping his version of the Sure-Thing Principle,[36] and adding an appropriate version of our Strong Comparative Probability Axiom, we could presumably obtain a joint derivation of countably additive subjective probability and probabilistically sophisticated non-expected utility preferences.

Dept. of Economics, University of California-San Diego, La Jolla, California 92093, U.S.A.

and

Dept. of Statistics, Tel-Aviv University, 69978 Tel-Aviv, Israel, and Dept. of Economics, The Ohio State University, Columbus, Ohio 43210, U.S.A.

Manuscript received July, 1990; final revision received December, 1991.

APPENDIX

PROOFS OF THEOREMS

PROOF OF THEOREM 1: Consider a setting $\{\mathscr{S}, \mathscr{E}, \mathscr{X}, \mathscr{A}, \succeq\}$ satisfying Axioms P1, P3, P4*, P5, and P6. Our proof consists of five steps. In Step 1 we obtain a probability measure $\mu(\cdot)$ on $\mathscr{E}$. In Step 2 we show that an event E is non-null if and only if $\mu(E) > 0$. In Step 3 we show that $\mu(\cdot)$ represents the comparative likelihood relation $\succeq_l$ from Axiom P4*. In Step 4 we show that exchanging the outcomes on any pair of equally likely events leaves the individual indifferent. Finally, in Step 5 we show that if $f(\cdot)$ and $g(\cdot)$ imply the same probability distribution over outcomes, they are indifferent.

Step 1 (Derivation of $\mu(\cdot)$): By P5 there exist outcomes $z^* \succ z$ in $\mathscr{X}$. Consider the setting $\{\mathscr{S}, \mathscr{E}, \mathscr{X}^*, \mathscr{A}^*, \succeq^*\}$, where $\mathscr{X}^* = \{z^*, z\}$, $\mathscr{A}^* \subseteq \mathscr{A}$ is the set of acts with outcomes in $\{z^*, z\}$, and $\succeq^*$ is the restriction of $\succeq$ to $\mathscr{A}^*$. It is straightforward to verify that P1, P3, P4, P5, and P6 hold over $\{\mathscr{S}, \mathscr{E}, \mathscr{X}^*, \mathscr{A}^*, \succeq^*\}$. From Section 4.2, we have that P2 also holds in this two-outcome setting. From Savage's Theorem, there accordingly exists a unique, finitely additive, non-atomic probability measure $\mu(\cdot)$ on $\mathscr{E}$ such that for any $f(\cdot)$ and $g(\cdot)$ in $\mathscr{A}^*$, $f(\cdot) \succeq^* g(\cdot)$ (and hence $f(\cdot) \succeq g(\cdot)$) if and only if $\mu(f^{-1}(z^*)) \geqslant \mu(g^{-1}(z^*))$. From Savage (1954, p. 37, Concl. 7) it also follows that for any $P \in \mathscr{P}_0(\mathscr{X})$, there exists an act $f(\cdot) \in \mathscr{A}$ whose implied probability distribution over outcomes is P.

Step 2 (E is non-null if and only if $\mu(E) > 0$): The second-last sentence of Step 1 implies that $\mu(E) > 0$ if and only if $[z^* \text{ on } E;\ z \text{ on } \sim E] \succ [z \text{ on } E;\ z \text{ on } \sim E]$, which by P3 is equivalent to the condition that E is non-null with respect to $\succeq$ over the entire set $\mathscr{A}$.

[36] Arrow's assumption of "Conditional Preference" (1970, p. 49).

This content downloaded from
154.59.124.38 on Thu, 28 Aug 2025 13:46:36 UTC
All use subject to https://about.jstor.org/terms

Step 3 ($\mu(\cdot)$ represents $\succeq_l$): Let A and B be a pair of arbitrary events. It follows from P4*, the definition of $\succeq_l$, and Step 1 that

$$A \succeq_l B \quad \Leftrightarrow \quad \begin{bmatrix} z^* & \text{if } s \in A - B \\ z & \text{if } s \in B - A \\ z & \text{if } s \notin A \cup B \end{bmatrix} \succeq \begin{bmatrix} z & \text{if } s \in A - B \\ z^* & \text{if } s \in B - A \\ z & \text{if } s \notin A \cup B \end{bmatrix}$$
$$\Leftrightarrow \quad \mu(A - B) \geqslant \mu(B - A),$$

which, since $\mu(\cdot)$ is a probability measure, is equivalent to the condition $\mu(A) \geqslant \mu(B)$.

Step 4 (Exchanging outcomes on equally likely events leaves the individual indifferent): Let A and B be a pair of disjoint events satisfying $\mu(A) = \mu(B)$, and consider the acts

$$f(\cdot) = \begin{bmatrix} x & \text{if } s \in A \\ y & \text{if } s \in B \\ h(s) & \text{if } s \notin A \cup B \end{bmatrix} \quad \text{and} \quad g(\cdot) = \begin{bmatrix} y & \text{if } s \in A \\ x & \text{if } s \in B \\ h(s) & \text{if } s \notin A \cup B \end{bmatrix}.$$

From Step 3 we have $A \succeq_l B$ and $B \succeq_l A$ which, by definition of $\succeq_l$, implies $f(\cdot) \sim g(\cdot)$.

Step 5 (Acts $f(\cdot)$ and $g(\cdot)$ that imply the same probability distributions over outcomes are indifferent): Without loss of generality, we can assume that $f(\cdot)$ and $g(\cdot)$ assign strictly positive probability to each element of their common outcome set $\{x_1, \ldots, x_n\}$.[37] Define $G_i = g^{-1}(x_i)$ for $i = 1, \ldots, n$. We will construct a sequence of acts $f(\cdot) = f_0(\cdot), f_1(\cdot), f_2(\cdot) \cdots f_{n-1}(\cdot)$, such that $f_i(\cdot)$ is obtained from $f_{i-1}(\cdot)$ by exchanging outcomes on equally likely events, so that $f_i(\cdot)$ is indifferent to $f_{i-1}(\cdot)$, and such that

$$\begin{array}{lllll} f_1^{-1}(x_1) = G_1, & & & & \\ f_2^{-1}(x_1) = G_1, & f_2^{-1}(x_2) = G_2, & & & \\ f_3^{-1}(x_1) = G_1, & f_3^{-1}(x_2) = G_2, & f_3^{-1}(x_3) = G_3, & & \\ & \vdots & & & \\ f_{n-1}^{-1}(x_1) = G_1, & f_{n-1}^{-1}(x_2) = G_2, & f_{n-1}^{-1}(x_3) = G_3, & \cdots & f_{n-1}^{-1}(x_{n-1}) = G_{n-1}, \end{array}$$

which implies that $f_{n-1}^{-1}(x_n)$ must also equal G_n, so that $f_{n-1}(\cdot) = g(\cdot)$.

Construction of $f_i(\cdot)$ *from* $f_{i-1}(\cdot)$ *for* $i = 1, \ldots, n-1$:[38] We have from the construction that: (i) $f_{i-1}(\cdot)$ is indifferent to $f(\cdot)$; (ii) $f_{i-1}(\cdot)$ implies the same outcome probabilities as $f(\cdot)$, and hence as $g(\cdot)$; [(iii) $f_{i-1}(\cdot)$ satisfies $f_{i-1}^{-1}(x_1) = G_1, \ldots, f_{i-1}^{-1}(x_{i-1}) = G_{i-1}$]. Since Property (ii) above implies $\mu(f_{i-1}^{-1}(x_i)) = \mu(G_i)$, we have

$$\mu\left(G_i - f_{i-1}^{-1}(x_i)\right) = \mu\left(f_{i-1}^{-1}(x_i) - G_i\right).$$

[Since Property (iii) implies that $f_{i-1}(\cdot)$ does not take any of the values $x_1, x_2, \ldots, x_{i-1}$ on the sets $G_i, \ldots, G_n$,] we have that $\{G_i \cap f_{i-1}^{-1}(x_i), G_i \cap f_{i-1}^{-1}(x_{i+1}), \ldots, G_i \cap f_{i-1}^{-1}(x_n)\}$ is a partition of G_i, so that the above equality may be rewritten as

$$\mu\left(G_i \cap f_{i-1}^{-1}(x_{i+1})\right) + \mu\left(G_i \cap f_{i-1}^{-1}(x_{i+2})\right) + \cdots + \mu\left(G_i \cap f_{i-1}^{-1}(x_n)\right)$$
$$= \mu\left(f_{i-1}^{-1}(x_i) - G_i\right),$$

in other words, the total probability that $f_{i-1}(\cdot)$ assigns to the outcomes $x_{i+1}, \ldots, x_n$ *on* the set G_i equals the probability that $f_{i-1}(\cdot)$ assigns to the outcome x_i *off of* the set G_i (that is, to the set $f_{i-1}^{-1}(x_i) - G_i$). By the last sentence of Step 1, we can partition $f_{i-1}^{-1}(x_i) - G_i$ into sets $\{E_{i+1}^i, E_{i+2}^i, \ldots, E_n^i\}$ such that

$$\mu\left(G_i \cap f_{i-1}^{-1}(x_{i+1})\right) = \mu\left(E_{i+1}^i\right),$$
$$\mu\left(G_i \cap f_{i-1}^{-1}(x_{i+2})\right) = \mu\left(E_{i+2}^i\right),$$
$$\vdots$$
$$\mu\left(G_i \cap f_{i-1}^{-1}(x_n)\right) = \mu\left(E_n^i\right).$$

[37] By Step 2, we can achieve this condition by replacing each zero probability outcome in $f(\cdot)$ by one of its positive probability outcomes, and similarly for $g(\cdot)$.

[38] Statements enclosed in square brackets do not hold or are vacuous for the case $i = 1$.

This content downloaded from
154.59.124.38 on Thu, 28 Aug 2025 13:46:36 UTC
All use subject to https://about.jstor.org/terms

Since $f_{i-1}(\cdot)$ takes on the value x_{i+1} over the set $G_i \cap f_{i-1}^{-1}(x_{i+1})$ and the value x_i over the set E_{i+1}^i, and since these sets are equally probable, it follows from Step 4 that exchanging the outcomes on these sets will yield an act which is indifferent to $f_{i-1}(\cdot)$ and which implies the same probabilities of each outcome as $f_{i-1}(\cdot)$, and hence as $f(\cdot)$. Exchanging the outcomes x_{i+2} and x_i on the sets $G_i \cap f_{i-1}^{-1}(x_{i+2})$ and E_{i+2}^i, and so on, up to the outcomes x_n and x_i on the sets $G_i \cap f_{i-1}^{-1}(x_n)$ and E_n^i, yields the act $f_i(\cdot)$ defined by

$$f_i(s) = \begin{cases} x_1 & \text{for } s \in G_1, \\ [x_2 & \text{for } s \in G_2], \\ \vdots & \\ [x_{i-1} & \text{for } s \in G_{i-1}], \\ x_i & \text{for } s \in G_i, \\ x_{i+1} & \text{for } s \in E_{i+1}^i, \\ \vdots & \\ x_n & \text{for } s \in E_n^i, \\ f_{i-1}(s) & \text{for } s \notin \left(f^{-1}(x_1) \cup \cdots \cup f^{-1}(x_i)\right), \end{cases}$$

where (i) $f_i(\cdot)$ is indifferent to $f_{i-1}(\cdot)$, and hence indifferent to $f(\cdot)$; (ii) $f_i(\cdot)$ implies the the same outcome probabilities as $f(\cdot)$, and hence as $g(\cdot)$; (iii) $f_i(\cdot)$ satisfies $f_i^{-1}(x_1) = G_1$, $[f_i^{-1}(x_2) = G_2, \ldots,]$ $f_i^{-1}(x_i) = G_i$.

Repeating this construction for $i = 1, \ldots, n-1$ yields an act $f_{n-1}(\cdot) \sim f(\cdot)$ which satisfies

$$f_{n-1}^{-1}(x_1) = G_1, \quad f_{n-1}^{-1}(x_2) = G_2 \quad \ldots \quad f_{n-1}^{-1}(x_{n-1}) = G_{n-1}.$$

Since this implies $f_{n-1}^{-1}(x_n) = G_n$, $f_{n-1}(\cdot)$ equals the act $g(\cdot)$, so that we have $f(\cdot) \sim g(\cdot)$. *Q.E.D.*

PROOF OF THEOREM 2: (i) $\to$ (ii): Our proof consists of 6 steps. In Step 1 we define a preference relation $\succeq_p$ over probability distributions induced by the relation $\succeq$ over acts. In Step 2 we show that $\succeq_p$ is monotonic with respect to stochastic dominance. In Step 3 we show that $\succeq_p$ satisfies mixture continuity. (Our definitions of monotonicity with respect to stochastic dominance and mixture continuity for a preference relation are given in footnotes 17 and 18.) In Steps 4 and 5 we construct a preference functional $V(\cdot)$ over probability distributions that represents $\succeq_p$ and exhibits the required properties. In Step 6 we show that the preference functional $\mathscr{V}(\cdot)$ over acts represents the relation $\succeq$ on $\mathscr{A}$.

Step 1 (Definition of $\succeq_p$): Let $\mu(\cdot)$ be the probability measure derived in Theorem 1. For each distribution $P = (x_1, p_1; \ldots; x_m, p_m)$ in $\mathscr{P}_0(\mathscr{X})$, let $\psi(P) = \psi(x_1, p_1; \ldots; x_m, p_m)$ be the set of acts $f(\cdot)$ which satisfy $\mu(f^{-1}(x_i)) = p_i$ for $i = 1, \ldots, m$.[39] By Theorem 1, all acts in the set $\psi(x_1, p_1; \ldots; x_m, p_m)$ are indifferent. Accordingly, the preference relation $\succeq$ over acts in $\mathscr{A}$ induces a complete, reflexive, and transitive preference relation $\succeq_p$ over distributions in $\mathscr{P}_0(\mathscr{X})$, where $P^* \succeq_p P$ if and only if $f^*(\cdot) \succeq f(\cdot)$ for all $f^*(\cdot) \in \psi(P^*)$ and $f(\cdot) \in \psi(P)$. We define the relations $\succ_p$ and $\sim_p$ in the usual manner.

Step 2 ($\succeq_p$ is monotonic with respect to stochastic dominance): Let P and Q be probability distributions in $\mathscr{P}_0(\mathscr{X})$ such that P first order stochastically dominates Q with respect to $\succeq_x$. We can write the combined support of P and Q as $\{\{z_{ij}\}_{j=1}^{k_i}\}_{i=1}^n$, where $z_{i1} \sim z_{ij}$ for all i and j, and where $z_{11} \prec z_{21} \cdots \prec z_{n1}$. Given this, we can write $P = (\cdots; z_{ij}, p_{ij}; \cdots)$ and $Q = (\cdots; z_{ij}, q_{ij}; \cdots)$, where some p_{ij} and q_{ij} may be zero. Stochastic dominance implies

$$\sum_{i=1}^{m} \sum_{j=1}^{k_i} p_{ij} \leqslant \sum_{i=1}^{m} \sum_{j=1}^{k_i} q_{ij} \qquad \text{for} \quad m = 1, \ldots, n.$$

From the Proof of Theorem 1, there exists an act $f(\cdot)$ which implies the distribution P. By Axiom P3, we can successively replace each of its outcomes z_{ij} by z_{i1} to obtain the act $\hat{f}(\cdot) = [z_{11} \text{ on } F_1; z_{21} \text{ on } F_2; \ldots; z_{n1} \text{ on } F_n] \sim f(\cdot)$, where $\mu(F_i) = \sum_{j=1}^{k_i} p_{ij}$ for $i = 1, \ldots, n$.

[39] Note that an act $f(\cdot)$ in $\psi(x_1, p_1; \ldots; x_m, p_m)$ could have outcomes outside of $\{x_1, \ldots, x_m\}$, provided they only occur on zero probability sets.

This content downloaded from
154.59.124.38 on Thu, 28 Aug 2025 13:46:36 UTC
All use subject to https://about.jstor.org/terms

By nonatomicity of $\mu(\cdot)$ and the above displayed inequality, we can construct a partition $\{G_1,\ldots,G_n\}$ of $\mathscr{S}$ such that $\mu(G_i)=\sum_{j=1}^{k_i}q_{ij}$ and $\bigcup_{i=1}^{m}F_i\subseteq\bigcup_{i=1}^{m}G_i$ for $m=1,\ldots,n$. Define the act $\hat{g}(\cdot)=[z_{11}$ on G_1; z_{21} on $G_2;\ldots;z_{n1}$ on $G_n]$.

Since $\hat{f}^{-1}(z_{i1})=F_i\subseteq\bigcup_{j=1}^{i}G_j=\hat{g}^{-1}(\{z_{11},z_{21},\ldots,z_{i1}\})$ for each i, it follows that wherever $\hat{f}(\cdot)$ yields the outcome z_{i1}, $\hat{g}(\cdot)$ either yields z_{i1} or an outcome strictly less preferred than z_{i1}, which by repeated application of Axiom P3 implies that $\hat{f}(\cdot)\succeq\hat{g}(\cdot)$.

By nonatomicity, we can partition each set G_i into sets $\{G_{i1},\ldots,G_{ik_i}\}$ such that $\mu(G_{ij})=q_{ij}$ for $j=1,\ldots,k_i$. By Axiom P3, we can take the act $\hat{g}(\cdot)$ and replace z_{i1} by z_{ij} over each set G_{ij}, to obtain the act $g(\cdot)=[\ldots;z_{ij}$ on $G_{ij};\ldots]\sim\hat{g}(\cdot)\preceq\hat{f}(\cdot)\sim f(\cdot)$. Since $g(\cdot)$ implies the distribution $(\ldots;z_{ij},q_{ij};\ldots)=Q$, we have that $P\succeq_{\not p}Q$.

If P *strictly* stochastically dominates Q, we have $\sum_{i=1}^{r}\sum_{j=1}^{k_i}p_{ij}<\sum_{i=1}^{r}\sum_{j=1}^{k_i}q_{ij}$ for some $r\in\{1,\ldots,n-1\}$. This implies that the set $(\bigcup_{j=1}^{r}G_j)-(\bigcup_{j=1}^{r}F_j)=(\bigcup_{j=1}^{r}G_j)\cap(\bigcup_{j=r+1}^{n}F_j)$ has positive probability, and hence that there exists a nonnull set on which $\hat{f}(\cdot)$ yields an outcome strictly preferred to z_{r1} but where $\hat{g}(\cdot)$ yields z_{r1} or an outcome strictly less preferred than z_{r1}. By Axiom P3 this implies $\hat{f}(\cdot)\succ\hat{g}(\cdot)$, so that $f(\cdot)\succ g(\cdot)$ and hence $P\succ_{\not p}Q$.

Step 3 ($\succeq_{\not p}$ is mixture continuous): Let P, Q, and R be arbitrary probability distributions in $\mathscr{P}_0(\mathscr{X})$. Without loss of generality, we can write $P=(z_1,p_1;\ldots;z_n,p_n)$ and $Q=(z_1,q_1;\ldots;z_n,q_n)$ where $z_1\preceq\cdots\preceq z_n$. We will demonstrate that the sets $\Lambda_{\succ}=\{\lambda\in[0,1]\,|\,\lambda\cdot P+(1-\lambda)\cdot Q\succ_{\not p}R\}$ and $\Lambda_{\prec}=\{\lambda\in[0,1]\,|\,\lambda\cdot P+(1-\lambda)\cdot Q\prec_{\not p}R\}$ are both open in $[0,1]$.

Let λ^* be an arbitrary element of $\Lambda_{\succ}$. We will show that there exists some $\varepsilon>0$ such that $\lambda\in(\lambda^*-\varepsilon,\lambda^*+\varepsilon)\cap[0,1]$ implies $\lambda\cdot P+(1-\lambda)\cdot Q\succ_{\not p}R$. By the proof of Theorem 1, there exist acts $f(\cdot)$ and $g(\cdot)$ in $\mathscr{A}$ which imply the distributions $\lambda^*\cdot P+(1-\lambda^*)\cdot Q$ and R respectively. Since $f(\cdot)\succ g(\cdot)$, Axiom P6 implies there exists a partition $\{A_{2,1},\ldots,A_{2,n_2}\}$ of $\mathscr{S}$ such that $[z_1$ if $s\in A_{2,i}$; $f(s)$ if $s\notin A_{2,i}]\succ g(\cdot)$ for each $A_{2,i}\in\{A_{2,1},\ldots,A_{2,n_2}\}$. If $\mu(f^{-1}(z_2))>0$, $\{A_{2,1},\ldots,A_{2,n_2}\}$ must contain some set A_2^* whose intersection with $f^{-1}(z_2)$ has positive measure. If $\mu(f^{-1}(z_2))=0$, let A_2^* be any set in $\{A_{2,1},\ldots,A_{2,n_2}\}$. We thus have $[z_1$ if $s\in A_2^*$; $f(s)$ if $s\notin A_2^*]\succ g(\cdot)$, where either $\mu(A_2^*\cap f^{-1}(z_2))>0$ or else $\mu(f^{-1}(z_2))=0$.

Applying P6 to the preceding preference ranking, there exists a partition $\{A_{3,1},\ldots,A_{3,n_3}\}$ such that $[z_1$ if $s\in A_2^*\cup A_{3,i}$; $f(s)$ if $s\notin A_2^*\cup A_{3,i}]\succ g(\cdot)$ for each $A_{3,i}\in\{A_{3,1},\ldots,A_{3,n_3}\}$. Again, $\{A_{3,1},\ldots,A_{3,n_3}\}$ must contain some set A_3^* such that $[z_1$ if $s\in A_2^*\cup A_3^*$; $f(s)$ if $s\notin A_2^*\cup A_3^*]\succ g(\cdot)$, where either $\mu(A_3^*\cap f^{-1}(z_3))>0$ or else $\mu(f^{-1}(z_3))=0$. Proceeding similarly yields sets $\{A_2^*,\ldots,A_n^*\}$ such that $h(\cdot)=[z_1$ if $s\in A_2^*\cup\cdots\cup A_n^*$; $f(s)$ if $s\notin A_2^*\cup\cdots\cup A_n^*]\succ g(\cdot)$, and for each $i=2,\ldots n$, either $\mu(A_i^*\cap f^{-1}(z_i))>0$ or else $\mu(f^{-1}(z_i))=0$. The act $h(\cdot)$ implies the probability distribution $H=(z_1,\rho_1;\ldots;z_n,\rho_n)$ with $H\succ_{\not p}R$ and where for each $i=1,\ldots,n$, either (a) $\rho_i\leqslant\mu(f^{-1}(z_i)-A_i^*)<\mu(f^{-1}(z_i))=\lambda^*\cdot p_i+(1-\lambda^*)\cdot q_i$, or else (b) $\rho_i=\mu(f^{-1}(z_i))=0$. Pick $\varepsilon>0$ so that it is strictly less than $(\lambda^*\cdot p_i+(1-\lambda^*)\cdot q_i)-\rho_i$ for all $i=2,\ldots,n$ that satisfy condition (a).

Let λ be an arbitrary element of $(\lambda^*-\varepsilon,\lambda^*+\varepsilon)\cap[0,1]$ and consider the distribution $\lambda\cdot P+(1-\lambda)\cdot Q=(z_1,(\lambda\cdot p_1+(1-\lambda)\cdot q_1);\ldots;z_n,(\lambda\cdot p_n+(1-\lambda)\cdot q_n))$. For each index $i=2,\ldots,n$ satisfying condition (a) we have

$$\begin{aligned}\lambda\cdot p_i+(1-\lambda)\cdot q_i&\geqslant\lambda^*\cdot p_i+(1-\lambda^*)\cdot q_i-\varepsilon\cdot|p_i-q_i|\\&\geqslant\lambda^*\cdot p_i+(1-\lambda^*)\cdot q_i-\varepsilon>\rho_i,\end{aligned}$$

and for each $i=2,\ldots,n$ satisfying condition (b) we have

$$\lambda\cdot p_i+(1-\lambda)\cdot q_i\geqslant 0=\rho_i,$$

which implies that the distribution $\lambda\cdot P+(1-\lambda)\cdot Q$ first order stochastically dominates the distribution H and hence that $\lambda\cdot P+(1-\lambda)\cdot Q\succ_{\not p}R$, so $\lambda\in\Lambda_{\succ}$. This implies that the set $\Lambda_{\succ}$ is open in $[0,1]$. A similar argument establishes that $\Lambda_{\prec}$ is open in $[0,1]$.

Step 4 (Construction of $V(\cdot)$ on a set of the form $\{P\in\mathscr{P}_0(\mathscr{X})\,|\,\delta_{\bar{x}}\succeq_{\not p}P\succeq_{\not p}\delta_{\underline{x}}\}$ for some $\bar{x}\succ_x\underline{x}$): (Henceforth we shall use the symbol δ_x for the degenerate probability distribution $(x,1)$.) For any P such that $\delta_{\bar{x}}\succeq_{\not p}P\succeq_{\not p}\delta_{\underline{x}}$, the fact that $\succeq_{\not p}$ satisfies mixture continuity and monotonicity with respect to stochastic dominance implies there exists a unique $\lambda_P\in[0,1]$ such that $P\sim_{\not p}\lambda_P\cdot\delta_{\bar{x}}+(1-\lambda_P)\cdot\delta_{\underline{x}}$, in which case define $V(P)=\lambda_P$ (thus $V(\delta_{\bar{x}})=1$ and $V(\delta_{\underline{x}})=0$). For any distributions P and Q such that $\delta_{\bar{x}}\succeq_{\not p}P\succeq_{\not p}(\succ_{\not p})Q\succeq_{\not p}\delta_{\underline{x}}$, we have

$$V(P)\cdot\delta_{\bar{x}}+(1-V(P))\cdot\delta_{\underline{x}}\sim_{\not p}P\succeq_{\not p}(\succ_{\not p})Q\sim_{\not p}V(Q)\cdot\delta_{\bar{x}}+(1-V(Q))\cdot\delta_{\underline{x}},$$

which by monotonicity with respect to stochastic dominance implies $V(P)\geqslant(>)V(Q)$. From this it follows that $V(P)\geqslant(>)V(Q)\Rightarrow P\succeq_{\not p}(\succ_{\not p})Q$. Thus, $V(\cdot)$ represents $\succeq_{\not p}$ over the set $\{P\in$

This content downloaded from
154.59.124.38 on Thu, 28 Aug 2025 13:46:36 UTC
All use subject to https://about.jstor.org/terms

$\mathscr{P}_0(\mathscr{X})|\delta_{\bar{x}} \succeq_{\not p} P \succeq_{\not p} \delta_{\underline{x}}\}$, and clearly inherits the properties of monotonicity with respect to stochastic dominance and mixture continuity from $\succeq_{\not p}$.

Step 5 (Construction of $V(\cdot)$ when there exist no best and no worst outcome): We will first construct a sequence of outcomes $\cdots x_{-2} \prec_x x_{-1} \prec_x x_0 \prec_x x_1 \prec_x x_2 \ldots$ such that for each $x \in \mathscr{X}$ there is a unique integer k such that $x_k \preceq_x x \prec_x x_{k+1}$. By Axiom P5, there exist outcomes $x_0 \prec_x x_1$. Define $\xi(x_0) = 0$ and $\xi(x_1) = 1$. For each outcome $x \succ_x x_1$, monotonicity with respect to stochastic dominance and mixture continuity imply there exists a unique $\lambda \in (0,1)$ such that $\delta_{x_1} \sim_{\not p} (1-\lambda) \cdot \delta_x + \lambda \cdot \delta_{x_0}$, in which case define $\xi(x) = \lambda + 1 \in (1,2)$. For each $x \prec_x x_0$, there exists a unique $\lambda \in (0,1)$ such that $\delta_{x_0} \sim_{\not p} (1-\lambda) \cdot \delta_{x_1} + \lambda \cdot \delta_x$, in which case define $\xi(x) = \lambda - 1 \in (-1,0)$. Within each of these cases and across cases, monotonicity with respect to stochastic dominance ensures that $x^* \succeq_x x$ if and only if $\xi(x^*) \geqslant \xi(x)$.

Consider the set $\xi(\mathscr{X}) = \{\xi(x)|x \subseteq \mathscr{X}\} \in (-1,2)$. Since $\xi(\mathscr{X})$ is bounded, it possesses a supremum $\xi_{\sup}$ and an infimum $\xi_{\inf}$. Since there is no best or worst outcome, $\xi(\mathscr{X})$ does not contain either $\xi_{\sup}$ or $\xi_{\inf}$. Accordingly, $\xi(\mathscr{X})$ contains an infinite sequence $1 = \xi_1 < \xi_2 < \xi_3 < \cdots$ converging to $\xi_{\sup}$ and an infinite sequence $0 = \xi_0 > \xi_{-1} > \xi_{-2} > \ldots$ converging to $\xi_{\inf}$. For all integers $i \neq 0,1$, choosing an outcome x_i such that $\xi(x_i) = \xi_i$ yields a sequence $\ldots x_{-2} \prec_x x_{-1} \prec_x x_0 \prec_x x_1 \prec_x x_2, \cdots$, with the desired property.

For any distribution $P \in \mathscr{P}_0(\mathscr{X})$, Axiom P3 implies that P is weakly less preferred than its most preferred outcome and weakly preferred to its least preferred outcome, which implies that there exists a unique integer k such that $\delta_{x_{k+1}} \succ_{\not p} P \succeq_{\not p} \delta_{x_k}$. This in turn implies that there exists a unique $\lambda \in [0,1)$ such that $P \sim_{\not p} \lambda \cdot \delta_{x_{k+1}} + (1-\lambda) \cdot \delta_{x_k}$. Define $V(P) = k + \lambda \in [k, k+1)$.

Let P and Q be distributions in $\mathscr{P}_0(\mathscr{X})$ such that $P \succeq_{\not p} Q$. If $\delta_{x_{k+1}} \succ_{\not p} P \succeq_{\not p} Q \succeq_{\not p} \delta_{x_k}$ for some k, then an argument identical to that of Step 4 implies that $V(P) \geqslant V(Q)$. If $P \succeq_{\not p} \delta_{x_k} \succ_{\not p} Q$ for some k, then by construction of $V(\cdot)$ we have $V(P) \geqslant k > V(Q)$.

Let P and Q be distributions in $\mathscr{P}_0(\mathscr{X})$ such that $V(P) \geqslant V(Q)$. If $k+1 > V(P) \geqslant V(Q) \geqslant k$ for some k, then an argument identical to that of Step 4 implies that $P \succeq_{\not p} Q$. If $V(P) \geqslant k > V(Q)$ for some k, then by construction of $V(\cdot)$ we must have $P \succeq_{\not p} \delta_{x_k} \succ_{\not p} Q$.

Thus, $V(\cdot)$ represents $\succeq_{\not p}$ over $\mathscr{P}_0(\mathscr{X})$, and again inherits the properties of monotonicity with respect to stochastic dominance and mixture continuity from $\succeq_{\not p}$. The construction and representation proofs in the cases where there is a worst outcome but not a best outcome, or a best outcome but not a worst outcome, follow easily.

Step 6 ($\mathscr{V}(\cdot)$ represents $\succeq$): Pick acts $f(\cdot), g(\cdot) \in \mathscr{A}$ with outcome sets $\{y_1, \ldots, y_r\}$ and $\{z_1, \ldots, z_s\}$. By Step 1 we have that $f(\cdot) \succeq g(\cdot)$ if and only if $(y_1, \mu(f^{-1}(y_1)); \ldots; y_r, \mu(f^{-1}(y_r))) \succeq_{\not p} (z_1, \mu(g^{-1}(z_1)); \ldots; z_s, \mu(g^{-1}(z_s)))$, which by Steps 4 and 5 is equivalent to

$$\begin{aligned}\mathscr{V}(f(\cdot)) &= V\big(y_1, \mu\big(f^{-1}(y_1)\big); \ldots; y_r, \mu\big(f^{-1}(y_r)\big)\big)\\ &\geqslant V\big(z_1, \mu\big(g^{-1}(z_1)\big); \ldots; z_s, \mu\big(g^{-1}(z_s)\big)\big) = \mathscr{V}(g(\cdot)).\end{aligned}$$

(ii) $\to$ (i): P1 (Ordering): This follows since there is a real-valued representation of $\succeq$.

P3 (Eventwise Monotonicity): If $x \succeq y$, then $[x$ if E; $h(s)$ if $\sim E] \succeq [y$ if E; $h(s)$ if $\sim E]$ follows from monotonicity with respect to stochastic dominance. Conversely, assume $f(\cdot) = [x$ if E; $h(s)$ if $\sim E] \succeq g(\cdot) = [y$ if E; $h(s)$ if $\sim E]$ for some $x, y, h(\cdot)$, and non-null E, but that $y \succ x$. Since E is non-null, it cannot have zero probability. Clearly, $f(\cdot)$ and $g(\cdot)$ imply the same probability of each outcome *other than* x or y. Since it is also clear that $\mu(f^{-1}(x)) > \mu(g^{-1}(x))$ while $\mu(f^{-1}(y)) < \mu(g^{-1}(y))$, monotonicity with respect to stochastic dominance implies that $g(\cdot) \succ f(\cdot)$, which is a contradiction.

P4* (Strong Comparative Probability): Pick disjoint events A and B, outcomes $x^* \succ x$ and $y^* \succ y$, and acts $g(\cdot)$ and $h(\cdot)$ such that $g_1(\cdot) = [x^*$ if A; x if B; $g(s)$ if $\sim (A \cup B)] \succeq g_2(\cdot) = [x$ if A; x^* if B; $g(s)$ if $\sim (A \cup B)]$. Clearly, $g_1(\cdot)$ and $g_2(\cdot)$ imply the same probability of each outcome *other than* x^* or x. Since $\mu(g_1^{-1}(x^*)) = \mu(A) + \mu(g^{-1}(x^*) - (A \cup B))$ and $\mu(g_2^{-1}(x^*)) = \mu(B) + \mu(g^{-1}(x^*) - (A \cup B))$, monotonicity with respect to stochastic dominance implies that $\mu(A) \geqslant \mu(B)$. This fact and a similar argument imply that the act $[y^*$ if A; y if B; $h(\cdot)$ if $\sim (A \cup B)]$ is weakly preferred to $[y$ if A; y^* if B; $h(\cdot)$ if $\sim (A \cup B)]$.

P5 (Nondegeneracy): Say all outcomes were indifferent. For arbitrary act $f(\cdot) = [x_1$ if E_1; x_2 if $E_2; \ldots; x_n$ if $E_n] \in \mathscr{A}$ and outcome $z \in \mathscr{X}$, repeated application of P3 yields $f(\cdot) \sim [z$ if E_1; x_2 if $E_2; \ldots; x_n$ if $E_n] \sim [z$ if E_1; z if $E_2; \ldots; x_n$ if $E_n] \sim [z$ if E_1; z if $E_2; \ldots; z$ if $E_n]$, which implies that all acts in $\mathscr{A}$ are indifferent. But this implies that all probability distributions in $\mathscr{P}_0(\mathscr{X})$ are indifferent, so $V(\cdot)$ is constant, which is a contradiction.

This content downloaded from
154.59.124.38 on Thu, 28 Aug 2025 13:46:36 UTC
All use subject to https://about.jstor.org/terms

P6 (Small Event Continuity): Pick arbitrary acts $f(\cdot) \succ g(\cdot)$ and outcome x. Let $\{z_1, \ldots, z_m\}$ be the union of the outcome sets of $f(\cdot)$ and $g(\cdot)$, let $P = (z_1, p_1; \ldots, z_m, p_m)$ and $Q = (z_1, q_1; \ldots, z_m, q_m)$ be their implied probability distributions over outcomes, and let $\bar{z}$ and $\underline{z}$ be the most and least preferred outcomes in $\{x\} \cup \{z_1, \ldots, z_m\}$. Since $\delta_{\bar{z}} \succeq_{\not p} P \succ_{\not p} Q \succeq_{\not p} \delta_{\underline{z}}$, mixture continuity and monotonicity with respect to stochastic dominance ensure there exists some sufficiently large integer n such that both $V((1-(1/n)) \cdot P + (1/n) \cdot \delta_{\underline{z}}) > V(Q)$ and $V(P) > V((1/n) \cdot \delta_{\bar{z}} + (1-(1/n)) \cdot Q)$. Since $\mu(\cdot)$ is nonatomic, we can partition each set $f^{-1}(z_i) \cap g^{-1}(z_j)$ into n equally probable events $\{E_{i,j,1}, \ldots, E_{i,j,n}\}$ (some of these sets, and hence partitions, may be empty). For $k = 1, \ldots, n$ define $A_k = \bigcup_{i=1}^m \bigcup_{j=1}^m E_{i,j,k}$, so that $\{A_1, \ldots, A_n\}$ forms a partition of $\mathscr{S}$, and so that $\mu(A_k) = (1/n)$, $\mu(f^{-1}(z_i) - A_k) = (1-(1/n)) \cdot p_i$ and $\mu(g^{-1}(z_j) - A_k) = (1-(1/n)) \cdot q_j$ for each i, j, and k.

We thus have that, for any $k = 1, \ldots, n$, the act $[\underline{z}$ if $s \in A_k$; $f(s)$ if $s \notin A_k]$ implies the probability distribution $(1/n) \cdot \delta_{\underline{z}} + (1-(1/n)) \cdot P$ which is strictly preferred to Q, which implies that $[\underline{z}$ if $s \in A_k$; $f(s)$ if $s \notin A_k] \succ g(\cdot)$. Similarly, for any $k = 1, \ldots, n$, the act $[\bar{z}$ if $s \in A_k$; $g(s)$ if $s \notin A_k]$ implies the probability distribution $(1/n) \cdot \delta_{\bar{z}} + (1-(1/n)) \cdot Q$ which is strictly less preferred than P, which implies that $f(\cdot) \succ [\bar{z}$ if $s \in A_k$; $f(s)$ if $s \notin A_k]$. *Q.E.D.*

Proof of Theorem 3: Defining $\mu_E(\cdot)$ and $V_{E,h(\cdot)}$ as in part (ii) of the Theorem, it follows immediately that $\mu_E(\cdot)$ is non-atomic and that $V_{E,h(\cdot)}(\cdot)$ is mixture continuous over $\mathscr{P}_0(\mathscr{X})$ and exhibits monotonicity with respect to stochastic dominance. Since $\mu(E) > 0$, monotonicity with respect to stochastic dominance implies $V_{E,h(\cdot)}(\delta_x) > V_{E,h(\cdot)}(\delta_y)$ for any $x \succ y$, so $V_{E,h(\cdot)}(\cdot)$ is non-constant on $\mathscr{P}_0(\mathscr{X})$.

For any subact $f(\cdot) \in \mathscr{A}_E$ with outcome set $\{x_1, \ldots, x_r\}$, we have

$$\begin{aligned}\mathscr{V}_{E,h(\cdot)}(f(\cdot)) &= V_{E,h(\cdot)}\big(x_1, \mu_E(f^{-1}(x_1)); \ldots; x_r, \mu_E(f^{-1}(x_r))\big)\\ &= V\big(x_1, \mu_E(f^{-1}(x_1)); \ldots; x_r, \mu_E(f^{-1}(x_r)) | \mu(E);\\ &\qquad z_1, \mu(h^{-1}(z_1))/\mu(\sim E); \ldots; z_k, \mu(h^{-1}(z_k))/\mu(\sim E)\big),\end{aligned}$$

which, by definition of the conditional preference functional $V(P|\rho; Q)$, equals

$$\begin{aligned}&V\big(x_1, \mu_E(f^{-1}(x_1)) \cdot \mu(E); \ldots; x_r, \mu_E(f^{-1}(x_r)) \cdot \mu(E);\\ &\qquad z_1, \mu(h^{-1}(z_1)); \ldots; z_k, \mu(h^{-1}(z_k))\big)\\ &= V\big(x_1, \mu(f^{-1}(x_1)); \ldots; x_r, \mu(f^{-1}(x_r));\\ &\qquad z_1, \mu(h^{-1}(z_1)); \ldots; z_k, \mu(h^{-1}(z_k))\big)\\ &= \mathscr{V}([f(s) \text{ for } s \in E;\ h(s) \text{ for } s \in \sim E]).\end{aligned}$$

By definition of the relation $\succeq|_{E,h(\cdot)}$, we have that $f(\cdot) \succeq|_{E,h(\cdot)} g(\cdot)$ if and only if

$$[f(s) \text{ for } s \in E\text{: } h(s) \text{ for } s \in \sim E] \succeq [g(s) \text{ for } s \in E;\ h(s) \text{ for } s \in \sim E],$$

which is equivalent to

$$\mathscr{V}([f(s) \text{ for } s \in E\text{: } h(s) \text{ for } s \in \sim E]) \succeq \mathscr{V}([g(s) \text{ for } s \in E\text{: } h(s) \text{ for } s \in \sim E]).$$

By the above, this is in turn equivalent to the condition

$$\mathscr{V}_{E,h(\cdot)}(f(\cdot)) \succeq \mathscr{V}_{E,h(\cdot)}(g(\cdot)). \qquad \textit{Q.E.D.}$$

REFERENCES

Allais, M. (1953): "Le Comportement de l'Homme Rationnel devant le Risque, Critique des Postulats et Axiomes de l'Ecole Américaine," *Econometrica*, 21, 503–546.

——— (1979): "The So-Called Allais Paradox and Rational Decisions Under Uncertainty," in Allais and Hagen (1979).

Allais, M., and O. Hagen (Eds.) (1979): *Expected Utility Hypotheses and the Allais Paradox*. Dordrecht, Holland: D. Reidel Publishing Co.

This content downloaded from
154.59.124.38 on Thu, 28 Aug 2025 13:46:36 UTC
All use subject to https://about.jstor.org/terms

ANSCOMBE, F., AND R. AUMANN (1963): "A Definition of Subjective Probability," *Annals of Mathematical Statistics*, 34, 199–205.

ARROW, K. (1965): *Aspects of the Theory of Risk Bearing*. Helsinki: Yrjö Jahnsson Säätiö.

——— (1970): *Essays in the Theory of Risk-Bearing*. Amsterdam: North-Holland Publishing Co.

——— (1974): "Optimal Insurance and Generalized Deductibles," *Scandinavian Actuarial Journal*, 1, 1–42. Reprinted in Arrow (1983–1985, Vol. 3).

——— (1983–1985): *Collected Papers of Kenneth J. Arrow, Volumes 1–6*. Cambridge, Mass.: Harvard University Press.

BECKER, S., AND F. BROWNSON (1964): "What Price Ambiguity? Or the Role of Ambiguity in Decision-Making," *Journal of Political Economy*, 72, 63–73.

BERNOULLI, D. (1738): "Specimen Theoriae Novae de Mensura Sortis," *Commentarii Academiae Scientiarum Imperialis Petropolitanae* V, 175–192. Translation: "Exposition of a New Theory on the Measurement of Risk," *Econometrica*, 22 (1954), 23–36.

CHATEAUNEUF, A. (1985): "On the Existence of a Probability Measure Compatible with a Total Preorder on a Boolean Algebra," *Journal of Mathematical Economics*, 14, 43–52.

CHATEAUNEUF, A., AND J. JAFFRAY (1984): "Archimedean Qualitative Probabilities," *Journal of Mathematical Psychology*, 28, 191–204.

CHIPMAN, J. (1960): "Stochastic Choice and Subjective Probability," in *Decisions, Values and Groups, Vol. I.*, ed. by D. Willner. New York: Pergamon Press.

CHOQUET, G. (1953–54): "Theory of Capacities," *Annales de l'Institut Fourier*, 5, 131–295.

DAVIDSON, D., AND P. SUPPES (1956): "A Finitistic Axiomatization of Utility and Subjective Probability," *Econometrica*, 24, 264–275.

DE FINETTI, B. (1937): "La Prévision: Ses Lois Logiques, ses Sources Subjectives," *Annales de l'Institute Henri Poincaré*, 7, 1–68. English translation: "Foresight: Its Logical Laws, Its Subjective Sources" in Kyburg and Smokler (1980).

——— (1949): "La 'Logica del Plausible' Secondo la Concezione di Polya," *Atti della XLII Riunione della Società Italiana per il Progresso delle Scienze*, 227–236.

DEMPSTER, A. (1968): "A Generalization of Bayesian Inference," *Journal of the Royal Statistical Society (Series B)*, 30, 205–232.

ELLSBERG, D. (1961): "Risk, Ambiguity and the Savage Axioms," *Quarterly Journal of Economics*, 75, 643–669. Reprinted in Gärdenfors and Sahlin (1988).

FERREIRA, P. (1972): "On Subjective Probability and Expected Utilities," *Annals of Mathematical Statistics*, 43, 928–933

FISHBURN, P. (1967): "Preference-Based Definitions of Subjective Probability," *Annals of Mathematical Statistics*, 38, 1605–1617.

——— (1969a): "A General Theory of Subjective Probabilities and Expected Utilities," *Annals of Mathematical Statistics*, 40, 1419–1429.

——— (1969b): "Weak Qualitative Probability on Finite Sets," *Annals of Mathematical Statistics*, 40, 2118–2126.

——— (1970): *Utility Theory for Decision Making*. New York: John Wiley and Sons.

——— (1981): "Subjective Expected Utility: A Review of Normative Theories," *Theory and Decision*, 13, 139–199.

——— (1982): *The Foundations of Expected Utility*. Dordrecht, Holland: D. Reidel Publishing Co.

——— (1983a): "Ellsberg Revisited: A New Look at Comparative Probability," *Annals of Statistics*, 11, 1047–1059.

——— (1983b): "A Generalization of Comparative Probability on Finite Sets," *Journal of Mathematical Psychology*, 27, 298–310.

——— (1984): "SSB Utility Theory and Decision-Making Under Uncertainty," *Mathematical Social Sciences*, 8, 253–285.

——— (1986): "The Axioms of Subjective Probability," *Statistical Science*, 1, 335–345.

——— (1987): "Reconsiderations in the Foundations of Decision Under Uncertainty," *Economic Journal*, 97, 825–841.

——— (1988): *Nonlinear Preference and Utility Theory*. Baltimore: Johns Hopkins Univ. Press.

——— (1989): "Nontransitive Measurable Utility for Decision Under Uncertainty," *Journal of Mathematical Economics*, 18, 187–207.

FISHBURN, P., AND I. LAVALLE (1987a): "A Nonlinear, Nontransitive and Additive-Probability Model for Decisions Under Uncertainty," *Annals of Statistics*, 15, 830–844.

——— (1987b): "State-Dependent SSB Utility," *Economics Letters*, 25, 21–25.

This content downloaded from
154.59.124.38 on Thu, 28 Aug 2025 13:46:36 UTC
All use subject to https://about.jstor.org/terms

——— (1988): "Transitivity is Equivalent to Independence for States—Additive SSB Utilities," *Journal of Economic Theory*, 44, 202–208.

FISHBURN, P., AND R. VICKSON (1978): "Theoretical Foundations of Stochastic Dominance," in *Stochastic Dominance: An Approach to Decision Making Under Risk*, ed. by G. Whitmore and M. Findlay. Lexington, Mass.: Heath.

GÄRDENFORS, P., AND N.-E. SAHLIN (EDS.) (1988): *Decision, Probability, and Utility: Selected Readings*. Cambridge: Cambridge Univ. Press.

GILBOA, I. (1985): "Subjective Distortions of Probabilities and Non-Additive Probabilities," Working Paper No. 18–85, Foerder Institute for Economic Research, Tel-Aviv University.

——— (1987): "Expected Utility with Purely Subjective Non-Additive Probabilities," *Journal of Mathematical Economics*, 16, 65–88.

GRANT, S. (1991): "Subjective Probability Without Monotonicity," manuscript, Dept. of Economics, Harvard University.

GUL, F. (1989): Private correspondence.

HACKING, I. (1967): "Slightly More Realistic Personal Probability," *Philosophy of Science*, 34, 311–325. Reprinted in Gärdenfors and Sahlin (1988).

HAZEN, G. (1987): "Subjectively Weighted Linear Utility," *Theory and Decision*, 23, 261–282.

HERSTEIN, I., AND J. MILNOR (1953): "An Axiomatic Approach to Measurable Utility," *Econometrica*, 21, 291–297.

HEY, J., AND P. LAMBERT (EDS.) (1987): *Surveys in the Economics of Uncertainty*. Oxford: Basil Blackwell, Ltd.

KADANE, J., AND R. WINKLER (1988): "Separating Probability Elicitation from Utilities," *Journal of the American Statistical Association*, 83, 357–363.

KAHNEMAN, D., AND A. TVERSKY (1979): "Prospect Theory: An Analysis of Decision Under Risk," *Econometrica*, 47, 263–291. Reprinted in Gärdenfors and Sahlin (1988).

KARNI, E. (1991): "A Definition of Subjective Probabilities with State-Dependent Preferences," manuscript, Dept. of Economics, Johns Hopkins University.

KARNI, E., AND D. SCHMEIDLER (1991): "Utility Theory with Uncertainty," in *Handbook of Mathematical Economics, Vol. 4.*, ed. by W. Hildenbrand and H. Sonnenschein. Amsterdam: North-Holland Publishing Co.

——— (1992): "On The Uniqueness of Subjective Probability," *Economic Theory*, forthcoming.

KARNI, E., D. SCHMEIDLER, AND K. VIND (1983): "On State Dependent Preferences and Subjective Probabilities," *Econometrica*, 51, 1021–1032.

KELLER, R. (1985): "The Effects of Problem Representation on the Sure-Thing and Substitution Principles," *Management Science*, 31, 738–751.

KOOPMAN, B. (1940a): "The Bases of Probability," *Bulletin of the American Mathematical Society*, 46, 763–774. Reprinted in Kyburg and Smokler (1980).

——— (1940b): "The Axioms and Algebra of Intuitive Probability," *Annals of Mathematics*, 41, 269–292.

——— (1941): "Intuitive Probability and Sequences," *Annals of Mathematics*, 42, 169–187.

KRAFT, C., J. PRATT, AND A. SEIDENBERG (1959): "Intuitive Probability on Finite Sets," *Annals of Mathematical Statistics*, 30, 408–419.

KREPS, D. (1988): *Notes on the Theory of Choice*. Boulder, Colorado: Westview Press.

KREPS, D., AND E. PORTEUS (1979): "Temporal von Neumann-Morgenstern and Induced Preferences," *Journal of Economic Theory*, 20, 81–109.

KYBURG, H., AND H. SMOKLER (EDS.) (1980): *Studies in Subjective Probability, 2nd Ed.* Huntington, New York: Robert E. Krieger Publishing Co.

LINDLEY, D. (1982): "Scoring Rules and the Inevitability of Probability," *International Statistical Review*, 50, 1–26.

LUCE, R. D. (1967): "Sufficient Conditions for the Existence of a Finitely Additive Probability Measure," *Annals of Mathematical Statistics*, 38, 780–786.

——— (1968): "On the Numerical Representation of Qualitative Conditional Probability," *Annals of Mathematical Statistics*, 39, 481–491.

LUCE, R. D., AND P. SUPPES (1965): "Preference, Utility and Subjective Probability," in *Handbook of Mathematical Psychology, Vol. III*, ed. by R. D. Luce, R. Bush, and E. Galanter. New York: John Wiley and Sons.

MACCRIMMON, K., AND S. LARSSON (1979): "Utility Theory: Axioms Versus 'Paradoxes,'" in Allais and Hagen (1979).

MACHINA, M. (1982): "'Expected Utility' Analysis Without the Independence Axiom," *Econometrica*, 50, 277–323.

This content downloaded from
154.59.124.38 on Thu, 28 Aug 2025 13:46:36 UTC
All use subject to https://about.jstor.org/terms

Machina, M. (1983): "Generalized Expected Utility Analysis and the Nature of Observed Violations of the Independence Axiom," in *Foundations of Utility and Risk Theory with Applications*, ed. by B. Stigum and F. Wenstøp. Dordrecht, Holland: D. Reidel Publishing Co. Reprinted in Gärdenfors and Sahlin (1988).

——— (1987): "Choice Under Uncertainty: Problems Solved and Unsolved," *Journal of Economic Perspectives*, 1, 121–154.

——— (1989): "Dynamic Consistency and Non-Expected Utility Models of Choice Under Uncertainty," *Journal of Economic Literature*, 27, 1622–1668.

Malmnäs, P.-E. (1981): *From Qualitative to Quantitative Probability*. Stockholm: Almqvist & Wiksell International.

Markowitz, H. (1959): *Portfolio Selection: Efficient Diversification of Investments*. New Haven: Yale Univ. Press.

Marschak, J. (1950): "Rational Behavior, Uncertain Prospects, and Measurable Utility," *Econometrica*, 18, 111–141. ("Errata," 18, 312.)

Moskowitz, H. (1974): "Effects of Problem Representation and Feedback on Rational Behavior in Allais and Morlat-Type Problems," *Decision Sciences*, 5, 225–242.

Mossin, J. (1969): "A Note on Uncertainty and Preferences in a Temporal Context," *American Economic Review*, 59, 172–174.

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

Quiggin, J. (1982): "A Theory of Anticipated Utility," *Journal of Economic Behavior and Organization*, 3, 323–343.

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

Ramsey, F. P. (1931): "Truth and Probability," in *Foundations of Mathematics and other Logical Essays*. London: K. Paul, Trench, Trubner, and Co. Reprinted in Ramsey (1978), Kyburg and Smokler (1980), and Gärdenfors and Sahlin (1988).

——— (1978): *Foundations: Essays in Philosophy, Logic, Mathematics and Economics*. Atlantic Highlands, New Jersey: Humanities Press.

Samuelson, P. (1952): "Probability, Utility, and the Independence Axiom," *Econometrica*, 20, 670–678.

Savage, L. J. (1951): "The Theory of Statistical Decision," *Journal of the American Statistical Association*, 46, 55–67.

——— (1954): *The Foundations of Statistics*. New York: John Wiley and Sons. Revised and Enlarged Edition, New York: Dover Publications, 1972.

——— (1961): "The Foundations of Statistics Reconsidered," in *Proceedings of the Fourth Berkeley Symposium on Mathematics and Probability*. Berkeley: Univ. of California Press.

——— (1967): "Difficulties in the Theory of Personal Probability," *Philosophy of Science*, 34, 305–310.

Schmeidler, D. (1982): "Subjective Probability without Additivity," Working Paper, Foerder Institute for Economic Research, Tel Aviv University.

——— (1989): "Subjective Probability and Expected Utility without Additivity," *Econometrica*, 57, 571–587.

Segal, U. (1984): "Nonlinear Decision Weights with the Independence Axiom," manuscript, Dept. of Economics, University of California, Los Angeles.

Shafer, G. (1976): *A Mathematical Theory of Evidence*. Princeton: Princeton Univ. Press.

Shimony, A. (1967): "Amplifying Personal Probability Theory: Comments on L. J. Savage's 'Difficulties in the Theory of Personal Probability'" *Philosophy of Science*, 34, 326–332.

Slovic, P., and A. Tversky (1974): "Who Accepts Savage's Axiom?" *Behavioral Science*, 19, 368–373.

Spence, M., and R. Zeckhauser (1972): "The Effect of the Timing of Consumption Decisions and the Resolution of Lotteries on the Choice of Lotteries," *Econometrica*, 40, 401–403.

Sugden, R. (1986): "New Developments in the Theory of Choice Under Uncertainty," *Bulletin of Economic Research*, 38, 1–24. Reprinted in Hey and Lambert (1987).

——— (1992): "An Axiomatic Foundation for Regret Theory," *Journal of Economic Theory*, forthcoming.

Villegas, C. (1964): "On Qualitative Probability σ-Algebras," *Annals of Mathematical Statistics*, 35, 1787–1796.

——— (1967): "On Qualitative Probability," *American Mathematical Monthly*, 74, 661–669.

This content downloaded from
154.59.124.38 on Thu, 28 Aug 2025 13:46:36 UTC
All use subject to https://about.jstor.org/terms

von Neumann, J., and O. Morgenstern (1944): *Theory of Games and Economic Behavior*. Princeton: Princeton Univ. Press. 2nd Edition, 1947. 3rd Edition, 1953.

Wakker, P. (1981): "Agreeing Probability Measures for Comparative Probability Structures," *Annals of Statistics*, 9, 658–662.

——— (1989a): *Additive Representations of Preferences: A New Foundation of Decision Analysis*. Dordrecht, Holland: Kluwer Academic Publishers.

——— (1989b): "Continuous Subjective Expected Utility With Non-Additive Probabilities," *Journal of Mathematical Economics*, 18, 1–27.

Weber, M., and C. Camerer (1987): "Recent Developments in Modelling Preferences Under Risk," *OR Spektrum*, 9, 129–151.

Yaari, M. (1969): "Some Remarks on Measures of Risk Aversion and On Their Uses," *Journal of Economic Theory*, 1, 315–329.

——— (1987): "The Dual Theory of Choice Under Risk," *Econometrica*, 55, 95–115.

Zadeh, L. (1978): "Fuzzy Sets as a Basis for a Theory of Probability," *Fuzzy Sets and Systems*, 1, 3–28.

This content downloaded from
154.59.124.38 on Thu, 28 Aug 2025 13:46:36 UTC
All use subject to https://about.jstor.org/terms