SUBJECTIVE PROBABILITY WITHOUT MONOTONICITY:
OR HOW MACHINA'S MOM MAY ALSO BE
PROBABILISTICALLY SOPHISTICATED

BY SIMON GRANT

If an agent's preferences over subjectively uncertain acts are consistent with his or her having a subjective probability distribution over the states of nature, then those preferences can induce consistent preferences over "objectively" risky lotteries. Such "probabilistically sophisticated" behavior thus allows us to treat decision making under situations of uncertainty in an analogous manner to those under risk. This paper first characterizes exactly what probabilistic sophistication entails for an agent's beliefs about the likelihood of states of nature. Secondly, it presents characterizations of probabilistically sophisticated individuals whose induced lottery preferences obey neither the Independence Axiom (unlike Savage (1954, 1972)) nor a monotonicity property that shares some of the nature of Independence (unlike Machina and Schmeidler (1992)).

KEYWORDS: Subjective probability, uncertainty, non-expected utility, monotonicity.

1. INTRODUCTION

ONE OF THE GREAT ACHIEVEMENTS of Savage's (1972) theory of choice under uncertainty was the joint derivation of an agent's preferences for outcomes and beliefs over states of the world that conformed to the properties of a mathematical probability measure. That is, even though it may be meaningless to attach "objective" or "widely accepted" probabilities to states of the world, given a set of axioms on the agent's preferences for acts, which are mappings from states of the world to outcomes, the agent behaves as though he or she assigns utilities to outcomes and probabilities to states. Moreover, the preferences for acts can be represented by the functional which assigns to each act its expected utility using those outcome utilities and state probabilities. That is, if we map acts to lotteries over outcomes using this subjective probability measure, then the induced preferences over lotteries conform to the expected utility theory of von Neumann-Morgenstern. Hence such an agent has been labelled a subjective expected utility maximizer. Notice in particular that if two acts are mapped to the same distribution over outcomes, then they will have the same expected utility and hence must have been viewed as equally desirable by the agent.

In recent years there has been a growing body of literature characterizing lottery preferences that do not satisfy one or more of the axioms of expected utility theory. This has led Machina and Schmeidler (1992, p. 747) (hereafter

\footnote{I would like to thank Atsushi Kajii and Ben Polak for very fruitful discussions and comments on this work. I am particularly grateful for their help in correcting errors and clarifying my thoughts and expressions. This work has also benefited from the comments and suggestions of Jerry Green, Mark Machina, Andreu Mas-Colell, Maninay Sengupta, and in particular, Jurgen Eichberger, Stephen King, a co-editor, and three anonymous referees. Naturally, I remain solely responsible for any remaining errors or omissions.}
referred to as M-S (1992)) to ask:

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

The answer seems to be that from the agent's preferences over acts, one should be able to derive a probability measure over states by which the set of acts could be mapped onto the set of lotteries over outcomes. Moreover, the agent should be indifferent between any pair of acts which are mapped using this probability measure to the same lottery over outcomes. Thus the induced preferences over the lotteries in conjunction with this probability measure would fully characterize the underlying preferences over acts. Such an agent will be termed probabilistically sophisticated.

This definition of probabilistic sophistication differs from the formal one that appeared in M-S (1992). They also required that the induced lottery preferences of a probabilistically sophisticated decision maker satisfy a continuity property as well as a monotonicity property that they dubbed "monotonicity with respect to stochastic dominance." But as the discussion in the paragraph above demonstrates, there is no logical reason to associate such structural requirements on the lottery preferences with probabilistically sophisticated beliefs. Thus the first aim of this paper is to clarify M-S's (1992) concept of probabilistically sophisticated preferences and disassociate it from extraneous properties of the induced lottery preferences.

Clearly Savage's subjective expected utility maximizer is probabilistically sophisticated. M-S (1992) also characterize an agent who is probabilistically sophisticated but whose induced preferences over lotteries need not conform to expected utility theory. That is, as the preferences do not necessarily satisfy Savage's P2 (Sure-Thing Principle) the induced preferences over lotteries need not exhibit the Independence Axiom. However, since P3 (Eventwise Monotonicity) is retained, the induced preferences do obey a substitution axiom that Grant, Kajii, and Polak (1992) have dubbed the Axiom of Degenerate Independence (ADI). It requires that a movement of a probability mass from one outcome to another outcome be judged by the induced lottery preferences as an improvement if and only if the latter outcome is preferred to the former.\(^2\)

The second purpose of this paper is to axiomatize decision-makers who are probabilistically sophisticated and yet whose induced preferences over lotteries do not necessarily satisfy Independence nor ADI. The motivation for this endeavor mirrors that of M-S (1992). First, if one's goal is to achieve a separation of an agents' subjective beliefs from their risk preferences, then, as M-S (1992) argue, such a separation should be achieved with as few restrictions on risk preferences as possible. Secondly, although ADI is consistent with the

\(^2\)A simple induction argument shows that a lottery preference relation satisfies ADI if and only if it is monotonic with respect to first-order stochastic dominance. Hence by retaining P3, M-S (1992) get their required monotonicity property for the probabilistically sophisticated preferences that they characterize.
letter (and, partly, the spirit) of relaxing Independence, it too is a substitution axiom that is, in some circumstances, an inappropriate restriction on an agent's risk preferences. Machina's (1989, p. 1643) example of a Mom with an indivisible treat to allocate to one of her two children is one such instance. Thirdly, in situations where a subjective expected utility maximizer undertakes an unobservable auxiliary action between the time of choosing an act and the resolution of the uncertainty, his or her observable risk preferences may fail to satisfy not only Independence, as M-S (1992) noted, but also ADI.

After first establishing notation and the basic framework, Section 2 presents the formal definition of probabilistically sophisticated preferences employed in this paper. The connection between Savage's postulate P2 (Sure-Thing Principle) and risk preferences exhibiting Independence is noted, and it is shown that probabilistically sophisticated preferences over acts will induce risk preferences that exhibit ADI if and only if the preferences over acts satisfy Savage's postulate P3 (Eventwise Monotonicity).

Particular examples illustrating the second and third motivations for considering risk preferences that fail to exhibit ADI are developed in Section 3. Although the underlying preferences over acts, corresponding to the examples of Section 3, fail to satisfy P3, I show in Section 4 that they satisfy appropriate conditional eventwise monotonicity axioms. These conditional monotonicity properties are then exploited to illustrate how an individual's beliefs about the relative likelihood of events can be inferred from his or her preferences over acts. This leads naturally to the definition of another axiom that is necessary for such beliefs to induce risk preferences that are consistent with the underlying preferences over acts. In conjunction with the usual ordering and nondegeneracy axioms and a slight strengthening of Savage's continuity postulate P6 (Small Event Continuity), this axiom along with one of the relaxations of P3 (Eventwise Monotonicity) implies probabilistically sophisticated beliefs, and induced risk preferences that need only satisfy a conditional monotonicity property.

In the final section I return to the first motivation mentioned above and address the issue of whether it is possible to characterize probabilistically sophisticated preferences without requiring the induced preferences to exhibit any property save some form of continuity. I conjecture that in the Savage framework of purely subjective uncertainty such a characterization is not possible since the ability to deduce the beliefs that an individual holds about the relative likelihood of events from his or her preferences over acts appears to rely crucially upon some form of eventwise monotonicity holding, albeit only "locally" and/or "conditionally," for those preferences. Proofs for the results stated in the text are collected together in an Appendix.

2. THE SAVAGE FRAMEWORK AND PROBABILISTIC SOPHISTICATION

I follow M-S (1992) and adopt Savage's (1972) setting of purely subjective uncertainty where the objects of choice are "acts" mapping states of nature to outcomes.
Savage Set-up

$\mathcal{S} = \{s, t, \ldots\}$ a set of states.  
$\mathcal{E} = 2^\mathcal{S} = \{A, B, C, \ldots\}$ the set of events (i.e., all subsets of $\mathcal{S}$).  
$\mathcal{X} = \{x, y, z, \ldots\}$ a set of outcomes or consequences.  
$\mathcal{A} = \{f, g, h, \ldots\}$ the set of finite-outcome acts on $\mathcal{S}$. ($f$ is said to be a finite-outcome act if its outcome set $f(\mathcal{S}) = \{f(s) | s \in \mathcal{S}\}$ is finite.)

$x$ (resp. $y$, $z$) will be used to denote both the outcome $x$ (resp. $y$, $z$) in $\mathcal{X}$ and the constant act $f(\mathcal{S}) = \{x\}$ (resp. $g(\mathcal{S}) = \{y\}$, $h(\mathcal{S}) = \{z\}$).

$\succeq$ is a (binary) relation over ordered pairs of acts, representing the agent’s preferences.

$\succeq$ and $\sim$ correspond to strict preference and indifference, respectively.

$\succeq_x$ is the relation over ordered pairs of outcomes obtained from $\succeq$ for constant acts (i.e., $x \succeq_x y \Leftrightarrow x \succeq y$).

Throughout this paper I will make the following assumption:

P1 (Ordering): $\succeq$ is complete, reflexive, and transitive.

For a given relation, $\succeq$, an event $E$ is said to be null if for any pair of acts, $f$ and $g$, which differ only on $E$, we have $f \sim g$.

Let

$$\mathcal{P}_0(\mathcal{X}) = \left\{ P = (x_1, p_1, \ldots, x_n, p_n) \left| \sum_{i=1}^n p_i = 1, x_i \in \mathcal{X}, p_i \geq 0 \right. \right\},$$

be the set of simple (i.e. finite outcome) probability distributions (one-stage) lotteries over outcomes. Let $\delta_x$ denote the degenerate lottery $P = (x, 1)$.

A probability measure, $\mu$, over $\mathcal{E}$, can be used to map each act in $\mathcal{A}$ to a lottery in $\mathcal{P}_0(\mathcal{X})$ in the following way: $f \mapsto \mu(f^{-1})$. For example, the act $g = [x$ if $s \in A; y$ if $s \notin A]$ would be mapped to the lottery $(x, \mu(A)); y, 1 - \mu(A))$, as $g^{-1}(x) = A, g^{-1}(y) = \mathcal{S}/A$, and $g^{-1}(z) = \emptyset$ for all $z$ not in $\{x, y\}$. In order for this mapping to be onto, we require the state space to be infinitely divisible and the probability measure to be nonatomic.

This mapping induces preferences over lotteries, denoted $\succeq_p$, from the underlying preferences over acts, $\succeq$, in the following manner. If $\mu$ maps acts $f$ and $g$ to lotteries $P$ and $Q$ respectively, then $f \succeq g$ implies $P \succeq_p Q$. Formally:

$$\succeq_p = \{(P, Q) \in \mathcal{P}_0(\mathcal{X}) \times \mathcal{P}_0(\mathcal{X}) | \exists (f, g) \in \succeq, P = \mu(f^{-1}) \text{ and } Q = \mu(g^{-1}) \},$$

The probability measure, $\mu$, representing the individual’s beliefs about the likelihood of events, and the lottery preferences $\succeq_p$, which $\mu$ induces, usefully characterize the individual’s underlying preferences over acts, $\succeq$, if we can reconstruct $\succeq$ from knowledge of $\mu$ and $\succeq_p$ alone. The ability to achieve such a separation of beliefs from risk preferences defines a probabilistically sophisticated individual.
DEFINITION (Probabilistic Sophistication): $\succeq$ is *probabilistically sophisticated* if there exists a unique finitely additive nonatomic probability measure $\mu$ on $\mathcal{A}$, inducing a relation, $\succeq_P$, over lotteries, such that: for all $P, Q$ in $\mathcal{P}_0(\mathcal{A})$, and all $f, g$ in $\mathcal{A}$, $P \succeq_P Q$, $\mu \circ f^{-1} = P$ and $\mu \circ g^{-1} = Q$ implies $f \succeq g$.

This definition is more general than the one formally presented in M-S (1992, p. 755), as they require the induced lottery preferences to respect the following notion of (*nondimensional*) monotonicity.\(^3\)

DEFINITION: Given a complete pre-order $\succeq_x$ over outcomes, $P = (x_1, p_1; \ldots; x_n, p_n)$ weakly first order stochastically dominates (FSD) $Q = (y_1, q_1; \ldots; y_m, q_m)$ with respect to $\succeq_x$ if $\Sigma_{i \mid x_i \succeq_x x_j} p_i \leq \Sigma_{i \mid y_i \succeq_y y_j} q_i$ for all $x \in \mathcal{A}$, and if in addition this expression holds with strict inequality for some $y \in \mathcal{A}$, then $P$ strictly FSD $Q$ with respect to $\succeq_x$.

$\succeq_P$ respecting this notion of first order stochastic dominance simply means that $P \succeq_P (\succeq_P) Q$ whenever $P$ weakly (strictly) FSD $Q$ with respect to $\succeq_x$.

I contend that the advantage of the definition presented above over the one employed by M-S (1992) is that it characterizes probabilistically sophisticated beliefs without imposing any further structural requirements on the induced lottery preferences.

If $\succeq$ is a complete pre-order, it is straightforward to show that probabilistic sophistication is equivalent to $\mu \circ f^{-1} = \mu \circ g^{-1}$ implying $f \sim g$. Informally speaking, this says that how an outcome in an act is viewed by an individual depends only on the individual’s assessed *likelihood* of the event in which that outcome arises, and *not* on the particular event itself. Essentially this requires that the individual be indifferent between any two acts that differ only on two equally likely disjoint events $A$ and $B$, where one has outcome $x$ on $A$ and $y$ on $B$, while the other has $y$ on $A$ and $x$ on $B$. I dub this property of the preferences and the probability measure representing beliefs *Event Independence*.\(^4\)

EI (Event Independence): *For all pairs of disjoint events A and B, outcomes x and y, and act h,*

$$\mu(A) = \mu(B) \quad \text{implies}$$

$$
\begin{bmatrix}
x & \text{if } s \in A \\
y & \text{if } s \in B \\
h(s) & \text{if } s \notin (A \cup B)
\end{bmatrix}
\sim
\begin{bmatrix}
y & \text{if } s \in A \\
x & \text{if } s \in B \\
h(s) & \text{if } s \notin (A \cup B)
\end{bmatrix}.
$$

\(^3\)Fishburn and Vickson (1978, Sec. 2.21, p. 97).

\(^4\)I thank Ben Polak for motivating this discussion and his help with defining what constitutes event independence in this context.
It follows immediately from Savage’s Theorem 5.2.1 (1972, p. 70) and M-S’s Theorem 1 (1992, p. 765) that both their axiomatizations of preferences over acts satisfy this definition of probabilistic sophistication. But Savage’s Theorem 5.2.2 (1972, p. 72) shows further that the lottery preferences induced using these probabilistically sophisticated beliefs also exhibit Independence. As M-S (1992) note, this property of the induced lottery preferences arises from the underlying preferences over acts satisfying P2 (Sure Thing Principle).

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

\[
\begin{pmatrix}
  f(s) & \text{if } s \in E \\
  g(s) & \text{if } s \notin E
\end{pmatrix} \succeq \begin{pmatrix}
  f^*(s) & \text{if } s \in E \\
  g(s) & \text{if } s \notin E
\end{pmatrix} \Rightarrow \begin{pmatrix}
  f(s) & \text{if } s \in E \\
  h(s) & \text{if } s \notin E
\end{pmatrix} \succeq \begin{pmatrix}
  f^*(s) & \text{if } s \in E \\
  h(s) & \text{if } s \notin E
\end{pmatrix}
\]

It states that if two acts differ on an event $E$ but are identical on the complementary event, $E^c$, then the agent’s preference between these two acts does not depend on what appears in the event where the acts agree. But this “separability” of preferences across mutually exclusive events is precisely the subjective uncertainty analog of Independence.

Although M-S (1992) dispense with P2, they do retain Savage’s monotonicity postulate P3.

P3 (Eventwise Monotonicity): For all outcomes $x$ and $y$, non-null events $E$ and acts $h$,

\[
\begin{pmatrix}
  x & \text{if } s \in E \\
  h(s) & \text{if } s \notin E
\end{pmatrix} \succeq \begin{pmatrix}
  y & \text{if } s \in E \\
  h(s) & \text{if } s \notin E
\end{pmatrix} \Leftrightarrow x \succeq y.
\]

P3 can be interpreted as saying that the preference for receiving outcome $x$ to receiving outcome $y$ conditional on event $E$ obtaining is determined solely by the unconditional preference for receiving $x$ for sure to receiving $y$ for sure and is unaffected by the act $h$ that determines the outcome if the complement of $E$ obtains. Although this “separability” of conditional preference for outcomes does not imply Independence for the induced lottery preferences, it does imply that induced lottery preferences satisfy the following substitution axiom.

**Axiom of Degenerate Independence (ADI):** For all simple lotteries $P$, outcomes $x$ and $y$, and $\alpha$ in $(0,1)$,

\[x \succeq y \text{ if and only if } \alpha \delta_x + (1 - \alpha)P \succeq_{P} \alpha \delta_y + (1 - \alpha)P.\]

**Proposition 2.1:** If $\succeq$ is a complete pre-order and is probabilistically sophisticated with respect to the finitely additive nonatomic probability measure $\mu$ on $\mathcal{E}$, $\succeq$ satisfies P3 if and only if $\succeq_{P}$ exhibits ADI.
SUBJECTIVE PROBABILITY

It seems to me that to answer the question first posed by M-S (1992) and quoted in the introduction above, one should endeavor to dispense with extraneous requirements on the induced lottery preferences such as ADI as well as Independence. Hence we should investigate whether it is possible to axiomatize probabilistically sophisticated preferences without recourse to invoking either P3 or P2. Moreover, as I demonstrate in the next section there are natural instances where an individual’s risk preferences would not satisfy ADI but where we would not necessarily wish to rule out the possibility that his or her underlying preferences over acts were probabilistically sophisticated.

It is interesting to note that a simple induction argument shows that a preference ordering over lotteries satisfies ADI if and only if it respects first order stochastic dominance. Hence by assuming P3, M-S (1992) automatically get respect of first order stochastic dominance for the lottery preferences induced from the probabilistically sophisticated preferences that they characterize.

3. WHY LOTTERY PREFERENCES MAY NOT EXHIBIT ADI

The nature of the Independence Axiom and its implications for decision making under risk have been discussed extensively in the literature. ADI, although formally weaker than Independence, is also a substitution axiom that restricts preferences over lotteries. By insisting that an individual’s induced lottery preferences obey ADI we are allowing the individual to distinguish between outcomes in a lottery only in terms of those outcomes’ desirability “assessed” when they are received for sure. In many situations this is no doubt a reasonable restriction to place on lottery preferences and has the advantage that we can characterize the individual’s behavior under risk by his or her induced preferences for lotteries over the indifference sets of outcomes—a space of univariate distributions.

In the introduction I cited two instances where we might reasonably find risk preferences failing to agree with the monotonicity property implied by ADI. I shall present two examples that serve both to illustrate the intuitions behind these failures of ADI, and that will, in the next section, aid in the motivation of the axioms that characterize probabilistically sophisticated preferences over acts leading to these forms of risk preferences.

The first example is of an individual who not only cares about the final outcome after the uncertainty has been resolved, but is also concerned with the process by which the outcomes are allocated. Recall Machina’s (1989) example of a Mom who has an indivisible treat which she can give to either her daughter, Abigail, or her son, Benjie. We are told she is indifferent between the outcome of Abigail or Benjie receiving the treat but prefers the “even-chance” lottery to all other lotteries that allocate the treat to either child.

As the outcome set consists of just two outcomes (i.e., either Abigail or Benjie getting the treat) we can identify each lottery by a number \( p_A \) in the unit interval, denoting the probability that the treat is allocated to Abigail. Intuitively, we might expect Mom to hold the view that the more “variable” the lottery is, the “fairer” it is as a procedure for allocating the treat. Hence we might reasonably expect her preferences for these two outcome lotteries to be representable by any monotonic transformation of the function:

\[
V(p_A) = p_A(1 - p_A).
\]

It follows immediately from Proposition 2.1 and the fact that \( V(\cdot) \) has a unique interior maximum on \([0, 1]\), that Mom’s preferences over these two-outcome acts (defined on some state space) cannot both be probabilistically sophisticated and satisfy P3. Yet there seems no logical reason, a priori, why Mom could not have such risk preferences and beliefs about the likelihood of events that conformed to a probability measure.

Following the example discussed in Diamond (1967), upon which the Machina Mom example is based, these preferences can be alternatively interpreted as a particular instance of a social ordering over lotteries of social outcomes that admits a representation by a nonutilitarian social welfare function. To see this let \( I \) represent the number of individuals in society. Each social outcome can be identified by an \( I \)-dimensional vector that denotes each individual’s welfare for that outcome, that is, \( X \subseteq \mathbb{R}^I_+ \). A social welfare function representing the social ordering over lotteries of social outcomes for a society of expected utility maximizers can then be expressed as follows:

\[
V(P) = W\left( \sum_{m=1}^{n} p_m x_m^1, \ldots, \sum_{m=1}^{n} p_m x_m^I \right).
\]

It is utilitarian (in the Harsanyi (1955) sense) if and only if \( W(u^1, \ldots, u^I) = F(\sum_{m=1}^{I} u^m) \). In the Mom example, \( I = 2 \), \( X = \{(1, 0), (0, 1)\} \), the domain of \( W \) is \( \{(u_A, u_B) \in \mathbb{R}^2_+ | u_A + u_B = 1\} \) and \( W(u_A, u_B) = u_A u_B \) which is clearly not utilitarian.

In a recent paper Epstein and Segal (1992) have proposed a set of axioms for a social ordering over lotteries that incorporates this concept of ex ante randomization as a means for enhancing the fairness of the procedure by which the social outcome is chosen. For our purposes the most interesting feature of their representation theorem is that the function, \( W \), derived from their postulates is necessarily strictly quasi-concave. That is, if \( u = (u^1, \ldots, u^I) \) and \( v = (v^1, \ldots, v^I) \) are distinct vectors and \( W(u) = W(v) \), then \( W(\alpha u + [1 - \alpha] v) > W(u) \) for any \( \alpha \) in \((0, 1)\). Note from the definition of \( V \) above, it follows that \( V \) will then also be strictly quasi-concave in probabilities. But a social ordering that admits such a representation may well violate ADI since if for some pair of distinct outcomes \( x \) and \( y \), \( \lambda \) in \((0, 1)\) and some lottery \( P \) we have \( V(\lambda \delta_x + [1 - \lambda]P) = V(\lambda \delta_y + [1 - \lambda]P) \), then from the strict quasi-concavity of \( W \) it follows
that \( V(\alpha \delta_x + (1-\alpha)\delta_y) + [1-\lambda]P) > V(\lambda \delta_x + [1-\lambda]P) \) for all \( \alpha \) in \( (0,1) \). Hence, in general, a social ordering over subjectively uncertain acts (that map states to social outcomes), that embodies such a concern for the ex ante fairness of the procedure by which the social outcome is determined, may well fail to satisfy P3. But yet again there is no reason, a priori, why such a social preference ordering should not admit beliefs about the likelihood of events that conform to a probability measure.

The second example is of a Savage subjective expected utility maximizer who must choose an act and then undertake some auxiliary action before the uncertainty is resolved. Following M-S (1992), let the individual’s von Neumann-Morgenstern utility function be \( u \) and subjective probability measure be \( \mu \). His or her subjective expected utility for a given choice of a finite outcome act \( f \) and action \( \alpha \), from a set of auxiliary actions \( D \), is

\[
\hat{V}(f, \alpha) = \sum_i \mu(f^{-1}(x_i)) \cdot u(x_i, \alpha).
\]

Hence the ranking of acts is made on the basis of the induced functional

\[
\mathcal{W}(f) = \max_{\alpha \in D} \left\{ \sum_i \mu(f^{-1}(x_i)) \cdot u(x_i, \alpha) \right\}.
\]

As M-S (1992) point out, the optimal choice of the auxiliary action depends upon an act, \( f \), only through its outcomes \( f(\mathcal{X}) = \{x_1, \ldots, x_n\} \) and their respective probabilities \( \{\mu(f^{-1}(x_1)), \ldots, \mu(f^{-1}(x_n))\} \).

Thus, the preference functional representing the risk preferences induced by the subjective probability measure \( \mu \) is given by:

\[
(2) \quad \text{for } p = (x_1, p_1, \ldots, x_n, p_n) \quad V(P) = \max_{\alpha \in D} \left\{ \sum_{i=1}^n p_i \cdot u(x_i, \alpha) \right\}.
\]

As is well-known (see for instance Machina (1984)) such a functional need only be convex in probabilities, that is for all \( P, Q \) in \( \mathcal{P}_0(\mathcal{X}) \) and \( \lambda \) in \( (0,1) \) it follows that \( V(\lambda P + [1-\lambda]Q) \leq \lambda V(P) + [1-\lambda]V(Q) \). What is less appreciated and is contrary to what M-S (1992) implicitly claim, is that a functional such as defined in (2) may fail to respect first order stochastic dominance with respect to \( \succeq_X \).

To see this consider the case where there are just two possible outcomes, \( x_A \) and \( x_B \), and two possible auxiliary actions, \( a \) and \( b \). Hence, we can identify each lottery by the probability it places on \( x_A \), denoted by \( p_A \), and the preference functional in equation (2) can be re-expressed as

\[
V(p_A) = \max \left\{ \left[ p_A u(x_A, a) + (1-p_A) u(x_B, a) \right], \left[ p_A u(x_A, b) + (1-p_A) u(x_B, b) \right] \right\}.
\]

Now, suppose that if the individual knew for sure that he or she would receive \( x_A \) (resp. \( x_B \), then he or she would strictly prefer to undertake auxiliary action \( a \) (resp. \( b \)). Suppose, in addition, having undertaken action \( a \) (resp. \( b \), the
individual would strictly prefer to receive outcome $x_A$ (resp. $x_B$). That is, $u(x_A, a) > u(x_A, b), u(x_B, b) > u(x_B, a), u(x_A, a) > u(x_B, a)$, and $u(x_B, b) > u(x_A, b)$. Letting $\Delta u(x_A) = u(x_A, a) - u(x_A, b)$ and $\Delta u(x_B) = u(x_B, b) - u(x_B, a)$, $V(p_A)$ can now be rewritten as

$$V(p_A) = \begin{cases} p_A u(x_A, b) + (1 - p_A) u(x_B, b) & \text{if } p_A < \Delta u(x_B)/[\Delta u(x_A) + \Delta u(x_B)], \\ p_A u(x_A, a) + (1 - p_A) u(x_B, a) & \text{if } p_A \geq \Delta u(x_B)/[\Delta u(x_A) + \Delta u(x_B)], \end{cases}$$

which is clearly nonmonotonic in $p_A$. The intuition is straightforward; whatever the relative values of $x_A$ and $x_B$, assessed when each is received for sure (and thus having undertaken the appropriate auxiliary action), given a small chance of receiving $x_A$, the better action to take is $b$ in which case $x_B$ is, under these conditions, preferred to $x_A$. Conversely, given a relatively large chance of receiving $x_A$, the better action to take is $a$ and thus $x_A$ now becomes the preferred outcome.

Proponents of monotonicity could argue (echoing their proponents of expected utility “cousins”) that what we are calling an outcome in the last two examples corresponds to many different outcomes which Mom and respectively, the individual in the delayed resolution of uncertainty case, do not value equally. But as Machina (1989) argues, distinguishing for Mom each outcome by the probability distribution in which it arises makes the theory vacuous. In the delayed resolution case, it is feasible that the outcome set could be extended to include all possible outcome and auxiliary decision pairs. But as Machina notes: “this level of description may be below the usual level at which economists typically operate with or can observe . . .”

4. PROBABILISTIC SOPHISTICATION FOR MACHINA’S MOM AND A DECISION-MAKER FACING DELAYED RESOLUTION OF UNCERTAINTY

As was demonstrated in the last section a probabilistically sophisticated Mom (or social planner) concerned with ex ante fairness and a Savage expected utility maximizer undertaking an unobservable action before the resolution of uncertainty may have induced (observable) risk preferences that violate ADI and hence their underlying (observable) preference relation over subjectively uncertain acts may fail to satisfy P3. However, they can be seen to satisfy suitable relaxations of P3.

---

6Machina (1989, p. 1663) (italics in original). Similarly Kreps and Porteus (1979, p. 83) have noted that “the obvious difficulty with this approach is that such complete models may become overburdened with detail and analytically intractable.”
Consider first a social ordering over acts, \( \succeq \), that embodies a notion of concern for ex ante fairness such as displayed by Machina's Mom. For an event \( E \), a pair of social outcomes \( x \) and \( y \), and an act \( h \), we can interpret \([x \text{ if } E; h(s) \text{ if } E^c] > [y \text{ if } E; h(s) \text{ if } E^c]\) as saying that outcome \( x \) is preferred to outcome \( y \) conditional on event \( E \) obtaining and act \( h \) occurring if the complement of event \( E \) obtains. P3 implies that for any other act \( g \), \([x \text{ if } E; g(s) \text{ if } E^c] > [y \text{ if } E; g(s) \text{ if } E^c]\). Note, however, that there is no reason, given the concern for fairness embodied in the (social) ordering, that the preference for \( x \) over \( y \) conditional on event \( E \) obtaining should be invariant to whatever appears if the complement of \( E \) obtains. For example, returning explicitly to the Machina's Mom example and letting \( x_A \) (respectively, \( x_B \)) denote the outcome where Abigail (respectively, Benjie) receives the treat, we would reasonably expect for any non-null event \( E \) that

\[
[x_A \text{ if } E; x_B \text{ if } E^c] > [x_B \text{ if } E; x_B \text{ if } E^c] \quad \text{but} \quad [x_B \text{ if } E; x_A \text{ if } E^c] > [x_A \text{ if } E; x_A \text{ if } E^c].
\]

P3 also implies that for any non-null subset of \( E \), say \( D \), with \( E \setminus D \) also non-null, that:

\[
[x \text{ if } E; h(s) \text{ if } E^c] > [x \text{ if } D; y \text{ if } E \setminus D; h(s) \text{ if } E^c] > [y \text{ if } E; h(s) \text{ if } E^c].
\]

But even this weaker conditional monotonicity condition need not hold. Replacing the conditionally preferred \( x \) for \( y \) on the event \( D \) in the act \([y \text{ if } E; h(s) \text{ if } E^c] \) will improve the desirability of the act not only for the usual intuitive "monotonic" reasons but also for "fairness" reasons by way of the "mixing" of \( x \) and \( y \) on the event \( E \). These two effects enable us to infer that \([x \text{ if } D; y \text{ if } E \setminus D; h(s) \text{ if } E^c]\) but the reinforcement of the "fairness" benefit on the usual monotonicity benefit precludes us from ruling out the possibility that

\[
[x \text{ if } D; y \text{ if } E \setminus D; h(s) \text{ if } E^c] > [x \text{ if } E; h(s) \text{ if } E^c].
\]

Again returning to the Machina Mom example, say that for three non-null disjoint events \( A, B, \) and \( C \) that form a partition of \( S \), Mom views \( A \) as "equally likely" as \( B \cup C \). Then it is quite reasonable for Mom to express \([x_A \text{ if } A \cup B; x_B \text{ if } C] > x_B \) and \([x_A \text{ if } A; x_B \text{ if } B \cup C] > x_B \) but \([x_A \text{ if } A; x_B \text{ if } B \cup C] > [x_A \text{ if } A \cup B; x_B \text{ if } C] \). This leads us naturally to the following weakening of P3 for a social planner such as Machina's Mom.

**P3** (Conditional Upper Eventwise Monotonicity): For all pairs of non-null, disjoint events \( A \) and \( B \), all outcomes \( x \) and \( y \), and all acts \( h \):

\[
\begin{align*}
x & \quad \text{if } s \in (A \cup B) \\
\[h(s) & \quad \text{if } s \notin (A \cup B) > (\succeq) \quad \begin{cases} 
y & \quad \text{if } s \in (A \cup B) \\
h(s) & \quad \text{if } s \notin (A \cup B) \end{cases}
\end{align*}
\]
implies
\[
\begin{bmatrix}
  x & \text{if } s \in A \\
  y & \text{if } s \in B \\
  h(s) & \text{if } s \not\in (A \cup B)
\end{bmatrix} \succ (\succeq) \begin{bmatrix}
  y & \text{if } s \in (A \cup B) \\
  h(s) & \text{if } s \not\in (A \cup B)
\end{bmatrix}.
\]

Returning to the example of a subjective expected utility maximizer who undertakes an unobservable auxiliary action between the time of choosing an act and the resolution of the uncertainty, again we should not expect the preference between two outcomes conditional on an event obtaining to be invariant to whatever occurs if the complement of that event obtains. The reason is simply that what occurs if that complementary event obtains may affect the choice of the optimal auxiliary action which in turn may affect the relative desirability of the two outcomes.

If \([x \text{ if } E; h(s) \text{ if } E^c] \succ [y \text{ if } E; h(s) \text{ if } E^c]\), then whatever optimal auxiliary actions are chosen for either act it readily follows for this individual that for any non-null subset of \(E\), say \(C\), \([x \text{ if } E; h(s) \text{ if } E^c] \succ [y \text{ if } C; x \text{ if } E \setminus C; h(s) \text{ if } E^c]\). However, as was illustrated for the risk preferences of such an individual in the previous section, given the optimal auxiliary action for the act \([y \text{ if } E; h(s) \text{ if } E^c]\) has been taken, the outcome \(x\) may not be preferred to \(y\) if \(E\) obtains. Hence for an event \(C\) which is a "large" enough subset of \(E\) we cannot rule out the possibility that \([y \text{ if } E; h(s) \text{ if } E^c] \succ [y \text{ if } C; x \text{ if } E \setminus C; h(s) \text{ if } E^c]\). Thus we can only require for a Savage expected utility maximizer facing a situation of delayed resolution of uncertainty that his or her observable act-preferences satisfy the following alternative weakening of P3.

**P3\textsuperscript{CL} (Conditional Lower Eventwise Monotonicity):** For all pairs of non-null, disjoint events \(A\) and \(B\), all outcomes \(x\) and \(y\), and all acts \(h\):

\[
\begin{bmatrix}
  x & \text{if } s \in (A \cup B) \\
  h(s) & \text{if } s \not\in (A \cup B)
\end{bmatrix} \succ (\succeq) \begin{bmatrix}
  y & \text{if } s \in (A \cup B) \\
  h(s) & \text{if } s \not\in (A \cup B)
\end{bmatrix}
\]

\[7\text{Note that no direct inference can be drawn about the individual's relative ranking of}
\[
\begin{bmatrix}
  x & \text{if } s \in (A \cup B) \\
  h(s) & \text{if } s \not\in (A \cup B)
\end{bmatrix}
\]

and

\[
\begin{bmatrix}
  x & \text{if } s \in A \\
  y & \text{if } s \in B \\
  h(s) & \text{if } s \not\in (A \cup B)
\end{bmatrix};
\]

\[\text{hence the qualifier Upper in the name of the axiom.}\]
implies
\[
\begin{bmatrix}
x & \text{if } s \in (A \cup B) \\
h(s) & \text{if } s \notin (A \cup B)
\end{bmatrix} \succ (\succsim) \begin{bmatrix}
x & \text{if } s \in A \\
y & \text{if } s \in B \\
h(s) & \text{if } s \notin (A \cup B)
\end{bmatrix}^8.
\]

Since we know from Proposition 2.1 that probabilistically sophisticated preferences that satisfy P3 have induced risk preferences that exhibit ADI, it is of interest to note what property for the induced risk preferences are implied by P3CU and P3CL respectively. It is straightforward to show that they imply the following conditional monotonicity properties.

**Definition (Conditional Upward Monotonicity):** For all \( x, y \in \mathcal{X} \), all \( P \) in \( \mathcal{P}_0(\mathcal{X}) \), and all \( \lambda, \gamma \in (0, 1] \):
\[
\lambda \delta_x + (1 - \lambda) P \succ_P (\succeq_P) \lambda \delta_y + (1 - \lambda) P
\]
implies
\[
\lambda (\gamma \delta_x + [1 - \gamma] \delta_y) + (1 - \lambda) P \succ_P (\succeq_P) \lambda \delta_y + (1 - \lambda) P.
\]

**Definition (Conditional Downward Monotonicity):** For all \( x, y \in \mathcal{X} \), all \( P \) in \( \mathcal{P}_0(\mathcal{X}) \), and all \( \lambda, \gamma \in (0, 1] \):
\[
\lambda \delta_x + (1 - \lambda) P \succ_P (\succeq_P) \lambda \delta_y + (1 - \lambda) P
\]
implies
\[
\lambda \delta_x + (1 - \lambda) P \succ_P (\succeq_P) \lambda (\gamma \delta_x + [1 - \gamma] \delta_y) + (1 - \lambda) P.
\]

**Proposition 4.1:** Given \( \succeq \) is a complete pre-order and is probabilistically sophisticated with respect to the finitely additive nonatomic probability measure \( \mu \) on \( \mathcal{E} \), \( \succeq \) satisfies P3CU (respectively, P3CL) if and only if \( \succeq_P \) exhibits Conditional Upward (respectively, Downward) Monotonicity.

Both these conditional monotonicity properties obviously hold for an expected utility maximizer as he or she can be represented by a functional that is linear in the probabilities. It is also immediately apparent that conditional upward (respectively, downward) monotonicity will be exhibited by risk preferences that can be represented by a functional that is quasi-concave (respectively,

---

Footnote 8: Note that no direct inference can be drawn about the individual's relative ranking of
\[
\begin{bmatrix}
y & \text{if } s \in (A \cup B) \\
h(s) & \text{if } s \notin (A \cup B)
\end{bmatrix}
\]
and
\[
\begin{bmatrix}
x & \text{if } s \in A \\
y & \text{if } s \in B \\
h(s) & \text{if } s \notin (A \cup B)
\end{bmatrix};
\]
hence the qualifier *Lower* in the name of the axiom.
quasi-convex) in probabilities. Thus these are appropriate properties for the non-ADI risk preferences discussed in Section 3 above to exhibit.

For preferences that satisfy P3, such as Savage’s subjective expected utility maximizer and M-S’s probabilistically sophisticated non-expected utility maximizer, it is possible to deduce the individual’s beliefs over relative likelihood of events from simple betting preferences. That is, if the individual prefers outcome \( x \) to outcome \( y \), then we can infer that he or she believes event \( A \) is as likely as event \( B \), if and only if:

\[
\begin{bmatrix}
  x & \text{if } s \in A \\
  y & \text{if } s \not\in A \\
\end{bmatrix} \succ
\begin{bmatrix}
  x & \text{if } s \in B \\
  y & \text{if } s \not\in B \\
\end{bmatrix}.
\]

For preferences that fail to satisfy P3 such an inference about the relative likelihood of events cannot be drawn simply from a preference existing between the above pair of two-outcome gambles since the conditional ranking of two outcomes no longer necessarily accords with their unconditional ranking.\(^9\) However, \( P3^{\text{CU}} \) and \( P3^{\text{CL}} \) can be shown to imply a conditional eventwise monotonicity property which then allows us from the conditional preference of one outcome over another to infer the relative likelihood of certain events from the preference between conditional gambles involving those two outcomes.

**Proposition 4.2:** Given any pair of outcomes, \( x \) and \( y \); any pair of non-null, disjoint events \( A \) and \( B \); and any act \( g \) for which

\[
\begin{bmatrix}
  x & \text{if } s \in A \cup B \\
  g(s) & \text{if } s \not\in A \cup B \\
\end{bmatrix} \succ
\begin{bmatrix}
  x & \text{if } s \in A \\
  y & \text{if } s \in B \\
  g(s) & \text{if } s \not\in A \cup B \\
\end{bmatrix}
\]

\[
\succ
\begin{bmatrix}
  y & \text{if } s \in A \cup B \\
  g(s) & \text{if } s \not\in A \cup B \\
\end{bmatrix};
\]

\( \succ \) satisfies \( P3^{\text{CU}} \) (respectively, \( P3^{\text{CL}} \)) implies for all non-null events \( C \subseteq A \) (respectively, \( B \)), and acts \( f \) with \( f(\mathcal{I}) \subseteq \{x, y\} \)

\[
\begin{bmatrix}
  x & \text{if } s \in C \\
  f(s) & \text{if } s \in A \setminus C \\
  y & \text{if } s \in B \\
  g(s) & \text{if } s \not\in A \cup B \\
\end{bmatrix} \succ
\begin{bmatrix}
  y & \text{if } s \in C \\
  f(s) & \text{if } s \in A \setminus C \\
  y & \text{if } s \in B \\
  g(s) & \text{if } s \not\in A \cup B \\
\end{bmatrix}
\]

\(^9\)I.e., the ranking obtained by comparing the two constant acts.
SUBJECTIVE PROBABILITY

\[ \begin{pmatrix}
  x & \text{if } s \in C \\
  f(s) & \text{if } s \in B \setminus C \\
  x & \text{if } s \in A \\
  g(s) & \text{if } s \notin A \cup B
\end{pmatrix}
\begin{pmatrix}
  y & \text{if } s \in C \\
  f(s) & \text{if } s \in B \setminus C \\
  x & \text{if } s \in A \\
  g(s) & \text{if } s \notin A \cup B
\end{pmatrix} \]

That is, given the hypothesis of Proposition 4.2, P3\textsuperscript{CU} (respectively, P3\textsuperscript{CL}) implies that eventwise monotonicity holds on event \( A \) (respectively, \( B \)) for acts with outcomes \( x \) and/or \( y \), given \( y \) if \( B \) (respectively, \( x \) if \( A \)) obtains and that act \( g \) determines the outcome if neither \( A \) nor \( B \) obtains.

An intuition for this result can be gleaned by supposing there exists a function \( v \) that represents preferences over the set of acts

\[ \mathcal{A}^* \equiv \{ f \in \mathcal{A} \mid f(A \cup B) \subseteq \{ x, y \}, f(s) = g(s) \text{ if } s \notin A \cup B \} \]

Moreover, suppose these preferences are probabilistically sophisticated with beliefs represented by the probability measure \( \mu \). If we define \( v : [0, 1] \to \mathbb{R} \) by \( v(\lambda) = v(f) \) for \( f \in \mathcal{A}^* \) and \( \lambda = \mu(f^{-1}(x) \cap [A \cup B]) / \mu(A \cup B) \), then P3\textsuperscript{CU} (respectively, P3\textsuperscript{CL}) implies Conditional Upward (respectively, Downward) Monotonicity for the induced risk preferences which in turn implies that \( v \) is “single-peaked” (respectively, “single-dipped”). Letting \( \lambda^* = \mu(A) / \mu(A \cup B) \), we have from the hypothesis of Proposition 4.2 that \( v(0) < v(\lambda^*) < v(1) \) and thus “single-peakedness” (respectively, “single-dippedness”) implies that \( v \) is monotonically increasing in \( \lambda \) at least for all \( \lambda \leq (\text{respectively, } \geq) \lambda^* \). (See Figure 1.)

\[\text{Figure 1a.} \text{— Conditionally upward monotonic risk preferences.}\]

\[\text{Figure 1b.} \text{— Conditionally downward monotonic risk preferences.}\]
A further implication of probabilistically sophisticated preferences that satisfy P3\(^{\text{CU}}\) or P3\(^{\text{CL}}\) is that

\[
\begin{bmatrix}
    x & \text{if } s \in A \cup B \\
g(s) & \text{if } s \notin A \cup B
\end{bmatrix} \succ
\begin{bmatrix}
    x & \text{if } s \in A \\
y & \text{if } s \in B \\
g(s) & \text{if } s \notin A \cup B
\end{bmatrix}
\]

\[
\sim
\begin{bmatrix}
y & \text{if } s \in A \\
x & \text{if } s \in B \\
g(s) & \text{if } s \notin A \cup B
\end{bmatrix}
\]

\[
\succ
\begin{bmatrix}
y & \text{if } s \in A \cup B \\
g(s) & \text{if } s \notin A \cup B
\end{bmatrix}
\]

implies that event \( A \) must be assessed as least as likely as event \( B \). Again an intuition for this property can be gleaned from the function \( v \) defined above. Referring to Figure 1 we see that by the single peakedness or dippeness of \( v \), \( v(1) > v(\mu(A)/\mu(A \cup B)) \geq v(\mu(B)/\mu(A \cup B)) > v(0) \) can only hold if \( \lambda^* = \mu(A)/\mu(A \cup B) \geq 0 \cdot 5 \geq \mu(B)/\mu(A \cup B) = 1 - \lambda^* \).

One might be then tempted to suppose that having discerned that event \( A \) is deemed as least as likely as event \( B \), that if outcome \( w \) is conditionally preferred to outcome \( z \) if \( A \cup B \) obtains given some act \( h \) determines the outcome if \( (A \cup B)^c \) obtains, then the subact with \( w \) on \( A \) and \( z \) on \( B \) should be weakly preferred to the subact with \( z \) on \( A \) and \( w \) on \( B \) given that the same act \( h \) determines the outcome if neither \( A \) nor \( B \) obtains. That is,

\[
\begin{bmatrix}
w & \text{if } s \in A \cup B \\
h(s) & \text{if } s \notin A \cup B
\end{bmatrix} \succ
\begin{bmatrix}
z & \text{if } s \in A \cup B \\
h(s) & \text{if } s \notin A \cup B
\end{bmatrix}
\]

\[
\Rightarrow
\begin{bmatrix}
w & \text{if } s \in A \\
z & \text{if } s \in B \\
h(s) & \text{if } s \notin A \cup B
\end{bmatrix} \sim
\begin{bmatrix}
z & \text{if } s \in A \\
w & \text{if } s \in B \\
h(s) & \text{if } s \notin A \cup B
\end{bmatrix}.
\]

Note, however, even given the preference between the first pair of acts, neither P3\(^{\text{CU}}\) nor P3\(^{\text{CL}}\) on their own guarantees that \([w \text{ if } A; z \text{ if } B; h(s) \text{ if } (A \cup B)^c]\) lies in preference between that pair of acts and so the second preference relation need not follow.

On the other hand, as discussed in Section 2, a necessary requirement for probabilistically sophisticated preferences and the probability measure that represents the beliefs of the likelihood of events is that they exhibit Event
Independence. Hence probabilistically sophisticated preferences that satisfy either of the conditional monotonicity axioms P3\textsuperscript{CU} or P3\textsuperscript{CL} should also satisfy the following axiom.

\textbf{P4\textsuperscript{CE} (Strong Conditional Equivalent Probability):} For all pairs of disjoint events \(A\) and \(B\), outcomes \(w, x, y,\) and \(z\), and acts \(g\) and \(h\),

\[
\begin{bmatrix}
x \\
g(s) \end{bmatrix} \text{ if } s \in A \cup B \quad \succ \begin{bmatrix}
x \\
g(s) \end{bmatrix} \text{ if } s \notin A \cup B
\]

\begin{bmatrix}
y \\
g(s) \end{bmatrix} \text{ if } s \in A \cup B \quad \sim \begin{bmatrix}
y \\
x \\
g(s) \end{bmatrix} \text{ if } s \notin A \cup B
\end{bmatrix} \quad \succ \begin{bmatrix}
y \\
g(s) \end{bmatrix} \text{ if } s \notin A \cup B

implies

\begin{bmatrix}
w \\
z \\
h(s) \end{bmatrix} \text{ if } s \in A \cup B \quad \sim \begin{bmatrix}
z \\
w \\
h(s) \end{bmatrix} \text{ if } s \notin A \cup B
\end{bmatrix}

The final two axioms for the preferences over acts, ensure that the relation is not trivial and has the right continuity properties so that the beliefs of the likelihood of events can be represented by a probability measure.\textsuperscript{10}

\textbf{P5 (Nondegeneracy):} There exist acts \(f\) and \(g\) such that \(f \succ g\).

\textbf{P6\textsuperscript{†} (Small Event Continuity):} For any acts \(f \succ g\) and outcome \(x\), there exists a finite set of events \(\{A_1, \ldots, A_n\}\) forming a partition of \(\mathcal{A}\), such that for all \(i\), and for all \(B_i \subseteq A_i\),

\[
f \succ \begin{bmatrix}
x \\
g(s) \end{bmatrix} \text{ if } s \in A_i \cup B_i \quad \text{ and } \quad \begin{bmatrix}
x \\
f(s) \end{bmatrix} \text{ if } s \notin A_i \cup B_i
\]

\textbf{THEOREM I:} If the preference relation \(\succeq\) over \(\mathcal{A}\) satisfies the axioms: P1 (Ordering), P4\textsuperscript{CE} (Strong Conditional Equivalent Probability), P5 (Nondegeneracy), P6\textsuperscript{†} (Small Event Continuity), and one of P3\textsuperscript{CU} (Conditional Upper Monotonicity) or P3\textsuperscript{CL} (Conditional Lower Monotonicity), then \(\succeq\) is probabilistically sophisticated.

Note P6\textsuperscript{†} is a slight strengthening of Savage’s postulate P6. However, P6 in conjunction with P3 implies P6\textsuperscript{†} and the intuitive basis for the motivation of the two axioms is identical. Essentially it ensures the state space is sufficiently rich

\textsuperscript{10}The reader is referred to Savage (1972, Section 3.2) for a detailed discussion of these continuity issues.
to allow $\succeq$ to be "continuous" with respect to changes in acts on "sufficiently" small events. Moreover, in conjunction with the other axioms, it also implies that the induced risk preferences satisfy the following continuity property.

**Mixture Continuity for Two-Outcome Sublotteries:** For all pairs of outcomes $x$ any $y$ in $\mathcal{X}$, all $\gamma \in (0,1]$, and distributions $P$ and $R$ in $\mathcal{P}_0(\mathcal{X})$, the sets

$$A^{[\gamma](x,y):P}_{\sim R} = \{ \lambda \in [0,1] | \gamma(\lambda \delta_x + [1-\lambda] \delta_y) + (1-\gamma)P > R \} \quad \text{and}$$

$$A^{[\gamma](x,y):P}_{R \succ} = \{ \lambda \in [0,1] | R > \gamma(\lambda \delta_x + [1-\lambda] \delta_y) + (1-\gamma)P \}$$

are open.

As the following result demonstrates, this continuity property is sufficient for a complete, reflexive, and transitive binary relation over $\mathcal{P}_0(\mathcal{X})$ to have a utility representation that is mixture continuous for two outcome sublotteries utility.\textsuperscript{11,12}

**Theorem II:** The following conditions on a preference relation $\succeq$ over $\mathcal{A}$ are equivalent:

(i) $\succeq$ satisfies: P1 Ordering, P3\textsuperscript{CU} Conditional Upper Monotonicity (resp. P3\textsuperscript{CL} Conditional Lower Monotonicity), P4\textsuperscript{CE} Strong Conditional Equivalent Probability, P5 Nondegeneracy, P6\textsuperscript{S} Small Event Continuity.

(ii) There exists a unique finitely additive, nonatomic probability measure $\mu$ on $\mathcal{A}$ and a nonconstant, preference functional $V(P) \equiv V(x_1, p_1, \ldots, x_n, p_n)$ on $\mathcal{P}_0(\mathcal{X})$ which is mixture continuous for two-outcome sublotteries and exhibits Conditional Upward (respectively, Downward) Monotonicity, such that the relation $\succeq$ over acts can be represented by the preference functional

$$V(f) = V(x_1, \mu(f^{-1}(x_1)); \ldots; x_n, \mu(f^{-1}(x_n)))$$

where $\{x_1, \ldots, x_n\} = f(\mathcal{A})$.

Thus Theorem II provides formal characterizations of probabilistically sophisticated preferences, that include social orderings that embody an ex ante concern for fairness such as the Machina Mom example, as well as the observable preferences over acts of a Savage subjective expected utility maximizer undertaking an unobservable auxiliary action in a situation of delayed resolution of uncertainty.

\textsuperscript{11} $V: \mathcal{P}_0(\mathcal{X}) \to \mathbb{R}$ is said to be mixture continuous for two-outcome sublotteries, if for any pair of outcomes, $x$ and $y$, any $\gamma$ in $(0,1]$, and any pair of lotteries, $P$ and $R$, the sets $\{ \lambda \in [0,1] | V(\lambda \delta_x + [1-\lambda] \delta_y) + (1-\gamma)P > V(R) \}$ and $\{ \lambda \in [0,1] | V(R) > V(\lambda \delta_x + [1-\lambda] \delta_y) + (1-\gamma)P \}$ are open.

\textsuperscript{12} Moreover, mixture continuity for two outcome sublotteries is in a sense the weakest continuity property that guarantees that a preference ordering for simple lotteries over an arbitrary outcome set admits a functional representation.
5. WITH HOW MUCH MONOTONICITY MAY WE DISPENSE?

Just as the examples of Section 3 provided illustrations of induced risk preferences for probabilistically sophisticated individuals that need not satisfy ADI, one can imagine situations in which these induced risk preferences need satisfy neither of the conditional monotonicity properties implied by $P_3^{\text{CU}}$ and $P_3^{\text{CL}}$. For instance, suppose Machina's Mom also has available to her an unobservable choice of an auxiliary action between the time of choosing an act and the resolution of the uncertainty. If different choices for that auxiliary action affect the relative desirability for Mom of the two outcomes of either child receiving the treat, then her observable risk preferences may well exhibit neither conditional upward nor downward monotonicity.

Ideally then, it would be nice to characterize probabilistically sophisticated preferences without requiring the induced risk preferences to exhibit any specific properties save perhaps some form of continuity such as two-outcome mixture continuity that, as is demonstrated in the proof of Theorem II, admits a functional representation of risk preferences for simple lotteries over an arbitrary outcome set. I conjecture, however, that in a purely subjective uncertainty framework as employed in this paper, such a characterization may well prove impossible to attain. My reason stems from the critical role that monotonicity plays in allowing us to infer from the individual's preferences between certain pairs of acts his or her beliefs of the relative likelihood of events.

The method of proof in Theorem I consists of proving the following two statements:

1. There exists a non-null event $B$, and there exist outcomes $x$ and $y$ and an act $g$, such that conditional on $g(s)$ for $s \notin B$, the ordering satisfies the M-S (1992) axioms (in particular, eventwise monotonicity) on $B$ for acts $f$ with $f(s) \in \{x, y\}$, for all $s \in B$.

2. There exists a partition $\{D_1, \ldots, D_n\}$ of the complement of the set $B$ from 1, such that the M-S (1992) axioms can be extended from $B$ to any of the events $B \cup D_i$ for acts $f$ with $f(s) \in \{x, y\}$ for all $s \in B \cup D_i$, and $f(s) = g(s)$ otherwise.

Applying M-S (1992)'s result, the first statement implies that we can construct a (conditional) probability measure that represents the individual's conditional beliefs about the relative likelihoods of subsets of the event $B$. The second step enables us to extend this measure to the events $B \cup D_i$ from which we can "piece together" the unconditional probability measure that represents the individual's beliefs about the likelihood of every event.

Hence one possible direction to extend the characterization of probabilistically sophisticated preferences is to require eventwise monotonicity (and analogous local versions of the other M-S axioms) only to hold on a collection of events whose union is $\mathcal{F}$ and whose intersection is non-null.

Notice, however, that the two-outcome mixture continuity property exhibited by the induced risk preferences of a probabilistically sophisticated individual is not by itself sufficient to ensure that eventwise monotonicity of the underlying preferences over acts need hold "locally" anywhere for any pair of outcomes.
Consider, for example, an individual whose beliefs over the set of events can be represented by a nonatomic finitely additive probability measure, \( \mu \), and whose risk preferences for lotteries over a pair of outcomes are represented by a continuous function \( \nu : [0, 1] \to \mathbb{R} \), with \( \nu(0) = 0 \), \( \nu(1) = 1 \) and which is nowhere monotonic.\(^{13}\) Although the beliefs are probabilistically sophisticated, such erratic risk preferences seem to preclude the possibility of inferring the beliefs from the underlying act preferences. In particular, the inability to deduce from any observed preference patterns that the agent believes two events are equally likely precludes the formulation of an axiom necessary (and, in conjunction with the usual axioms of ordering, nondegeneracy and continuity, sufficient) to guarantee EI (Event Independence) and hence probabilistic sophistication.

Thus to sum up, unless someone can devise an alternative means to determine which set of events an individual views as equally likely that does not require even the most “conditional” and “local” form of eventwise monotonicity for the underlying preferences over acts, then a full characterization of probabilistically sophisticated preferences that induce (two-outcome) mixture continuous risk preferences appears unobtainable.

However, as I demonstrated in the previous section, for particular examples of nonatomic risk preferences that are economically relevant, “conditional” or “local” monotonicity may still obtain, enabling one to adapt the betting preference procedure to elicit beliefs and define the conditions that are both necessary and sufficient to ensure that such beliefs accord with probabilistically sophisticated behavior.

Economics Program, Research School of Social Sciences, The Australian National University, Canberra, ACT 0200, Australia

Manuscript received May, 1991; final revision received February, 1994.

APPENDIX

**Proof of Proposition 2.1:** Choose a non-null event \( E \) (i.e. \( \mu(E) > 0 \)), two outcomes \( x \) and \( y \), and an act \( h \). By probabilistic sophistication we have:

(i) \( x \geq y \iff \delta_x \succeq_p \delta_y \), and

(ii) \[
g_1 = \begin{bmatrix}
  x & \text{if } s \in E \\
  h(s) & \text{if } s \not\in E
\end{bmatrix} \succeq \begin{bmatrix}
  y & \text{if } s \in E \\
  h(s) & \text{if } s \not\in E
\end{bmatrix} = g_2
\]

\(\iff \mu \circ g_1^{-1} = \mu(E)\delta_x + [1 - \mu(E)]H \succeq_p \mu(E)\delta_y + [1 - \mu(E)]H = \mu \circ g_2^{-1}\)

where \( H = (x_1, p_1, \ldots, x_n, p_n) \) with \( p_i = [1 - \mu(E)]^{-1}\mu(h^{-1}(x_i) \setminus E) \). Hence ADI implies P3 (EM).

Choose now a simple lottery \( H = (x_1, p_1, \ldots, x_n, p_n) \). As the mapping from finite acts to simple lotteries using \( \mu \) is onto, there exists an act \( h \) such that \( p_i = [1 - \mu(E)]^{-1}\mu(h^{-1}(x_i) \setminus E) \) for each \( i \). Hence from (i) and (ii) it now follows that P3(EM) implies ADI.

Q.E.D.

**Proof of Proposition 4.1:** Follows from an analogous construction of acts to the one employed in the proof of Proposition 2.1 above.

\(^{13}\) For an example of such a function see, for instance, Gelbaum and Olmsted (1964, p. 29).
SUBJECTIVE PROBABILITY

Proof of Proposition 4.2: I shall prove the implication for P3\text{CL} (an analogous argument can be employed to prove the implication for P3\text{CU}). It follows trivially if $B \setminus C$ is null. Fix non-null $C \subset B$ (with $B \setminus C$ also non-null) and act $f$ with $f(x') \leq (x, y)$. Let $D' = f^{-1}(y) \cap B \setminus C$ and $D'' = f^{-1}(x) \cap B \setminus C$. Thus,

\[
\begin{bmatrix}
    x & \text{if } s \in C \\
    f(s) & \text{if } s \in B \setminus C \\
    x & \text{if } s \in A \\
    g(s) & \text{if } s \notin A \cup B
\end{bmatrix}
\begin{bmatrix}
    x & \text{if } s \in C \\
    y & \text{if } s \in D' \\
    x & \text{if } s \in A \cup D'' \\
    g(s) & \text{if } s \notin A \cup B
\end{bmatrix}
\]

and

\[
\begin{bmatrix}
    y & \text{if } s \in C \\
    f(s) & \text{if } s \in B \setminus C \\
    x & \text{if } s \in A \\
    g(s) & \text{if } s \notin A \cup B
\end{bmatrix}
\begin{bmatrix}
    y & \text{if } s \in C \\
    y & \text{if } s \in D' \\
    x & \text{if } s \in A \cup D'' \\
    g(s) & \text{if } s \notin A \cup B
\end{bmatrix}
\]

As

\[
\begin{bmatrix}
    x & \text{if } s \in A \\
    y & \text{if } s \in B \\
    g(s) & \text{if } s \notin A \cup B
\end{bmatrix}
\begin{bmatrix}
    y & \text{if } s \in A \\
    g(s) & \text{if } s \notin A \cup B
\end{bmatrix}
\]

it follows from P3\text{CL} that

\[
\begin{bmatrix}
    x & \text{if } s \in C \\
    y & \text{if } s \in D' \\
    x & \text{if } s \in A \cup D'' \\
    g(s) & \text{if } s \notin A \cup B
\end{bmatrix}
\begin{bmatrix}
    x & \text{if } s \in A \\
    y & \text{if } s \in B \\
    g(s) & \text{if } s \notin A \cup B
\end{bmatrix}
\]

and

\[
\begin{bmatrix}
    y & \text{if } s \in C \\
    y & \text{if } s \in D' \\
    x & \text{if } s \in A \cup D'' \\
    g(s) & \text{if } s \notin A \cup B
\end{bmatrix}
\begin{bmatrix}
    x & \text{if } s \in A \\
    y & \text{if } s \in B \\
    g(s) & \text{if } s \notin A \cup B
\end{bmatrix}
\]

Hence again by P3\text{CL}

\[
\begin{bmatrix}
    x & \text{if } s \in C \\
    y & \text{if } s \in D' \\
    x & \text{if } s \in A \cup D'' \\
    g(s) & \text{if } s \notin A \cup B
\end{bmatrix}
\begin{bmatrix}
    y & \text{if } s \in C \\
    y & \text{if } s \in D' \\
    x & \text{if } s \in A \cup D'' \\
    g(s) & \text{if } s \notin A \cup B
\end{bmatrix}
\]

Q.E.D.

Proof of Theorem 1: The theorem is shown for the case where P3\text{CL} (Conditional Lower Eventwise Monotonicity) holds. A similar construction can be used where P3\text{CU} (Conditional Upper Eventwise Monotonicity) holds instead. The basic idea of the proof is to show that there exists a set of acts which differ only on one event, for which the hypothesis of M-S (1992)'s Theorem 1 holds on that event. This allows us to infer the existence of a probability measure that represents the individual's beliefs over the relative likelihoods of subevents of that event. Step 4 shows how the continuity of the individual's preferences over acts implied by P6 is enabled for us to extend this measure in a consistent way to the whole state space. The final step then shows how the individual is indifferent between acts which are mapped to the same lottery using this measure and thus the induced lottery preferences are consistent with the underlying preferences over acts.

Step 1:

Claim 1: There exist outcomes $x$ and $y$, event $E$, and act $g$ such that

\[
\begin{bmatrix}
    x & \text{if } s \in E \\
    g(s) & \text{if } s \notin E
\end{bmatrix}
> \begin{bmatrix}
    y & \text{if } s \in E \\
    g(s) & \text{if } s \notin E
\end{bmatrix}
\]

**Proof:** Assume not. That is for every pair of outcomes \( x' \) and \( y' \), event \( E' \) and act \( g' \) \( [x' \text{ if } E; g'(s) \text{ if } E'] \sim [y' \text{ if } E; g(s) \text{ if } E] \). Hence for any \( f \) and \( h \), with \( f(S) = (x_1, \ldots, x_m) \) and \( h(S) = (y_1, \ldots, y_n) \), let \( (C_1, \ldots, C_m) \) and \( (D_1, \ldots, D_n) \) be partitions of \( S \) with \( C_i = f^{-1}(x_i) \) and \( D_j = h^{-1}(y_j) \). Let \( (E_1, \ldots, E_p) \) be the coarsest common refinement of \( (C_1, \ldots, C_m) \) and \( (D_1, \ldots, D_n) \). Then

\[
\begin{align*}
  f \sim & \begin{bmatrix} h(s) & \text{ if } s \in E_1 \\ f(s) & \text{ if } s \not\in S \setminus E_1 \end{bmatrix} \sim \begin{bmatrix} h(s) & \text{ if } s \in E_1 \cup E_2 \\ f(s) & \text{ if } s \not\in (E_1 \cup E_2) \end{bmatrix} \\
  \sim & \cdots \sim \begin{bmatrix} h(s) & \text{ if } s \in S \setminus E_p \\ f(s) & \text{ if } s \in E_p \end{bmatrix} \sim h.
\end{align*}
\]

A contradiction of P5 (Nondegeneracy).

Q.E.D.

**Step 2:**

**Step 2a:**

**Claim 2a:** For the pair of outcomes, \( x \) and \( y \), event \( E \) and act \( g \) from Step 1 above, there exists two disjoint non-null events \( A \) and \( B, A \cup B = E \) such that

\[
\begin{bmatrix}
  x & \text{ if } s \in A \\
  x & \text{ if } s \in B \\
  g(s) & \text{ if } s \not\in A \cup B
\end{bmatrix} \succ
\begin{bmatrix}
  x & \text{ if } s \in A \\
  y & \text{ if } s \in B \\
  g(s) & \text{ if } s \not\in A \cup B
\end{bmatrix} \succ
\begin{bmatrix}
  y & \text{ if } s \in A \\
  y & \text{ if } s \in B \\
  g(s) & \text{ if } s \not\in A \cup B
\end{bmatrix}.
\]

**Proof:** By P6\(^1\) (Small Event Continuity) there exists a finite partition of \( S \), \( (C_1, \ldots, C_n) \) such that for all \( i \), and all \( B_i \subseteq C_i \)

\[
\begin{bmatrix}
  x & \text{ if } s \in E \setminus B_i \\
  y & \text{ if } s \in B_i \\
  g(s) & \text{ if } s \not\in E \cup B_i
\end{bmatrix} \succ
\begin{bmatrix}
  y & \text{ if } s \in E \\
  g(s) & \text{ if } s \not\in E
\end{bmatrix}.
\]

As \( (C_1, \ldots, C_n) \) is a finite partition of \( S \), at least one element, say \( C_i \), has a non-null intersection with \( E \). Hence, let \( B = B_i \subseteq C_i \cap E \) and \( A = E \setminus B \). First preference relation then follows directly from P3\(^{cl}\).

Q.E.D.

**Step 2b:**

For the pair of outcomes, \( x \) and \( y \), the non-null and disjoint pair of events \( A \) and \( B \), and act \( g \), from Step 2a above it follows from Proposition 4.2 that for all non-null events \( C \subseteq B \), and acts \( f \) with \( f(S) \subseteq (x, y) \),

\[
\begin{bmatrix}
  x & \text{ if } s \in C \\
  f(s) & \text{ if } s \in B \setminus C \\
  x & \text{ if } s \in A \\
  g(s) & \text{ if } s \not\in A \cup B
\end{bmatrix} \succ
\begin{bmatrix}
  y & \text{ if } s \in C \\
  f(s) & \text{ if } s \in B \setminus C \\
  y & \text{ if } s \in A \\
  g(s) & \text{ if } s \not\in A \cup B
\end{bmatrix},
\]

that is, conditional eventwise monotonicity on \( B \) for \( x \) and \( y \), given \( x \) on \( A \) and \( g \) on \( S \setminus E \).
CLAIM 2b: For the pair of outcomes, \( x \) and \( y \), non-null and disjoint pair of events \( A \) and \( B \), and act \( g \), from Step 2a it follows that for all pairs of disjoint events \( C, D \subseteq B \) and acts \( f \) and \( h \), with \( f(\mathcal{S}), h(\mathcal{S}) \subseteq (x, y) \),

\[
\begin{bmatrix}
x & \text{if } C \\
y & \text{if } D \\
f(s) & \text{if } B \setminus (C \cup D) \\
x & \text{if } A \\
g(s) & \text{if } E^c
\end{bmatrix}
\approx
\begin{bmatrix}
y & \text{if } C \\
x & \text{if } D \\
f(s) & \text{if } B \setminus (C \cup D) \\
x & \text{if } A_c \\
g(s) & \text{if } E^c
\end{bmatrix}
\]

\[
\begin{bmatrix}
x & \text{if } C \\
y & \text{if } D \\
h(s) & \text{if } B \setminus (C \cup D) \\
x & \text{if } A \\
g(s) & \text{if } E^c
\end{bmatrix}
\Rightarrow
\begin{bmatrix}
y & \text{if } C \\
x & \text{if } D \\
h(s) & \text{if } B \setminus (C \cup D) \\
x & \text{if } A_c \\
g(s) & \text{if } E^c
\end{bmatrix}
\]

(i.e., Conditional Strong Comparative Probability on \( B \) for acts \( f \) with \( f(\mathcal{S}) \subseteq (x, y) \), given \( x \) on \( A \) and \( g \) on \( \mathcal{S} \setminus E \).

PROOF: (Note, to aid in following the reasoning of this and subsequent proofs, the changes in acts from one line of a proof to the next are highlighted in bold face.) By P6 and Proposition 4.2, there exists an event \( \hat{C} \subset C \) such that

\[
\begin{bmatrix}
x & \text{if } \hat{C} \\
y & \text{if } C \setminus \hat{C} \\
x & \text{if } D \\
f(s) & \text{if } B \setminus (C \cup D) \\
x & \text{if } A \\
g(s) & \text{if } E^c
\end{bmatrix}
\approx
\begin{bmatrix}
x & \text{if } \hat{C} \\
y & \text{if } C \setminus \hat{C} \\
x & \text{if } D \\
f(s) & \text{if } B \setminus (C \cup D) \\
x & \text{if } A \\
g(s) & \text{if } E^c
\end{bmatrix}
\]

Thus by P4CE (Strong Conditional Equivalent Probability),

\[
\begin{bmatrix}
x & \text{if } \hat{C} \\
y & \text{if } C \setminus \hat{C} \\
y & \text{if } D \\
h(s) & \text{if } B \setminus (C \cup D) \\
x & \text{if } A \\
g(s) & \text{if } E^c
\end{bmatrix}
\approx
\begin{bmatrix}
y & \text{if } \hat{C} \\
x & \text{if } C \setminus \hat{C} \\
x & \text{if } D \\
h(s) & \text{if } B \setminus (C \cup D) \\
x & \text{if } A \\
g(s) & \text{if } E^c
\end{bmatrix}
\]

and applying Proposition 4.2 again,

\[
\begin{bmatrix}
x & \text{if } \hat{C} \\
x & \text{if } C \setminus \hat{C} \\
y & \text{if } D \\
h(s) & \text{if } B \setminus (C \cup D) \\
x & \text{if } A \\
g(s) & \text{if } E^c
\end{bmatrix}
\approx
\begin{bmatrix}
y & \text{if } \hat{C} \\
y & \text{if } C \setminus \hat{C} \\
x & \text{if } D \\
h(s) & \text{if } B \setminus (C \cup D) \\
x & \text{if } A \\
g(s) & \text{if } E^c
\end{bmatrix}
\quad Q.E.D.
\]
Step 3: Following Steps 1–3 of M-S (1992)'s proof of their Theorem 1 (pp. 771–772), consider $\succeq^s$, defined as $\succeq$ restricted to $\mathcal{A}^s = \{ f \in \mathcal{A} | (f(B) \subseteq (x, y) : f(A) = x; f(s) = g(s) if s \notin E)$. In conjunction with P1, P5, and P6$^1$ holding over $\succeq^s$, Proposition 4.2 and Claim 2b imply that Savage's postulates hold in this setting. Hence we can construct $\succeq^p$ for all pairs of subsets of $B$ from "betting preferences" of obtaining $x$ or $y$ on subsets of $B$ conditional on $x$ occurring if $A$ obtains and act $g$ determining the outcome if a state not in $A \cup B$ obtains. That is, for all $C, D \subseteq B$, $C \succeq^p D$ if and only if

$$
\begin{align*}
\begin{bmatrix}
 x & if s \in C \\
 y & if s \in B \setminus C \\
x & if s \in A \\
g(s) & if s \notin E
\end{bmatrix}
& \quad \succ
\begin{bmatrix}
 x & if s \in D \\
 y & if s \in B \setminus D \\
x & if s \in A \\
g(s) & if s \notin E
\end{bmatrix}
\end{align*}
$$

By Savage's Theorem, there accordingly exists a unique, finitely additive, nonatomic probability measure over subsets of $B, \mu_B$, that represents these beliefs conditioned on $B$.

Step 4:

CLAIM 4a: For the pair of outcomes, $x$ and $y$, the non-null and disjoint pair of events $A$ and $B$, and the act $g$ from Step 2a, there exists a finite partition of $B \setminus \{D_1, \ldots, D_n\}$, such that for all $i$: either $D_i \subseteq A$ or $g(D_i) = \{z\}$ where $z \in g(\mathcal{A} \setminus A \cup B)$, and for all $D_i' \subseteq D_i$,

$$
\begin{align*}
\begin{bmatrix}
 x & if s \in B \cup D_i' \\
 x & if s \in B \setminus D_i' \\
g(s) & if s \notin A \cup B \setminus D_i'
\end{bmatrix}
& \quad \succ
\begin{bmatrix}
 x & if s \in B \cup D_i' \\
 x & if s \in B \setminus D_i' \\
g(s) & if s \notin A \cup B \setminus D_i'
\end{bmatrix}
\end{align*}
$$

PROOF: From Claim 2a it follows that there exists four non-null events $A_1, A_2, B_1,$ and $B_2$, with $A_1 \cup A_2 = A$ and $B_1 \cup B_2 = B$, such that

$$
\begin{align*}
 h_1 =
\begin{bmatrix}
 x & if A_1 \\
 x & if A_2 \\
x & if B_1 \\
x & if B_2 \\
g(s) & if \mathcal{A} \setminus A \cup B
\end{bmatrix}
& \quad \succ
\begin{bmatrix}
 x & if A_1 \\
 x & if A_2 \\
x & if B_1 \\
x & if B_2 \\
g(s) & if \mathcal{A} \setminus A \cup B
\end{bmatrix}
\end{align*}
$$

$$
\begin{align*}
 h_2 =
\begin{bmatrix}
 y & if B_1 \\
y & if B_2 \\
g(s) & if \mathcal{A} \setminus A \cup B
\end{bmatrix}
& \quad \succ
\begin{bmatrix}
 y & if B_1 \\
y & if B_2 \\
g(s) & if \mathcal{A} \setminus A \cup B
\end{bmatrix}
\end{align*}
$$

Let $(G_1, \ldots, G_m)$ be the finite partition over $\mathcal{A} \setminus A \cup B$ induced by $g$. That is, for all $G_i, g(G_i) = \{z\}$ where $z \in g(\mathcal{A} \setminus A \cup B)$, Applying P6$^1$ (Small Event Continuity) to pairs of acts: (a) $h_1$ and $h_2$ for outcome $x$; (b) $h_2$ and $h_3$ for outcome $y$; (c) $h_3$ and $h_4$ for outcome $y$; and (d) $h_4$ and $h_5$ for outcome $y$, and then taking the coarsest common refinement of the four partitions obtained with $(A, B, G_1, \ldots, G_m)$, we obtain the finite partition $(C_1, \ldots, C_n)$. Defining $D_i = C_i \cap \mathcal{A} \setminus B$, provides us
with a partition of $\mathcal{A} \setminus B$ such that for all $i$, and all $D_i \subseteq D_i$ we have

\[
\begin{array}{c}
x \quad \text{if } B \cup D_i \\
x \quad \text{if } A \setminus D_i \\
g(s) \quad \text{if } \mathcal{A} \setminus A \cup B \cup D_i \\
\end{array} \quad \succ h_2 \succ \\
\begin{array}{c}
y \quad \text{if } B \cup D_i' \\
x \quad \text{if } A \setminus D_i' \\
g(s) \quad \text{if } \mathcal{A} \setminus A \cup B \cup D_i' \\
\end{array} \quad \succ h_4 \succ \\
\begin{array}{c}
y \quad \text{if } B \cup D_i' \\
y \quad \text{if } A \cup D_i' \\
g(s) \quad \text{if } \mathcal{A} \setminus A \cup B \cup D_i' \\
\end{array}
\]

Q.E.D.

Hence by applying the results of Steps 2 and 3 we can obtain a relation $\geq_{\mu_B}^{BuDi}$ over all subsets of $B \cup D$, which is uniquely represented by the finitely additive probability measure, $\mu_{BuDi}$. Moreover, $\geq_{\mu_B}^{BuDi}$ and $\geq_B$ agree over $B$. As both relations (weakly) agree with the partial ordering of set inclusion, it suffices to show that $\geq_{BuDi}$ and $\geq_B$ agree over $B$.

CLAIM 4b: For all non-null events $C, F \subseteq B$, $C \sim_{BuDi}^{BuDi} F \Rightarrow C \sim_{B} F$.

PROOF: As $\mu_B$ (respectively, $\mu_{BuDi}$) is additive, $C \sim_{BuDi}^{BuDi} F \Rightarrow C \setminus F \sim_{BuDi}^{BuDi} F \setminus C$. But from the conditional eventwise monotonicity over $B$ (respectively, $B \cup D$) that follows from Proposition 4.2,

\[
C \setminus F \sim_{B} F \setminus C \Leftrightarrow
\begin{array}{c}
x \quad \text{if } C \setminus F \\
y \quad \text{if } C \cap F \\
x \quad \text{if } F \setminus C \\
y \quad \text{if } B \setminus (C \cup F) \\
x \quad \text{if } A \setminus D \\
g(s) \quad \text{if } (E \cup D_i)^c
\end{array} \quad \succ
\begin{array}{c}
\text{if } C \setminus F \\
\text{if } C \cap F \\
\text{if } F \setminus C \\
\text{if } B \setminus (C \cup F) \\
\text{if } A \setminus D \\
g(s) \quad \text{if } (E \cup D_i)^c
\end{array}
\]

Q.E.D.
Note from the last two-way implication of the above proof and \( \text{P4}^{\text{CE}} \) we have the following claim.

**CLAIM 4c:** For any pair of disjoint events \( F \) and \( G \) in \( B \cup D_1 \), any outcomes \( w, z, \) and act \( h \):

\[
F \sim_{I}^{\text{BuDi}} G \Rightarrow \left[ \begin{array}{l}
\text{w if } F; \text{ z if } G; \text{ h(s) if } (F \cup G)\wedge \nonumber \\
\text{z if } F; \text{ w if } G; \text{ h(s) if } (F \cup G)\wedge 
\end{array} \right]
\sim \left[ \begin{array}{l}
\text{z if } F; \text{ w if } G; \text{ h(s) if } (F \cup G)\wedge 
\end{array} \right].
\]

As \( (B, D_1, \ldots, D_n) \) forms a partition of \( \mathcal{A} \), we can construct \( \succeq_I \) for all events from the collection of conditional beliefs \( (\succeq_{I}^{\text{BuDi}}) \), \( \succeq_I \) inherits the properties of each \( \succeq_{I}^{\text{BuDi}} \), and thus can be represented by a unique finitely additive probability measure, \( \mu \).

**Step 5:**

**CLAIM 5a:** For all pairs of disjoint events, \( F \) and \( G \), all pairs of outcomes \( w \) and \( z \), and acts \( h \):

\[
\mu(F) = \mu(G) \quad \text{implies} \quad \left[ \begin{array}{l}
w \quad \text{if } s \in F \\
z \quad \text{if } s \in G \\
h(s) \quad \text{if } s \notin F \cup G 
\end{array} \right] \sim \left[ \begin{array}{l}
z \quad \text{if } s \in F \\
w \quad \text{if } s \in G \\
h(s) \quad \text{if } s \notin F \cup G 
\end{array} \right].
\]

**PROOF:** Let \( \{H_1, \ldots, H_m\} \) be the finite partition over \( \mathcal{A} \setminus F \cup G \) induced by \( h \). Consider the partition \( \{E_1, \ldots, E_m\} \) of \( \mathcal{A} \), which is the coarsest common refinement of \( (F, G, H_1, \ldots, H_m) \) and the partition \( \{B_1, D_1, \ldots, D_n\} \) from Step 4. As \( \mu \) is nonatomic, there exist equi-numbered finite partitions of \( F \) and \( G \), \( (F_1, \ldots, F_m) \) and \( (G_1, \ldots, G_m) \) such that for each \( k \):

(i) \( \mu(F_k) = \mu(G_k) < \mu(E^B) \) where \( E^B \) is in \( \{E_1, \ldots, E_m\} \) and \( E^B \subseteq B \).

(ii) \( F_k \subseteq E_j \) for some \( E_j \) in \( \{E_1, \ldots, E_m\} \).

(iii) \( G_k \subseteq E_j \), for some \( E_j \) in \( \{E_1, \ldots, E_m\} \).

Hence for each \( k \), there exists a \( B' \in E^B \) such that \( F_k \sim_{I}^{\text{BuDi}} B' \) for some \( D_i \) and \( G_k \sim_{I}^{\text{BuDi}} B' \) for some \( D_i \). By swapping the outcomes on \( F_k \) and \( B' \) followed by swapping the outcomes on \( B' \) and \( G_k \) and then swapping the outcomes on \( B' \) and \( F_k \) again, it follows from the thrice application of Claim 4c that indifference is preserved. The conclusion thus follows by employing this procedure for each \( k \).

**Q.E.D.**

**COROLLARY 5b:** For any two acts \( f \) and \( g \) in \( \mathcal{A} \), \( \mu \circ f^{-1} = \mu \circ g^{-1} \) implies \( f \sim g \).

**PROOF:** Since Claim 5a says exchanging outcomes on equally likely events leaves the individual indifferent, we can utilize M-S’s (1991) construction in Step 5 of the proof of their Theorem 1. That is, we can construct a finite sequence of acts \( f = f_0 \), \( f_1, \ldots, f_n = g \) such that \( f_{i+1} \) is obtained from \( f_i \) by exchanging the outcomes on a pair of equally likely events, which implies \( f = f_0 \sim f_1 \sim \cdots \sim f_n = g \).

**Q.E.D.**

**CLAIM 5c:** \( \succeq \) is probabilistically sophisticated with respect to \( \mu \).

**PROOF:** From the definition of \( \succeq_p \), for all \( \mu \circ f^{-1} \succeq_p \mu \circ g^{-1} \) there exist acts \( f' \) and \( g' \) such that \( \mu \circ f'^{-1} = \mu \circ f^{-1} \), \( \mu \circ g'^{-1} = \mu \circ g^{-1} \), and \( f' \succeq g' \). By Claim 5b, \( f \sim f' \) and \( g \sim g' \), hence by P1, \( f \succeq g \).

**Q.E.D.**

**PROOF OF THEOREM II:** (i) \( \Rightarrow \) (ii).

**Step 1:** Let \( \succeq_p \) be the preference relation over \( \mathcal{A} \) induced from \( \succeq \) using the probability measure \( \mu \), defined in the proof of Theorem I. As \( \mu \) is nonatomic, the mapping from acts to simple lotteries is onto. Hence P1 implies that \( \succeq_p \) is a complete, reflexive, and transitive relation.

**Step 2:**

**CLAIM 2:** \( \succeq_p \) satisfies Conditional Upper (resp. Lower) Monotonicity.
SUBJECTIVE PROBABILITY

Proof: Fix \(x, y\) in \(\mathcal{X}\), \(P\) in \(\mathcal{P}(\mathcal{X})\), and \(\lambda, \gamma\) in \((0, 1)\). Since mapping of acts in \(\mathcal{A}\), using \(\mu\), is onto \(\mathcal{P}(\mathcal{X})\), there exists disjoint events \(A\) and \(B\) and act \(h\) such that \(\mu(A) = \lambda\gamma, \mu(B) = \lambda(1 - \gamma)\), and \(\mu + h^{-1}\) conditional on \(s\) in \(A \cup B\) is \(P\). From definition of \(\succeq_P\) and fact that \(\succeq\) is probabilistically sophisticated (Theorem 1), we have

\[
\begin{bmatrix}
  x & \text{if } A \cup B \\
  y & \text{if } (A \cup B)^c
\end{bmatrix}
\succeq_P
\begin{bmatrix}
  y & \text{if } A \cup B \\
  h(s) & \text{if } (A \cup B)^c
\end{bmatrix}
\Rightarrow
(1 - \lambda)P \succ_P (\lambda\gamma\delta_x + (1 - \lambda)P;
\begin{bmatrix}
  x & \text{if } A \\
  y & \text{if } B
\end{bmatrix}
\succeq_P
\begin{bmatrix}
  y & \text{if } A \\
  h(s) & \text{if } (A \cup B)^c
\end{bmatrix}
\Rightarrow
\lambda\gamma\delta_x + (1 - \lambda)P \succ_P (\lambda\delta_y + (1 - \lambda)P;
\begin{bmatrix}
  x & \text{if } A \\
  y & \text{if } B
\end{bmatrix}
\succeq_P
\begin{bmatrix}
  x & \text{if } A \\
  y & \text{if } B
\end{bmatrix}
\Rightarrow
(1 - \lambda)P \succ_P (\lambda\gamma\delta_x + (1 - \lambda)P.

Hence \(P^3_{CU}\) (resp. \(P^3_{CL}\)) implies Conditional Upward (resp. Lower) Monotonicity. Q.E.D.

Step 3:

Claim \((\succeq_P\) is Mixture Continuous for Two-Outcome Sublotteries: For all \(\gamma \in (0, 1)\), outcomes \(x, y\) in \(\mathcal{X}\), lotteries \(P, R\) in \(\mathcal{P}(\mathcal{X})\), the sets

\[
A_{[\gamma; \gamma + \epsilon)]R} = \{ \lambda \in [0, 1] | \gamma(\lambda\delta_x + [1 - \lambda]\delta_y) + (1 - \gamma) \succ R \}
\]

\[A_{[\gamma, \gamma + \epsilon)]R} = \{ \lambda \in [0, 1] | R \succ \gamma(\lambda\delta_x + [1 - \lambda]\delta_y) + (1 - \gamma) \}
\]

are open.

Proof: Fix \(\gamma, x, y, P, R\) and let \(\lambda\) be an arbitrary element of \(A_{[\gamma; \gamma + \epsilon)]R}\). It suffices to show that there exists some \(\epsilon > 0\) such that \(\lambda\) in \((\lambda^x - \epsilon, \lambda^x + \epsilon) \cap [0, 1]\) implies \((\lambda\delta_x + [1 - \lambda]\delta_y) + (1 - \gamma)P \succ R\). From the definition of \(\succeq_P\) and the fact that \(\succeq\) is probabilistically sophisticated, there exist acts \(f\) and \(g\) in \(\mathcal{A}\), which imply distributions \(\gamma(\lambda\delta_x + [1 - \lambda]\delta_y) + (1 - \gamma)P\) and \(R\), respectively, and with \(f \succ g\). Axiom \(P^6\) implies that there exists a partition \(\{A_i^x, \ldots, A^n_x\}\) of \(\mathcal{X}\) such that for all \(i = 1, \ldots, n_x\) and all \(B_i^x \subseteq A_i^x\),

\[
\begin{bmatrix}
  x & \text{if } B_i^x \cap f^{-1}(y) \\
  f(s) & \text{elsewhere}
\end{bmatrix}
\succ g;
\]

and a partition \(\{A_j^y, \ldots, A^n_y\}\) of \(\mathcal{X}\) such that for all \(j = 1, \ldots, n_y\) and all \(B_j^y \subseteq A_j^y\),

\[
\begin{bmatrix}
  y & \text{if } B_j^y \cap f^{-1}(x) \\
  f(s) & \text{elsewhere}
\end{bmatrix}
\succ g.
Let
\[ e^x = \max_i \mu \left( A_i^x \cap f^{-1}(y) \right) , \quad e^y = \max_j \mu \left( A_j^y \cap f^{-1}(x) \right) \]
and
\[ e = \begin{cases} \min(e^x, e^y) & \text{if } \lambda \lambda^* (1 - \lambda^*) > 0, \\ e^x & \text{if } \lambda^* < 1, \\ e^y & \text{otherwise.} \end{cases} \]
Thus for any \( \lambda^- \) in \((\lambda^* - e, \lambda^*) \cap [0,1]\) the act
\[ f^\lambda = \begin{cases} x \text{ if } B^x & \\ f(s) \text{ elsewhere} & \end{cases} \]
(where \( B^x \subset A_i^x \cap f^{-1}(y) \) for some \( i \in \{1, \ldots, n_i\} \) with \( \mu(B^x) = \lambda^- - \lambda^- \) is preferred to the act \( g \) and implies the distribution \( \lambda(1 - \delta_x + (1 - \lambda^-) \delta_y) + (1 - \gamma)P \). Similarly, for any \( \lambda^+ \) in \((\lambda^*, \lambda^* + e) \cap [0,1]\) the act
\[ f^\lambda = \begin{cases} y \text{ if } B^y & \\ f(s) \text{ elsewhere} & \end{cases} \]
(where \( B^y \subset A_j^y \cap f^{-1}(x) \) for some \( j \in \{1, \ldots, n_j\} \) with \( \mu(B^y) = \lambda^+ - \lambda^- \) is preferred to the act \( g \) and implies the distribution \( \lambda(\gamma \delta_y + (1 - \lambda^+) \delta_x) + (1 - \gamma)P \). A similar argument can be employed to show that \( A_i^Q \neq \emptyset \).

**Q.E.D.**

**Step 4** (Existence of \( V \cdot \) on Set \((P \in \mathcal{P}_Q(\mathcal{X})| P \neq P \) ). It will suffice to show that there exists a countable subset \( \mathcal{P}_Q(\mathcal{X}) \subseteq \mathcal{P}_Q(\mathcal{X}) \), such that for all \( Q \) and \( R \) in \((P \in \mathcal{P}_Q(\mathcal{X})| P \neq P \) with \( R \neq Q \), there exists some \( Z \) in \( \mathcal{P}_Q(\mathcal{X}) \) with \( Q \subseteq P \) \( P \subseteq R \); i.e., \( \mathcal{P}_Q(\mathcal{X}) \cap (P \in \mathcal{P}_Q(\mathcal{X})| P \neq P \) is a countable \( \neq P \) dense subset of \((P \in \mathcal{P}_Q(\mathcal{X})| P \neq P \) (see Kreps (1988, Theorem 3.5, p. 25) or Debreu (1964, Proposition 5, p. 291)). Moreover, as \((P \in \mathcal{P}_Q(\mathcal{X})| P \neq P \) has a best and worst lottery, we can define for any pair of real numbers \( a < b \) a function \( V : (P \in \mathcal{P}_Q(\mathcal{X})| P \neq P \) \( P \neq P \rightarrow [a,b] \) that represents \( \mathcal{P}_Q(\mathcal{X}) \) restricted to \((P \in \mathcal{P}_Q(\mathcal{X})| P \neq P \) with \( V(P) = a \) and \( V(P) = b \).

Let \( f \) (resp. \( g \)) be an act that implies \( \tilde{P} \) (resp. \( P \) ). As \( P \) and \( \tilde{P} \) are simple lotteries, \( \tilde{P} \) can be obtained from \( P \) by a finite number of operations of moving a probability mass from an outcome in the range of \( f \) to an outcome in the range of \( g \). That is, there exist two finite sequences of lotteries, \( P_0, P_1, \ldots, P_n = \tilde{P} \) and \( T_1, \ldots, T_n \), where for all \( i = 1, \ldots, n \), \( P_i = \gamma \delta_{x_i} + (1 - \gamma)T_i \) \( P_{i-1} = \gamma \delta_{x_i} + (1 - \gamma)T_i \), with \( x_i \) in \( f(\mathcal{F}) \) and \( \gamma_i \) in \( g(\mathcal{G}) \).

Let \( Q \) denote the set of rational numbers. For each \( i = 1, \ldots, n \), define
\[ \mathcal{G}_i = \left\{ P \in \mathcal{P}_Q(\mathcal{X}) | P \neq P \ \text{and} \ \gamma_i \delta_{x_i} + (1 - \gamma_i)T_i \right\}, \]
\[ \mathcal{P}_i = \left\{ P \in \mathcal{P}_Q(\mathcal{X}) | P \neq P \ \text{and} \ \gamma_i \delta_{x_i} + (1 - \gamma_i)T_i \right\}, \]
\[ \mathcal{G} = \bigcup_{i=1}^n \mathcal{G}_i \text{ and } \mathcal{P} = \bigcup_{i=1}^n \mathcal{P}_i. \]

**Claim 4:** \((P \in \mathcal{P}_Q(\mathcal{X})| P \neq P \) \( \subseteq \bigcup_{i=1}^n \mathcal{G}_i). \)

**Proof:** Suppose not, i.e., there exists some \( Q \) in \((P \in \mathcal{P}_Q(\mathcal{X})| P \neq P \) but \( Q \) is not in \( \bigcup_{i=1}^n \mathcal{G}_i \). Thus for each \( i \), we have \( P \in \mathcal{P}_Q(\mathcal{X})| P \neq P \) \( \cap \mathcal{G}_i = \emptyset \). But as they have an empty intersection and are both open by mixture continuity for two-outcome sublotteries, one must be empty and the other equal to \( \{0,1\} \). For each \( i = 1, \ldots, n \), \( P_i = \gamma_i \delta_{x_i} + (1 - \gamma_i)T_i = \gamma_i \delta_{x_i} + (1 - \gamma_i)T_{i+1} \), thus, it follows for all \( i \),
\[ A_{\mathcal{G}_i} = A_{\mathcal{G}_{i+1}} \in [\gamma_{i+1};(\gamma_{i+1} + 1);T_{i+1}] \quad \text{and} \quad A_{\mathcal{P}_i} = A_{\mathcal{P}_{i+1}} \in [\gamma_{i+1};(\gamma_{i+1} + 1);T_{i+1}], \]
which in turn implies either \( Q \neq P \) or \( P \neq Q \), a contradiction.

**Q.E.D.**

Fix \( Q \) and \( R \) in \((P \in \mathcal{P}_Q(\mathcal{X})| P \neq P \) with \( Q \neq P \) \( R \). By Claim 4, for some \( i \) and \( j \) in \( \{1, \ldots, n\} \), \( Q \in \mathcal{G}_i \) and \( R \in \mathcal{G}_j \). By mixture continuity for two-outcome sublotteries, there exist \( \lambda \)
and $\lambda_1^2$ in $[0, 1]$ such that

$$
\gamma_i\left[\lambda_2^2 \delta_{s_i} + (1 - \lambda_2^2) \delta_{s_j}\right] + (1 - \gamma_i)T_i \sim_p Q \quad \text{and}
$$

$$
\gamma_i\left[\lambda_3^2 \delta_{s_i} + (1 - \lambda_3^2) \delta_{s_j}\right] + (1 - \gamma_i)T_j \sim_p R.
$$

Since $Q \succ_p R$ we have both $A_{Q^*_R}(x_1, y_1, k, T_i)$, and $A_{Q^*_R}(x_1, y_1, k, T_j)$ are not empty. If either $A_{Q^*_R}(x_1, y_1, k, T_i)$ or $A_{Q^*_R}(x_1, y_1, k, T_j)$ are not empty then we are done. For instance if $A_{Q^*_R}(x_1, y_1, k, T_i)$ is not empty, then by completeness of $\succ_p$ we have $A_{Q^*_R}(x_1, y_1, k, T_i) \cup A_{Q^*_R}(x_1, y_1, k, T_j) = [0, 1] = A_{Q^*_R}(x_1, y_1, k, T_i) \cup A_{Q^*_R}(x_1, y_1, k, T_j)$. Thus either (a) the intersection of $A_{Q^*_R}(x_1, y_1, k, T_i)$ or $A_{Q^*_R}(x_1, y_1, k, T_j)$ is a nonempty open set; or (b) the intersection of $A_{Q^*_R}(x_1, y_1, k, T_i) \cup A_{Q^*_R}(x_1, y_1, k, T_j)$ is a nonempty closed set. If (b) holds, then there exists $x'$ in $[0, 1]$ such that $R \succ_p \gamma_1 \lambda_3 \delta_{s_i} + (1 - \lambda_3) \delta_{s_j} + (1 - \gamma_i)T_i \succ_p Q$. Hence $R \succ_p Q$. This is a contradiction. Thus (a) holds and there exists $x' \in [0, 1]$ such that $R \succ_p \gamma_1 \lambda_3 \delta_{s_i} + (1 - \lambda_3) \delta_{s_j} + (1 - \gamma_i)T_i \succ_p R$. This is a nonempty open set in $[0, 1]$. Hence, $R \succ_p Q \succ_p \gamma_1 \lambda_3 \delta_{s_i} + (1 - \lambda_3) \delta_{s_j} + (1 - \gamma_i)T_i \succ_p R \cap Q$ is not empty and for $A_{Q^*_R}(x_1, y_1, k, T_i)$ is a nonempty open set in $[0, 1]$. Hence there exists $x'$ such that $R \succ_p Q$. This is a nonempty open set. Hence there exists $x' \in [0, 1]$ such that $R \succ_p Q$. This is a nonempty open set. Hence there exists $x' \in [0, 1]$ such that $R \succ_p Q$. This is a nonempty open set.

**Step 5 (Existence of $V(\cdot)$)** When there exists no Best nor Worst Lottery—joint work with Atsushi Kajii): By nondegeneracy and completeness of $\succ_p$, there exist two lotteries, $P_1, P_4$ and $P_3, P_4$, in $\mathcal{G}(\mathcal{Z})$, such that $P_3 \succ_p P_4$ and $P_3 \succ_p P_4$. Thus, by Step 4, there exists a function $V: Z \in \mathcal{G}(\mathcal{Z}) \rightarrow [0, 1]$, where $V(P_1, P_4) = 1/4$ and $V(P_3, P_4) = 3/4$, that represents $\succ_p$ restricted to the function's domain.

For any $P$ and $Q$ in $\mathcal{G}$ such that $P \succ_p Q$, define the (open) "interval" $I(Q, P) = \{R \in \mathcal{G}(\mathcal{Z}) : P \succ_p R \succ_p Q\}$. The method of proof is to show that the set $(Z \in \mathcal{G}(\mathcal{Z}) : P \succ_p Z)$ can be covered by a countable union of nested "intervals" and thus $\succ_p$ restricted to this set can be represented by a function whose range is a subset of $[3/4, 1]$. Similarly, we will then follow that $\succ_p$ restricted to $(Z \in \mathcal{G}(\mathcal{Z}) : P \succ_p Z)$ can be represented by a function whose range is a subset of $[0, 1/4]$.

Consider the set $U$ that consists of pairs of the form $(i^U, f^U)$ where $i^U$ is the countable union of a set of intervals $I((P_{3/4}, P_{1}), I((P_{3/4}, P_{1}), I((P_{3/4}, P_{1}), \ldots)$ with $P_i$ in $\mathcal{G}(\mathcal{Z})$ and $P_{i+1} \succ_p P_i$ for all $i = 1, 2, \ldots$, and $f^U: \mathcal{G}(\mathcal{Z}) \rightarrow [3/4, 1]$ represents $\succ_p$ on $i^U$. Note that from Step 4 it follows that for any pair of real numbers $a < b$, and for any interval $I(Q, P)$ there exists a function $V: Z \in \mathcal{G}(\mathcal{Z}) \rightarrow [a, b]$, with $V(Q) = a$ and $V(P) = b$ and which represents $\succ_p$ restricted to $I(Q, P)$. Hence for any $i^U$ we can find an appropriate $f^U$ by taking the limit of a sequence of functions $f^U_i$ defined as:

$$
f^U_i = \begin{cases} 
3/4 & \text{if } P_{3/4} \succ_p Z \\
V_i(Z) & \text{if } Z \in I_i(P_{3/4}, P^1) \\
7/8 & \text{if } Z \succ_p P^1 \\
15/16 & \text{if } Z \succ_p P^1 
\end{cases}
$$

$$
f^U_i = \begin{cases} 
3/4 & \text{if } P_{3/4} \succ_p Z \\
V_i(Z) & \text{if } Z \in I_i(P_{3/4}, P^1) \\
15/16 & \text{if } Z \succ_p P^1 
\end{cases}
$$

where $V_i: (Z \in \mathcal{G}(\mathcal{Z}) : P^i \succ_p Z \succ_p P_{3/4}) \rightarrow [3/4, 1 - (2^{-i-2})]$ represents $\succ_p$ on its domain, with
\( V(P_{3/4}) = 3/4 \) and \( V(P_i) = 1 - (2^{-i-2}) \) and \( f_1^U = f_{1-}^{U_{-1}} \) on \( I_{1-}(P_{3/4}, P_{1/4}) \). \( U \) can be endowed with the following ordering, \( \succ U \), defined as follows: \((I^U, f^U) \succ_U (I^U, f^U)\) if and only if \( I^U \supseteq I^U\).

**Claim 5a:** \( I^U \supseteq I^U \), implies for any \( f^U \), such that \((I^U, f^U) \in U \), there exists \( f^U \) such that \((I^U, f^U) \in U \) and \( f^U \) agrees with \( f^U \) on \( I^U \).

**Proof:** Take \( f^U = f^U \). Clearly, \( \succ U \) is a partial order. Thus, by the Hausdorff Maximal Principle (see Royden (1988, p. 25)), there exists a maximal linearly ordered (with respect to \( \succ U \)) set \( U^* \subset U \). Moreover, from Claim 5a we can choose elements in this set such that for any pair \((I^U, f^U)\) and \((I^U, f^U)\) in \( U^* \) with \( I^U \subseteq I^U \), \( f^U \) agrees with \( f^U \) on \( I^U \). Let \( I^U \) be the union of \( I^U \)'s in \( U^* \). Define \( f^{U*} \) naturally, i.e., \( f^{U*}(Q) = f^U(Q) \), where \( f^U \) is such that \((I^U, f^U) \) is in \( U^* \) and \( Q \) is in \( I^U \). Note that although there may be many \( f^U \)'s that satisfy the condition above, it does not matter because they coincide whenever they are evaluated at this same point.

**Claim 5b:** \( I^{U*} = (Z \in \mathcal{P}(\mathcal{A})) | Z \succ_P P_{3/4} \).

**Proof:** Suppose there is a \( Q \in (Z \in \mathcal{P}(\mathcal{A})) | Z \succ_P P_{3/4} \) that is not in \( I^{U*} \), i.e. \( Q \succ_P I^* \). Take \( \hat{Q} \) in \( I^{U*} \) and let \( \hat{V} : (Z \in \mathcal{P}(\mathcal{A})) | Z \succ_P \hat{P} \rightarrow \{ f^{U*}((\hat{P}) + 3/4), 1 \} \) represent \( \succ_P \) on \( I(P, Q) \), with \( V(\hat{P}) = f^{U*}((\hat{P}) + 3/4), V(Q) = 1 \). Define \( \bar{f}^U = I(P_{3/4}, Q) \) and

\[
\bar{f}^U(Z) = \begin{cases} 
\left( f^{U*}(\hat{P}) + 3/4 \right) & \text{if } \hat{P} \succ_P Z, \\
\hat{V}(Z) & \text{if } Q \succ_P Z \succ_P P, \\
1 & \text{if } Z \succ_P Q.
\end{cases}
\]

By construction \((\bar{f}^U, f^U) \succ U^* \) which contradicts maximality of \( U^* \).

Q.E.D.

Now consider the set \( L \), which consists of pairs of the form \((I^L, f^L)\) where \( I^L \) is the countable union of a set of intervals \([I_x(P_i, P_{1/4}), I_y(P_{1/4}), \ldots] \) with \( I_x(P_i, P_{1/4}) \subseteq \bigcup_{i=1}^{\infty} (P_{i+1}, P_{1/4}) \) for all \( i = 1, 2, \ldots \); and \( f^L : \mathcal{P}(\mathcal{A}) \rightarrow [0, 1/4] \) represents \( \succ_P \) on \( I^L \). By the same reasoning as above, such function exists for each \( I^L \), and the set \( L \) can be partially ordered using set inclusion over the \( I^L \)'s. By analogous reasoning as above, it follows that there exists a maximally linearly ordered subset, say \( L^* \), such that for any pair \((I^L, f^L)\) and \((I^L, f^L)\) in \( L^* \) with \( I^L \subseteq I^L \), \( f^L \) agrees with \( f^L \) on \( I^L \). Defining \( I^{L*} \) as the union of \( I^{L*} \), in \( L^* \) and \( f^{L*} \) as the function that represents \( \succ_P \) on \( I^{L*} \), a similar argument to one that proved Claim 5b can be used to show \((I^{L*} = (Z \in \mathcal{P}(\mathcal{A})) | P_{1/4} \succ_P Z) \).

Thus the function

\[
V(Z) = \begin{cases} 
V_0(Z) & \text{if } P_{1/4} \succ_P Z, \\
f^{L*}(Z) & \text{if } P_{3/4} \succ_P Z \succ_P P_{1/4} \text{ represents } \succ_P,
\end{cases}
\]

Q.E.D.

(ii) \( \Rightarrow \) (i)

**P1** (Ordering): This follows since there is a real-valued representation of \( \succ P \).

**P3\text{CU}** (Conditional Upward Monotonicity) (resp. **P4\text{CL}** (Conditional Downward Monotonicity)): This follows from the proof of Claim 2 above, which demonstrates that if probabilistic sophistication holds, then Conditional Upward (resp. Downward) Monotonicity implies **P3\text{CU}** (resp. **P3\text{CL}**).

**P4\text{CL}** (Strong Conditional Equivalence Probability): Pick disjoint events \( A \) and \( B \), outcomes \( w, x, y, z, \) and acts \( g \) and \( h \) such that \( g_1 = [x \text{ if } A \cup B; g(s) \text{ if } (A \cup B)^c] \geq g_2 = [x \text{ if } A; y \text{ if } B; g(s) \text{ if } (A \cup B)^c] \geq g_3 = [y \text{ if } A \cup B; g(s) \text{ if } (A \cup B)^c] \geq g_4 = [y \text{ if } A \cup B; g(s) \text{ if } (A \cup B)^c] \).

Clearly \( g_1, g_2, g_3, \) and \( g_4 \) imply the same probability on each outcome other than \( x \) and \( y \); thus \( \mu \ast g_2^{-1} \) and \( \mu \ast g_3^{-1} \) are mixtures of \( \mu \ast f^{-1} \) and \( \mu \ast g_2^{-1} \). Hence the indifference of \( g_2 \) and \( g_3 \) implies that \( \mu(A) = \mu(B) \). This fact combined with probabilistic sophistication implies that the act \( [w \text{ if } A; z \text{ if } B; h(s) \text{ if } (A \cup B)^c] \) is indifferent to the act \( [z \text{ if } A; w \text{ if } B; h(s) \text{ if } (A \cup B)^c] \).

**P5** (Nondegeneracy): This follows since \( V : \mathcal{A} \rightarrow \mathbb{R} \) is not constant.
P6\(^{+}\) (Small Event Continuity): Pick arbitrary acts \(f > g\) and outcome \(x\). Let \(P = (y_1, p_1; \ldots; y_n, p_n)\) and \(R = (z_1, r_1; \ldots; z_r, r_r)\) be distributions implied by \(f\) and \(g\). Let \(\{C_1, \ldots, C_T\}\) and \(\{D_1, \ldots, D_T\}\) be the partition of \(\mathcal{A}\) induced by \(f\) and \(g\), i.e. \(C_i = f^{-1}(y_i)\) and \(D_i = g^{-1}(z_i)\). Mixture continuity for two-outcome sublotteries of \(\succeq_p\) implies that for all \(y_i, \lambda \in [0, 1], p_\lambda = (1 - \lambda)\delta_{y_i}\) and \((1 - p_\lambda)\delta_x\) is open in \([0, 1]\), where \(P_{\lambda} = (y_1, p_1/(1-p_\lambda); \ldots; y_\lambda, p_{\lambda-1}/(1-p_{\lambda-1}); \ldots; y_T, p_T/(1-p_T))\). For each \(i = 1, \ldots, T\), let

\[
\lambda_{y_i} = \begin{cases} 
1 & \text{if } \{\lambda \in [0, 1] | p_\lambda \in [0, 1] \frac{\lambda \delta_x + (1-\lambda) \delta_y}{1-p_{\lambda-1}} \cap (1-p_\lambda) \delta_x \in \mathcal{A}\} = [0, 1], \\
\inf \{\lambda \in [0, 1] | R \succeq_p p_\lambda \in [0, 1] \frac{\lambda \delta_x + (1-\lambda) \delta_y}{1-p_{\lambda-1}} \cap (1-p_\lambda) \delta_x \in \mathcal{A}\} & \text{otherwise}. 
\end{cases}
\]

Let \(\epsilon_x = \min \{\lambda, p_\lambda\}\). As \(\mu\) is nonatomic, there exists a finite refinement of the partition \(\{C_1, \ldots, C_s\}, \{C_{s+1}, \ldots, C_{s+n}\}\) for which \(\mu(C_{i,j}) < \epsilon_x\) for all \(s = 1, \ldots, S, j = 1, \ldots, n\). Thus it follows from probabilistic sophistication that for all \(C_{i,j} \subset C_{i,j'}\), \(x\) if \(C_{i,j} \succeq C_{i,j'}\), \(f(x)\) if \(\mathcal{A}/C_{i,j} \succeq g(x)\) if \(\mathcal{A}/C_{i,j'}\). P6\(^{+}\) holds for coarsest common refinement of these two partitions.

Q.E.D.

REFERENCES


