SUBJECTIVE PROBABILITIES ON "SMALL" DOMAINS

Igor Kopylov*

Institute for Mathematical Behavioral Sciences, and

Department of Economics, University of California, Irvine, CA 92697

Abstract

The choice-theoretic derivation of subjective probability proposed by Savage (and generalized by Machina and Schmeidler) does not apply in settings like the Ellsberg Paradox, where choice behavior reflects a distinction between risk and ambiguity. This paper formulates two representation results—one for expected utility, the other for probabilistic sophistication—that derive subjective probabilities but only on a "small" domain of risky events. These events can be either specified exogenously or defined in terms of choice behavior; in the latter case, both the values and the domain of probability are subjective. The analysis identifies a mathematical structure—called a mosaic—that is appropriate for the domain of subjective probability. In contrast with an algebra or a σ-algebra, this structure is implied by the behavioral definitions of risky events.

Keywords: Ambiguity, subjective probability, expected utility, probabilistic sophistication, Ellsberg Paradox.

*This paper is based on a chapter of my Ph.D. thesis (University of Rochester, 2003). I owe special thanks to Larry Epstein for stimulating my interest in the subject and for advising me patiently throughout writing this paper. I am also grateful to Peter Wakker, Mark Machina, Peter Klibanoff, and Arthur Robson for helpful comments. All mistakes are my own.

1 Introduction

1.1 Objectives

Savage [20] provides foundations for the use of *subjective probability* in decision making. More precisely, he axiomatizes an *expected utility* representation for preference over uncertain prospects—acts defined on a set S of states of nature—in a way that does not rely on any extraneous randomization device. One component of this representation is a probabilistic belief assigned subjectively on the universal class of events Σ, such as the power set of all subsets of S.

Two well-known paradoxes challenge Savage's theory. In the paradox due to Allais [1], all events have explicit numerical probabilities, but typical preferences are not represented by expected utility. Motivated by the Allais paradox, Machina and Schmeidler [16] extend Savage's theory and characterize the use of subjective probabilities separately from the expected utility functional form. In their model of *probabilistic sophistication*, the decision maker ranks acts in two stages: first, she uses subjective probabilities to translate each act into a lottery (a distribution over outcomes), and then she ranks the induced lotteries but not necessarily via expected utility. Note that both in Savage's theory and in Machina–Schmeidler's extension, subjective probabilities are derived for all events in the universal class Σ.

In the other paradox, due to Ellsberg [4], the decision maker is unwilling to assign probabilities to all events in Σ. For example, consider an urn with balls of three possible colors: B, G, and R. Suppose that the decision maker is told only that the total number of balls is 90 and that $R = 30$. Then probabilities of events in the class

$$\mathcal{R}_0 = \{\emptyset,\ \{R\},\ \{B, G\},\ \{B, G, R\}\}$$

are known. The typical preference is to bet on events in $\mathcal{R}_0$ rather than on events with imprecisely known probabilities, for example,

(i) to bet on $\{R\}$ rather than on $\{B\}$ because the probability of $\{R\}$ is known to be $\frac{1}{3}$ while the probability of $\{B\}$ lies between 0 and $\frac{2}{3}$;

(ii) to bet on $\{B, G\}$ rather than on $\{R, G\}$ because the probability of $\{B, G\}$ is known to be $\frac{2}{3}$ while the probability of $\{R, G\}$ lies between $\frac{1}{3}$ and 1.

Such preference reversal is inconsistent with the use of any subjective probability measure p on all events because it implies that $p(\{R\}) > p(\{B\})$ but also $p(\{B\}) + p(\{G\}) > p(\{R\}) + p(\{G\})$. Thus the Ellsberg Paradox shows behavioral significance of the well-known distinction (Knight [13]) between *risk*, which can be represented by numerical probabilities, and *ambiguity*, which cannot.[1]

In the light of his findings, Ellsberg states that "both the predictive and normative use of the Savage or equivalent postulates might be improved by avoiding attempts to apply them in certain, specifiable circumstances where they do not seem acceptable." In other words, Ellsberg suggests that a theory of subjective probability should specify a distinction between risk and ambiguity and then apply only to risky acts and events. The main objective of this paper is to develop such a theory as a natural extension of Savage's and Machina–Schmeidler's results. In order to do so, we address several related questions:

1. How can risky events and risky acts be specified?

2. What mathematical structure does the class of risky events have?

3. How can Savage's (or Machina–Schmeidler's) axioms be adapted in order to characterize expected utility maximization (or probabilistic sophistication) on risky acts?

In some settings, one can specify an exogenous class $\mathcal{R} \subset \Sigma$ of risky events.[2] For example, one can take $\mathcal{R}$ to be the class of events where probabilities are given to the decision maker explicitly, such as $\mathcal{R} = \mathcal{R}_0$ in the Ellsberg Paradox. In general, such exogenous formulations may seem arbitrary because decision makers

[1] Knight uses *uncertainty* rather than *ambiguity*. We adopt Ellsberg's terminology, in which uncertainty is comprehensive and includes both risk and ambiguity.

[2] Exogenous formulations are used by Sarin and Wakker [19], by Zhang [23], and by Epstein [5].

may disagree about the identity of events to which they assign probabilities.[3] To address this concern, one can define *subjectively* risky events in terms of preference. More precisely, for each event $A \in \Sigma$, whether A is (subjectively) risky or (subjectively) ambiguous can be determined solely on the basis of the decision maker's preference over acts. This can be done via definitions due to Zhang [24] and to Epstein–Zhang [6]. Either of these definitions derives a class of subjectively risky events, written as $\mathcal{R}_Z \subset \Sigma$ and $\mathcal{R}_{EZ} \subset \Sigma$ respectively, in a way that rules out Ellsberg-type choice among subjectively risky acts. The difference between the two definitions is that the former also rules out Allais-type behavior while the latter does not. Accordingly, Zhang uses one of Savage's postulates to motivate his approach, while Epstein–Zhang employ a weaker postulate due to Machina–Schmeidler to motivate theirs.[4]

In order to derive subjective probabilities on a class of risky events, one needs to identify a mathematical structure of this class. Zhang [23] argues that this structure may be weaker than an algebra or *a fortiori*, a σ-algebra (used both by Savage and by Machina–Schmeidler). For example, consider an Ellsberg-type urn with four possible colors: B, G, R, and Y. Suppose that the decision maker is told only that the total number of balls is 100 and that $B + G = B + Y = 60$.

[3] For example, Ellsberg observes various responses to his paradox among "sophisticated and reasonable" people:

> There are those who do not violate the [Savage] axioms, or say they won't... (e.g., G. Debreu, R. Shlaiffer, P. Samuelson); these subjects tend to apply the axioms rather than their intuition, and when in doubt, to apply some form of the Principle of Insufficient Reason. Some violate the axioms cheerfully... (J. Marschak, N. Dalkey); others sadly but persistently, ... this group included L. J. Savage, when last tested by me.

[4] Nehring [17] and Ghirardato–Marinacci [8] propose other definitions of ambiguity. These definitions are less appropriate for a general theory of subjective probability because they are at least in part motivated by considerations other than Savage's (or Machina–Schmeidler's) behavioral postulates.

Then

$$\mathcal{R}_1 = \{\emptyset,\ \{B,G\},\ \{B,Y\},\ \{R,Y\},\ \{G,R\},\ \{B,G,R,Y\}\}$$

is the class of events for which probabilities are given explicitly. Accordingly, one could take $\mathcal{R} = \mathcal{R}_1$. Alternatively, one could use a class of subjectively risky events, either $\mathcal{R}_Z$ or $\mathcal{R}_{EZ}$; later we provide an example of preferences where $\mathcal{R}_Z = \mathcal{R}_{EZ} = \mathcal{R}_1$. However, $\mathcal{R}_1$ is not an algebra because it is not closed under intersections; for instance, $\{B,G\} \cap \{G,R\} = \{G\} \notin \mathcal{R}_1$. Instead of an algebra, a weaker structure—called a *mosaic*—can be derived for both $\mathcal{R}_Z$ and $\mathcal{R}_{EZ}$ from their definitions. This motivates modelling the use of subjective probability on mosaics.[5]

One contribution of this paper is to identify mosaics as a natural structure for the domain of subjective probability. Second, we formulate two main representation results—one for subjective expected utility (Theorem 3.1), the other for probabilistic sophistication (Theorem 4.1)—that derive subjective probabilities on a given mosaic $\mathcal{R} \subset \Sigma$. One can apply these results to $\mathcal{R} = \mathcal{R}_Z$ and to $\mathcal{R} = \mathcal{R}_{EZ}$ respectively. In this way, one can obtain a fully subjective theory of expected utility and a fully subjective theory of probabilistic sophistication. Both theories are constructive and derive the values and the domain of the subjective probabilistic belief from preference. However, the two theories use different definitions of risk and obtain different representations for preference over risky acts. The former delivers an expected utility representation on $\mathcal{R}_Z$-measurable acts, while the latter delivers probabilistic sophistication on $\mathcal{R}_{EZ}$-measurable acts.

The noted Theorems 3.1 and 4.1 extend Savage's and Machina–Schmeidler's results from σ-algebras to mosaics. For the most part, their axioms are retained; the main modification is required for Savage's P6 (used also by Machina–Schmeidler). In the proof, the fact that the class of risky events is a mosaic rather than a σ-algebra necessitates constructing subjective probabilities in a way different from

[5] Zhang [23, 24] and Epstein–Zhang [6] employ another structure, called a λ-system, that is motivated by the additive property of probability. This structure is more restrictive than a mosaic and, without additional assumptions, is not implied by the definitions of $\mathcal{R}_Z$ and $\mathcal{R}_{EZ}$ (see an example in Section 4.2).

Savage's. We provide an explicit formula that computes subjective probabilities for risky events from the decision maker's "preference to bet" on these events. This formula captures a simple intuition and is free of some of the limitations of Savage's approach.

Our results allow a lot of flexibility for applications. In particular, choice among ambiguous acts is not restricted by any conditions or *a fortiori,* by any parametric model, such as Choquet expected utility (Schmeidler [21]) or the multiple priors model (Gilboa and Schmeidler [9]). Also, there is freedom in the identity of the mosaic $\mathcal{R}$ where the use of subjective probability is axiomatized. In this paper, we focus on applications where $\mathcal{R} = \mathcal{R}_Z$ or $\mathcal{R} = \mathcal{R}_{EZ}$, but in general, $\mathcal{R}$ need not be equal to either of these domains.

1.2 Finite or Countable Additivity

Another strength of our model is that it does not use countable set operations, which are essential both for Savage's and for Machina–Schmeidler's results. These authors assume the universal class of events Σ to be a σ-algebra.[6] However, as Savage [p. 43] notes, it is peculiar that one should use countable unions of events in order to derive a finitely additive probability measure.

There is another reason why the use of σ-algebras is problematic for Savage's theory. Prior to ranking all Σ-measurable acts, the decision maker must *conceive* of all events in the class Σ. However, a σ-algebra, even when generated by simple events, often contains very complex ones. To illustrate, let each state of the world s be determined by an infinite sequence of coin tosses, each toss resulting in either heads or tails. Then $s \in S = \prod_{k=1}^{\infty}\{H_k, T_k\}$. For an arbitrary finite n, identify every $A \subset \prod_{k=1}^{n}\{H_k, T_k\}$ with the obvious event (subset in S). Call such a set A a cylinder. Let Σ be the algebra of cylinders, and let $\sigma(\Sigma)$ be the smallest σ-algebra that contains Σ. It is well-known that many events in $\sigma(\Sigma)$ are not readily obtained via combinations of cylinders. More precisely, for any countable

[6] Formally, Savage takes Σ to be the power set, but his analysis is unchanged for Σ an arbitrary σ-algebra.

ordinal γ, there are sets in $\sigma(\Sigma)$ that cannot be arrived at from Σ by a γ-sequence of set-theoretic operations, each operation being a complement, a countable union or a countable intersection (Billingsley [2, pp. 31–32]). In this sense, the class $\sigma(\Sigma)$ is substantially more complex than Σ.

Probability theorists avoid this complexity and do not construct measures directly on a σ-algebra; rather, the typical procedure is to describe the measure explicitly on "simple" events in a subalgebra and then to apply a measure extension theorem (recall the construction of the Lebesgue measure on the Borel σ-algebra). Similarly, the decision maker who conceives of all events in an algebra Σ may be unable or unwilling to conceive of some sets in $\sigma(\Sigma)$. Therefore, reliance on all events in a σ-algebra may be problematic for the normative theory of subjective probability. Our model alleviates this problematic aspect by relaxing Savage's primitives and by assuming Σ to be only a finitely additive algebra. This relaxation is possible, roughly, without changing Savage's axioms (compare with Gul [11] who uses axioms different from Savage's in order to allow S to be finite).

1.3 Outline

This paper proceeds as follows. Next we introduce a version of Savage's framework and define the notion of a mosaic. In Section 3, we axiomatize expected utility on $\mathcal{R}$-measurable acts for an arbitrary mosaic $\mathcal{R}$ (Theorem 3.1) and apply this result to $\mathcal{R} = \mathcal{R}_Z$ (Corollary 3.2). In Section 4, we axiomatize probabilistic sophistication on $\mathcal{R}$-measurable acts for an arbitrary mosaic $\mathcal{R}$ (Theorem 4.1) and apply this result to $\mathcal{R} = \mathcal{R}_{EZ}$ (Corollary 4.2). Proofs are sketched in Section 5 and are presented in detail in Appendix.

2 Preliminaries

We use a version of Savage's framework. Given are a set $S = \{s, \ldots\}$ of states of the world and a set $X = \{x, y, z, \ldots\}$ of outcomes or prizes; as in Savage, no structure is imposed on the sets S and X. Also given is an *algebra* $\Sigma =$

$\{A, B, C, D, E, \dots\}$ of subsets of S that are called *events*. By definition, the algebra Σ satisfies the following conditions:

$(*)\quad S \in \Sigma\,;$

$(**)\quad A \in \Sigma \quad \Rightarrow \neg A \in \Sigma\,;$[7] and

$(\alpha)\quad A_1 \in \Sigma$ and $A_2 \in \Sigma \quad \Rightarrow\ A_1 \cup A_2 \in \Sigma.$

It follows that Σ is closed under finite unions and intersections. Call a union $A = \cup_{i=1}^{n} A_i$ a *partition* if the events $A_1, A_2, \dots, A_n$ are disjoint.

The algebra Σ is a σ-*algebra* if it is closed also under countable set operations, that is, if it satisfies

$(\sigma)\quad A_i \in \Sigma$ for $i = 1, 2, \dots \quad \Rightarrow\ \bigcup_{i=1}^{\infty} A_i \in \Sigma.$

By taking the primitive class of events Σ to be an algebra, we relax the assumption that Σ is a σ-algebra used both by Savage and by Machina–Schmeidler.

Call a function $f : S \to X$ an *act* if it has finite range and if it is Σ-measurable, that is, if

$$f^{-1}(x) = \{s \in S : f(s) = x\} \in \Sigma \qquad \text{for all } x \in X.$$[8]

The restriction of the act f to an event $A \in \Sigma$ is called a *subact*. For notational simplicity, identify every $x \in X$ with the constant act yielding the outcome x in all states $s \in S$.

Denote by $\mathcal{F} = \{f, g, h, \dots\}$ the set of acts. Interpret each act $f \in \mathcal{F}$ as a physical action that results in the outcome $f(s)$ contingent on the realized state of the world $s \in S$. The decision maker's preference among such physical actions is given as a binary relation $\succeq$ on $\mathcal{F}$.

In addition to Savage's primitives S, X, Σ, and $\succeq$, take a class $\mathcal{R} \subset \Sigma$ of *risky events* as given and call all other events in Σ *ambiguous*. Say that $A = \cup_{i=1}^{n} A_i$ is a *risky partition* if the event A and all the events $A_1, \dots, A_n$ are risky. Assume that $\mathcal{R}$ satisfies the following conditions:

[7] $\neg A$ denotes the complement of the set A in S.

[8] In the literature, f is often called a *simple* act to reflect the fact that it has finite range. This terminology is redundant in our model because we do not use acts other than simple.

($*$) $S \in \mathcal{R}$;

($**$) $A \in \mathcal{R} \quad \Rightarrow \neg A \in \mathcal{R}$; and

(μ) $S = \cup_{i=1}^{m} S_i$ is a risky partition $\quad \Rightarrow S_i \cup S_j \in \mathcal{R}$ for all $i, j = 1 \ldots m$.

Call such $\mathcal{R}$ a *mosaic*. Note that the mosaic $\mathcal{R}$ is closed under unions of elements of any fixed risky partition of the universal event S; however, $\mathcal{R}$ may not be closed under arbitrary unions and intersections. The properties of a mosaic are intuitive if we think of $\mathcal{R}$ as a class of events where the decision maker assigns probabilities. Moreover, mosaics—unlike more restrictive structures—accommodate the behavioral definitions of risky events, due to Zhang [24] and Epstein–Zhang [6], that we adopt later.

Call an act $f \in \mathcal{F}$ *risky* if it is $\mathcal{R}$-measurable, that is, if $f^{-1}(x) \in \mathcal{R}$ for all $x \in X$; call f *ambiguous* otherwise. Denote by $\mathcal{G} \subset \mathcal{F}$ the set of risky acts. Note that every constant act $x \in X$ is risky because $S \in \mathcal{R}$ and $\emptyset = \neg S \in \mathcal{R}$.

The fact that $\mathcal{R}$ may not be closed under intersections makes the following notation useful. Given a collection of risky events $\mathcal{E} \subset \mathcal{R}$, let

$$\mathcal{R} \Cap \mathcal{E} = \{A \in \mathcal{R} : A \cap E \in \mathcal{R} \quad \text{for all } E \in \mathcal{E}\},$$

and let $\mathcal{G} \Cap \mathcal{E}$ be the set of $(\mathcal{R} \Cap \mathcal{E})$-measurable acts. In other words, $\mathcal{G} \Cap \mathcal{E}$ is the set of such risky acts that remain $\mathcal{R}$-measurable when restricted to an arbitrary event $E \in \mathcal{E}$. Note that if $\mathcal{R}$ is an algebra, then the above notation is redundant because $\mathcal{R} = \mathcal{R} \Cap \mathcal{E}$ and $\mathcal{G} = \mathcal{G} \Cap \mathcal{E}$ for each collection $\mathcal{E} \subset \mathcal{R}$.

Given acts $f, g \in \mathcal{F}$ and an event $A \in \Sigma$, denote by fAg the *composite act* that yields $f(s)$ if $s \in A$ and $g(s)$ if $s \in \neg A$. Note that fAg may be ambiguous even when the event A and the acts f and g are risky. For example, take distinct outcomes x and y and risky events $A, B \in \mathcal{R}$ such that $A \cap B \notin \mathcal{R}$. Then the acts $f = xBy$ and $g = y$ are risky but $fAg = x(A \cap B)y$ is not. In order for the composite act fAg to be risky, it is sufficient to require that $A \in \mathcal{R}$, $f \in \mathcal{G} \Cap \{A\}$ and $g \in \mathcal{G} \Cap \{\neg A\}$. Under these conditions, S can be partitioned into a finite number of risky events that have the form $f^{-1}(x) \cap A$ or $g^{-1}(x) \cap \neg A$ for $x \in X$. As $\mathcal{R}$ satisfies (μ), the disjoint union

$$(f^{-1}(x) \cap A) \cup (g^{-1}(x) \cap \neg A) = (fAg)^{-1}(x)$$

is a risky event for each $x \in X$. Thus the act fAg is risky.

3 Subjective Expected Utility

Savage axiomatizes an expected utility representation for preference over the $\mathcal{F}$ set of all acts. Next, we formulate his axioms P1–P6 (see Savage [20], Fishburn [7], and Kreps [15] for more detailed treatments of these axioms).

Axiom P1 (Ordering). $\succeq$ *is complete and transitive on* $\mathcal{F}$.

P1 is a standard rationality condition.

Axiom P2 (Sure-Thing Principle). *For all events* $A \in \Sigma$ *and for all acts* $f, g, h, h' \in \mathcal{F}$,

$$fAh \succeq gAh \quad \Rightarrow \quad fAh' \succeq gAh'. \tag{3.1}$$

P2 requires that preference is separable across mutually exclusive events and can be conditioned on any event A independently of the outcomes obtained on $\neg A$. Without loss of generality, this axiom can be simplified as follows: for all events $A \in \Sigma$, for all acts $f, g \in \mathcal{F}$ and for all outcomes $x, y \in X$,

$$fAx \succeq gAx \quad \Rightarrow \quad fAy \succeq gAy. \tag{3.2}$$

Obviously, invariance (3.2) is a special case of (3.1) where the acts h and h' are constant; on the other hand, both h and h' in (3.1) have finite range, and (3.1) follows from (3.2) by induction.

Axiom P3 (Eventwise Monotonicity). *For each event* $A \in \Sigma$*, at least one of the following statements holds simultaneously for all outcomes* $x, y \in X$ *and for all acts* $h \in \mathcal{F}$*:*

(i) $x \succeq y \quad \Leftrightarrow \quad xAh \succeq yAh$ *;*

(ii) $xAh \sim yAh$.

P3 postulates that preference over certain outcomes in X remains unchanged when conditioned on any event A that is not viewed as virtually impossible, and becomes degenerate when conditioned on any null event.

Axiom P4 (Weak Comparative Probability). *For all events $A, B \in \Sigma$, for all outcomes $x \succ x'$ and $z \succ z'$,*

$$xAx' \succeq xBx' \quad \Rightarrow \quad zAz' \succeq zBz'.$$

P4 requires that the preference to bet on the event A rather than on the event B is independent of the stakes involved in the bets and is based exclusively on the (subjective) relative likelihoods of A and B.

Axiom P5 (Non-degeneracy). *There exist outcomes x and x' such that $x \succ x'$.*

P5 needs no explanation.

Axiom P6 (Small Event Continuity). *For any outcome x and for any acts $f \succ g$, there exists a partition $S = \cup_{i=1}^{m} S_i$ such that for all $i = 1 \ldots m$, $xS_if \succ g$ and $f \succ xS_ig$.*

P6 requires that for any outcome x and for any pair of acts $f \succ g$, the universal set S can be partitioned into (sufficiently) small events $S_1, \ldots, S_m$. Here, the event S_i being small means that the strict preference $f \succ g$ is not reversed when the outcomes $f(s)$, or alternatively $g(s)$, are replaced by x for all $s \in S_i$. To motivate P6, suppose that the description of each state of the world includes infinitely many coin tosses. Identify every finite sequence of heads and tails $(s_1, \ldots, s_n) \in \prod_{k=1}^{n}\{H_k, T_k\}$ with the obvious event in S. It is intuitive that for a sufficiently large n, the decision maker views every sequence $(s_1, \ldots, s_n)$ as too unlikely to reverse the preference $f \succ g$ when $f(s)$, or alternatively $g(s)$, are replaced by x for all $s \in (s_1, \ldots, s_n)$.[9]

[9]This is intuitive even if the decision maker does not view the coin as fair or the coin tosses as independent. For example, the subjective probability measure p may be exchangeable.

Call a function $p : \Sigma \to [0,1]$ a *probability measure* if $p(S) = 1$ and p is finitely additive, that is, $p(A_1 \cup A_2) = p(A_1) + p(A_2)$ for all disjoint events $A_1, A_2 \in \Sigma$. Call a probability measure $p : \Sigma \to [0,1]$ *convex-ranged* if for each event $B \in \Sigma$, the range $\{p(A) : A \in \Sigma,\ A \subset B\}$ equals the interval $[0, p(B)]$.

Finally, assume that the algebra Σ of events is a σ-algebra. Even though this assumption is not required in order to state Savage's axioms, it is essential for his construction of subjective probabilities.

Theorem (Savage). *Let Σ be a σ-algebra. The following two statements are equivalent.*

1. *$\succeq$ satisfies P1, P2, P3, P4, P5, P6.*

2. *$\succeq$ is represented by expected utility*

$$U(f) = \sum_{x \in X} u(x) \cdot p\left(f^{-1}(x)\right) \qquad \text{for } f \in \mathcal{F}, \tag{3.3}$$

   where $u : X \to \mathbb{R}$ is a non-constant utility index, and $p : \Sigma \to [0,1]$ is a convex-ranged probability measure.

In this representation, the index u is unique up to a positive linear transformation, and the probability measure p is unique.

The decision maker as portrayed by (3.3) assigns probabilities $p(A)$ to all events $A \in \Sigma$, attaches utility indices $u(x)$ to all outcomes $x \in X$, and then ranks all acts $f \in \mathcal{F}$ via expected utility. Thus, axioms P1–P6 provide foundations for expected utility maximization on the set $\mathcal{F}$ of all acts. However, one of these axioms, the Sure-Thing Principle, is violated by Ellsberg-type behavior and hence, may not be universally acceptable. This motivates restricting P1–P6 and the associated expected utility representation to the domain of risk.

3.1 Expected Utility On Risky Acts

In this section we rewrite Savage's axioms on the set $\mathcal{G}$ of risky acts so that the rewritten axioms—most importantly, the Sure-Thing Principle—do not restrict choice among ambiguous acts but still deliver an expected utility representation on the set $\mathcal{G}$ and accordingly, a subjective probability measure on $\mathcal{R}$.

It is straightforward to reformulate P1, P4, and P5.

Axiom P1($\mathcal{R}$). *$\succeq$ is complete and transitive on $\mathcal{G}$.*

Note that P1 implies P1($\mathcal{R}$) and is equivalent to P1(Σ).

Axiom P4($\mathcal{R}$). *For all risky events $A, B \in \mathcal{R}$, for all outcomes $x \succ x'$ and $z \succ z'$,*

$$xAx' \succeq xBx' \quad \Rightarrow \quad zAz' \succeq zBz'.$$

Obviously, any bet on a risky event, such as xAx', is a risky act. Note that P4 implies P4($\mathcal{R}$) and is equivalent to P4(Σ).

Axiom P5($\mathcal{R}$). *There exist outcomes x and x' such that $x \succ x'$.*

P5($\mathcal{R}$) is identical to P5.

In order to rewrite the Sure-Thing Principle, one needs to ensure that all composite acts involved in invariance (3.2) are risky. For example, it would be inappropriate to require that for all risky events $A, B \in \mathcal{R}$, for all risky acts $f, g, \in \mathcal{G}$ and for all outcomes $x, y \in X$,

$$fAx \succeq gAx \quad \Rightarrow \quad fAy \succeq gAy$$

because some of the composite acts fAx, gAx, fAy, or gAy may be ambiguous. To guarantee that all of these acts are risky, it is sufficient to take f and g to be $\mathcal{R}$-measurable on A.

Axiom P2($\mathcal{R}$). *For all risky events $A \in \mathcal{R}$, for all risky acts $f, g \in \mathcal{G} \Cap \{A\}$ and for all outcomes $x, y \in X$,*

$$fAx \succeq gAx \quad \Rightarrow \quad fAy \succeq gAy. \tag{3.4}$$

P2($\mathcal{R}$) requires that preference is separable but only across mutually exclusive risky events. More precisely, P2($\mathcal{R}$) states that preference over acts that are $\mathcal{R}$-measurable on a risky event A and constant on $\neg A$ can be conditioned on A

independently of the outcome obtained on $\neg A$. This version of separability does not restrict choice among ambiguous acts. Note that P2 implies P2($\mathcal{R}$) and is equivalent to P2(Σ).

Rewrite P3 as follows.

Axiom P3($\mathcal{R}$). *For each risky event $A \in \mathcal{R}$, at least one of the following statements holds simultaneously for all outcomes $x, y \in X$ and for all acts $h \in \mathcal{G} \Cap \{\neg A\}$:*

(i) $x \succeq y \quad \Leftrightarrow \quad xAh \succeq yAh$;

(ii) $xAh \sim yAh$.

This axiom requires monotonicity but only for preference over risky acts. Note that P3 implies P3($\mathcal{R}$) and is equivalent to P3(Σ).

Finally, rewrite P6 as follows.

Axiom P6($\mathcal{R}$). *For any outcome x, for any finite collection of risky events $\mathcal{E} \subset \mathcal{R}$, and for any $\mathcal{E}$-measurable acts $f \succ g$, there exists a risky partition $\{S_1, \ldots, S_m\} \subset \mathcal{R} \Cap \mathcal{E}$ of S such that for all $i = 1 \ldots m$, $xS_if \succ g$, and $f \succ xS_ig$.*[10]

P6($\mathcal{R}$) postulates existence of (sufficiently) small events $S_1, \ldots, S_m$ but, unlike P6, requires that these events belong to a particular subclass $\mathcal{R} \Cap \mathcal{E} \subset \mathcal{R}$. If $\mathcal{R}$ is an algebra, then $\mathcal{R} \Cap \mathcal{E} = \mathcal{R}$; for example, P6($\Sigma$) is equivalent to P6. To motivate P6($\mathcal{R}$) in the general case when $\mathcal{R}$ is not an algebra, suppose that, given a finite collection $\mathcal{E} \subset \mathcal{R}$ and $\mathcal{E}$-measurable acts, the description of each state of the world includes infinitely many coin tosses that are viewed as risky and as independent of all events in $\mathcal{E}$.[11] Then it is intuitive that

[10] The composite acts xS_if and xS_ig are risky because $f, g \in \mathcal{G} \Cap \{\neg S_i\}$ for all $i = 1 \ldots m$. To prove this, partition S into a finite number of risky events that have the form $f^{-1}(z) \cap S_i$ for some $z \in X$ and some $i = 1 \ldots m$; here, $f^{-1}(z) \cap S_i \in \mathcal{R}$ because $f^{-1}(z) \in \mathcal{E}$ and $S_i \in \mathcal{R} \Cap \mathcal{E}$. Then by ($\mu$), for all $x \in X$ and for all $i = 1 \ldots m$, $f^{-1}(x) \cap \neg S_i = \cup_{j \neq i}(f^{-1}(x) \cap S_j) \in \mathcal{R}$.

[11] Here, independence means that observing results of coin tosses does not affect the decision maker's perception of the likelihoods of events in $\mathcal{E}$. It should be stressed that this intuitive notion of independence is used only to motivate P6($\mathcal{R}$) and is not a formal part of our model.

(i) for all sequences of coin tosses $(s_1, \ldots, s_n)$ and for all events $E \in \mathcal{E}$, the intersection $(s_1, \ldots, s_n) \cap E$ is a risky event because the decision maker, who assigns probabilities $p(s_1, \ldots, s_n)$ and $p(E)$, should assign probability $p(s_1, \ldots, s_n) \cdot p(E)$ to the event $(s_1, \ldots, s_n) \cap E$;

(ii) for n sufficiently large, the decision maker views each sequence $(s_1, \ldots, s_n)$ as too unlikely to reverse her preference $f \succ g$ when the outcomes $f(s)$, or alternatively $g(s)$, are replaced by x for all $s \in (s_1, \ldots, s_n)$.

Note that P6($\mathcal{R}$) restricts both the ranking of risky acts and the class $\mathcal{R}$ of risky events; in particular, P6($\mathcal{R}$) implies that $\mathcal{R}$ is infinite.

Call a function $p : \mathcal{R} \to [0, 1]$ a *probability measure* if for all risky partitions $S = \cup_{i=1}^m S_i$ of the universal event,

$$\sum_{i=1}^{m} p(S_i) = 1.$$

Call a probability measure $p : \mathcal{R} \to [0, 1]$ *finely ranged* if for any finite collection $\mathcal{E} \subset \mathcal{R}$ and for any $\varepsilon > 0$, there exists a risky partition $S = \cup_{i=1}^m S_i$ such that $S_i \in \mathcal{R} \Cap \mathcal{E}$ and $p(S_i) < \varepsilon$ for all $i = 1 \ldots m$.

Our first main result is

Theorem 3.1. *Let Σ be an algebra, and let $\mathcal{R} \subset \Sigma$ be a mosaic. Then the following two statements are equivalent.*

1. *$\succeq$ satisfies P1($\mathcal{R}$), P2($\mathcal{R}$), P3($\mathcal{R}$), P4($\mathcal{R}$), P5($\mathcal{R}$), P6($\mathcal{R}$).*

2. *$\succeq$ is represented on the set $\mathcal{G}$ by expected utility*

$$U(f) = \sum_{x \in X} u(x) \cdot p\left(f^{-1}(x)\right) \qquad \text{for } f \in \mathcal{G}, \tag{3.5}$$

where $u : X \to \mathbb{R}$ is a non-constant utility index, and $p : \mathcal{R} \to [0, 1]$ is a finely ranged probability measure.

In this representation, the index u is unique up to a positive linear transformation, and the probability measure p is unique.

The decision maker as portrayed by (3.5) assigns probabilities to all risky events $A \in \mathcal{R}$, attaches utility indices to all outcomes $x \in X$, and then ranks all risky acts $f \in \mathcal{G}$ via expected utility. Therefore, P1($\mathcal{R}$)–P6($\mathcal{R}$) provide foundations for the use of probabilities on risky events and for expected utility maximization on risky acts.

The Savage Theorem is a special case of Theorem 3.1, where $\mathcal{R} = \Sigma$ and Σ is a σ-algebra. In this case, both results use equivalent axioms and deliver the same representations.[12] Moreover, Theorem 3.1 implies that even if the algebra Σ does not satisfy (σ), Savage's axioms *taken "as is"* are necessary and sufficient for the preference $\succeq$ over Σ-measurable acts to be represented by expected utility. In other words, the assumption that Σ is a σ-algebra is not crucial for axiomatizing expected utility via P1–P6.

In general, Theorem 3.1 uses axioms that are parallel to Savage's counterparts but applies them only to risky acts and events. Most importantly, P2($\mathcal{R}$) does not postulate separability of preference across ambiguous events. Accordingly, in representation (3.5), only risky events are assigned subjective probabilities and only risky acts are ranked via expected utility. Note that choice among ambiguous acts is not restricted by any parametric utility representation such as Choquet expected utility (Schmeidler [21]) or maxmin expected utility (Gilboa and Schmeidler [9]).

In contrast with Savage's result, Theorem 3.1 delivers a subjective probability measure p that is not necessarily convex-ranged. The construction of such p requires a new approach that is sketched in Section 5.

3.2 Fully Subjective Expected Utility

Similar to the Savage Theorem, Theorem 3.1 is formulated for a given class of events. Accordingly, subjective probabilities in the expected utility representation (3.5) are derived on a mosaic $\mathcal{R}$ which is *exogenous* to the model. It may be

[12]The equivalence of representations (3.5) and (3.3) follows from the fact that in either of them, the probability measure p is unique and the utility index u is unique up to a positive linear transformation.

unclear how to specify $\mathcal{R}$ in settings where decision makers may disagree about the identity of events to which they assign probabilities. In order to address this concern, one can define risky events subjectively, that is, in terms of preference $\succeq$. The following definition, due to Zhang [24], is motivated by the Sure-Thing Principle.

Definition (Zhang). *Call an event $E \in \Sigma$ subjectively risky if for all outcomes $x, y \in X$ and for all acts $f, g \in \mathcal{F}$,*

(i) $xEf \succeq xEg \quad \Rightarrow \; yEf \succeq yEg$; *and*

(ii) $fEx \succeq gEx \quad \Rightarrow \; fEy \succeq gEy$.

Otherwise, call E subjectively ambiguous.

This definition takes (complementary) events E and $\neg E$ to be subjectively risky if preference is separable across these events and can be conditioned

(i) on $\neg E$ independently of the outcome obtained on E;

(ii) on E independently of the outcome obtained on $\neg E$.

Therefore, whether an event E is subjectively risky depends exclusively on the ranking of acts that are *constant* on E or on $\neg E$. This constancy reflects the intuition that an event being subjectively risky does not imply the same property for subsets of this event. Note that Zhang's definition remains intuitive even if the decision maker is not ambiguity averse as in the Ellberg Paradox.

Let $\mathcal{R}_Z \subset \Sigma$ denote Zhang's class of subjectively risky events and let $\mathcal{G}_Z \subset \mathcal{F}$ denote the associated set of subjectively risky ($\mathcal{R}_Z$-measurable) acts. Both $\mathcal{R}_Z$ and $\mathcal{G}_Z$ are uniquely derived from the preference $\succeq$. Moreover, if $\succeq$ is reflexive, then $\mathcal{R}_Z$ is a *mosaic*, and P2($\mathcal{R}_Z$) holds on $\mathcal{G}_Z$. (P2($\mathcal{R}_Z$) follows immediately from the definition of $\mathcal{R}_Z$.)

Show that $\mathcal{R}_Z$ satisfies the properties of a mosaic.

($*$) $S \in \mathcal{R}_Z$ because $\succeq$ is reflexive and hence, $x \succeq x$ for all $x \in X$.

($**$) By definition, $E \in \mathcal{R}_Z$ is equivalent to $\neg E \in \mathcal{R}_Z$.

(μ) Partition S into $m \geq 2$ subjectively risky events $\{S_1, \ldots, S_m\} \subset \mathcal{R}_Z$. Then for all outcomes $x, y \in X$ and acts $f, g \in \mathcal{F}$,

$$\begin{aligned}
x(S_1 \cup S_2)f \succeq x(S_1 \cup S_2)g \quad &\Rightarrow \quad xS_1(xS_2f) \succeq xS_1(xS_2g) \underset{S_1 \in \mathcal{R}_Z}{\Rightarrow} \\
yS_1(xS_2f) \succeq yS_1(xS_2g) \quad &\Rightarrow \quad xS_2(yS_1f) \succeq xS_2(yS_1g) \underset{S_2 \in \mathcal{R}_Z}{\Rightarrow} \\
yS_2(yS_1f) \succeq yS_2(yS_1g) \quad &\Rightarrow \quad y(S_1 \cup S_2)f \succeq y(S_1 \cup S_2)g.
\end{aligned}$$

A similar argument repeated for $S_3, \ldots, S_m \in \mathcal{R}_{EZ}$ shows that for all outcomes $x, y \in X$ and acts $f, g \in \mathcal{F}$,

$$x(\cup_{i=3}^m S_i)f \succeq x(\cup_{i=3}^m S_i)g \quad \Rightarrow \; y(\cup_{i=3}^m S_i)f \succeq y(\cup_{i=3}^m S_i)g,$$

or equivalently,

$$f(S_1 \cup S_2)x \succeq g(S_1 \cup S_2)x \quad \Rightarrow \; f(S_1 \cup S_2)y \succeq g(S_1 \cup S_2)y.$$

Thus, $S_1 \cup S_2 \in \mathcal{R}_Z$.

In general, the class $\mathcal{R}_Z$ need not be an algebra. To illustrate this point, we adopt Zhang's four-color setting. For simplicity, assume that X has only two elements, $x \succ x'$. Let $S = \{B, G, R, Y\}$, and suppose that the decision maker is told only that the total number of balls is 100, and that $B + G = B + Y = 60$. Based on this information about the composition of the urn, one can evaluate the following lower bounds for probabilities of events:

- 0.6 for $\{B, G, R\}$, $\{B, R, Y\}$, $\{B, G, Y\}$, $\{B, G\}$, and $\{B, Y\}$;
- 0.4 for $\{G, R, Y\}$, $\{G, R\}$, and $\{R, Y\}$;
- 0.2 for $\{B, R\}$ and $\{B\}$;
- 0 for $\{G, Y\}$, $\{G\}$, $\{R\}$, and $\{Y\}$.

Consider a ranking that reflects the above evaluations:

$$x \succ x\{B,G,R\}x' \sim x\{B,R,Y\}x' \sim x\{B,G,Y\}x' \sim x\{B,G\}x' \sim x\{B,Y\}x' \succ$$
$$x\{G,R,Y\}x' \sim x\{G,R\}x' \sim x\{R,Y\}x' \succ$$
$$x\{B,R\}x' \sim x\{B\}x' \succ$$
$$x\{G,Y\}x' \sim x\{G\}x' \sim x\{R\}x' \sim x\{Y\}x' \sim x'.$$

For this ranking, $\mathcal{R}_Z = \{\emptyset, \{B,G\}, \{B,Y\}, \{R,Y\}, \{G,R\}, \{B,G,R,Y\}\}$ is not closed under intersections. (In this example, $\mathcal{R}_Z$ coincides with the class of events for which probabilities are given explicitly, but in general, this need not be so.)

In order to obtain an expected utility representation on the mosaic $\mathcal{G}_Z$ of subjectively risky acts, one can use Theorem 3.1.

Corollary 3.2. *Let $\succeq$ be a reflexive binary relation on $\mathcal{F}$, and let $\mathcal{R}_Z$ be the class of subjectively risky events defined by Zhang. Then the following statements are equivalent.*

1. *$\succeq$ satisfies P1($\mathcal{R}_Z$), P3($\mathcal{R}_Z$), P4($\mathcal{R}_Z$), P5($\mathcal{R}_Z$), P6($\mathcal{R}_Z$).*

2. *$\succeq$ is represented on the set $\mathcal{G}_Z$ by expected utility*

$$U(f) = \sum_{x \in X} u(x) \cdot p\left(f^{-1}(x)\right) \qquad \text{for } f \in \mathcal{G}_Z, \tag{3.6}$$

where $u : X \to \mathbb{R}$ is a non-constant utility index, and $p : \mathcal{R}_Z \to [0,1]$ is a finely ranged probability measure.

In this representation, the index u is unique up to a positive linear transformation, and the probability measure p is unique.

Note that all components of representation (3.6)—the domains $\mathcal{R}_Z$ and $\mathcal{G}_Z$, the probability measure p, and the utility index u—are derived from preference. Therefore, Corollary 3.2 provides a *fully subjective* theory of expected utility. One special case of this theory is Savage's result; in this case, P2 implies that $\mathcal{R}_Z = \Sigma$ and that representation (3.6) holds on the set $\mathcal{F}$ of all acts, where it is equivalent to (3.3). More generally, the fully subjective theory applies when P2 does not hold. Then, preference is still separable across subjectively risky events, these events are assigned probabilities, and the associated subjectively risky acts are ranked via expected utility. Thus, axioms P1($\mathcal{R}_Z$), P3($\mathcal{R}_Z$), P4($\mathcal{R}_Z$), P5($\mathcal{R}_Z$), and P6($\mathcal{R}_Z$) provide foundations for the distinction between subjectively risky events in $\mathcal{R}_Z$ and all other (subjectively ambiguous) events.

In general, $\mathcal{R} = \mathcal{R}_Z$ is not the unique (or even the largest) mosaic $\mathcal{R}$ such that the preference $\succeq$ satisfies P1($\mathcal{R}$)–P6($\mathcal{R}$). The selection of $\mathcal{R}_Z$ rather than of any such $\mathcal{R}$ as *the* domain of subjectively risky events is motivated by Zhang's explicit and intuitive behavioral definition.

4 Probabilistic Sophistication

In order to accommodate Allais-type behavior, Machina and Schmeidler [16] extend Savage's theory. They model a probabilistically sophisticated decision maker who ranks acts in two stages: first, she uses subjective probabilities to reduce each act to a lottery—a distribution over outcomes—and then she ranks the induced lotteries via a *risk preference* which may have no expected utility representation.

The fundamental difference between expected utility maximization and probabilistic sophistication is that the former implies a strong form of separability across mutually exclusive events (such as P2 in Savage's framework), while the latter does not. Motivated by this observation, Machina and Schmeidler relax the Sure-Thing Principle.

Axiom P4* (Strong Comparative Probability). *For all partitions $E = A \cup B$, for all outcomes $x \succ x'$ and $z \succ z'$, and for all acts $h, h' \in \mathcal{F}$,*

$$(xAx')Eh \succeq (xBx')Eh \quad \Rightarrow \quad (zAz')Eh' \succeq (zBz')Eh'. \tag{4.1}$$

P4* states that the preference to bet on the event A rather than on the (disjoint) event B is independent of the stakes that are involved in such bets and of the outcomes that are obtained if neither A nor B occurs.[13] The invariable preference to bet on A rather than on B reflects exclusively the decision maker's belief that A is at least as probable as B. Accordingly, P4* does not rule out non-linear risk preferences.[14]

Savage's axioms other than P2 remain intuitive for probabilistically sophisticated behavior. This motivates adopting P1, P3, P4*, P5, P6 as an axiomatic foundation for probabilistic sophistication. Before stating Machina–Schmeidler's representation result, we need a few preliminaries.

A *lottery* $l : X \to [0,1]$ is a probability distribution that has a finite support in X. For every $Y \subset X$, let $l(Y) = \sum_{x\in Y} l(x)$. Denote by $\mathcal{L} = \{l, \dots\}$ the set of lotteries endowed with the metric

$$\|l - l'\| = \sum_{x\in X} |l(x) - l'(x)| \qquad \text{for } l, l' \in \mathcal{L}. \tag{4.2}$$

Define a *mixture* $\tau l + (1-\tau)l'$ of lotteries l and l' with a weight $\tau \in [0,1]$ by

$$(\tau l + (1-\tau)l')(x) = \tau l(x) + (1-\tau)l'(x) \qquad \text{for } x \in X.$$

Call a function $V : \mathcal{L} \to \mathbb{R}$ *mixture continuous* if for all lotteries $l, l' \in \mathcal{L}$, $V(\tau l + (1-\tau)l')$ is continuous with respect to the weight $\tau \in [0,1]$.

Given a ranking of outcomes $\succeq_X$, define a notion of first-order stochastic dominance in $\mathcal{L}$. For all $x \in X$, let

$$Y_x = \{y \in X : y \succeq_X x\}.$$

Say that a lottery l *weakly dominates* l', written $l \geqslant l'$, if $l(Y_x) \geq l'(Y_x)$ for all $x \in X$. Say that l *strictly dominates* l', written $l \gg l'$, if $l(Y_x) \geq l'(Y_x)$ for all

[13]The combination of P2 and P4 implies P4* as follows:

$$\begin{aligned}(xAx')Eh \succeq (xBx')Eh \quad &\underset{\text{P2}}{\Rightarrow} \quad (xAx')Ex' \succeq (xBx')Ex' \quad \underset{\text{P4}}{\Rightarrow} \\ &(zAz')Ez' \succeq (zBz')Ez' \quad \underset{\text{P2}}{\Rightarrow} \quad (zAz')Eh' \succeq (zBz')Eh'.\end{aligned}$$

Conversely, P4* implies P4 as a special case for $E = S$, but does not imply P2 (providing that X has at least three elements).

[14]P4* requires only that the risk preferences are monotonic.

$x \in X$ and $l(Y_x) > l'(Y_x)$ for some $x \in X$. Call a function $V : \mathcal{L} \to \mathbb{R}$ *weakly monotonic* if for all $l, l' \in \mathcal{L}$,

$$l \geqslant l' \quad \Rightarrow \quad V(l) \geq V(l'); \tag{4.3}$$

call V *strictly monotonic* if for all $l, l' \in \mathcal{L}$, V satisfies (4.3) and

$$l \gg l' \quad \Rightarrow \quad V(l) > V(l'). \tag{4.4}$$

Throughout, we will use first-order stochastic dominance and the associated notions of monotonicity without specifying the ranking $\succeq_X$ explicitly; this ranking will be clear from the context.

Given a probability measure $p : \Sigma \to [0,1]$ and an act $f \in \mathcal{F}$, define a lottery $[f]_p \in \mathcal{L}$ by

$$[f]_p(x) = p(f^{-1}(x)) \quad \text{for } x \in X;$$

say that $[f]_p$ is *induced by f*.

Finally, assume that the algebra Σ of events is a σ-algebra. Even though this assumption is not required in order to state P4* (and all the other axioms), it is essential for the construction of subjective probabilities used by Machina–Schmeidler.[15]

Theorem (Machina–Schmeidler). *Let Σ be a σ-algebra. Then the following two statements are equivalent.*

1. *$\succeq$ satisfies P1, P3, P4*, P5, P6.*

2. *$\succeq$ is represented by*

$$U(f) = V([f]_p) \qquad \text{for } f \in \mathcal{F}, \tag{4.5}$$

where $p : \Sigma \to [0,1]$ is a convex-ranged probability measure, and the utility function $V : \mathcal{L} \to \mathbb{R}$ is non-constant, strictly monotonic, and mixture continuous.

[15]The construction is basically the same as in Savage.

In this representation, the probability measure p is unique.

The decision maker as portrayed by (4.5) assigns subjective probabilities to all events $A \in \Sigma$ and translates all acts $f \in \mathcal{F}$ into lotteries $[f]_p$ which she ranks via a mixture continuous and strictly monotonic function V. In other words, V represents the decision maker's risk preference over lotteries. The Savage Theorem accommodates the special case where V has the expected utility form but in general, the risk preference may have no expected utility representation.[16] Thus, P1, P3, P4*, P5, P6 relax Savage's axioms and provide foundations for probabilistic sophistication on the set $\mathcal{F}$ of all acts. However, Ellsberg-type behavior is not probabilistically sophisticated and violates Strong Comparative Probability. This motivates restricting P4* and the associated characterization of probabilistic sophistication to the domain of risk.

4.1 Probabilistic Sophistication on Risky Acts

Next, we rewrite P4* for risky ($\mathcal{R}$-measurable) acts in order to reflect the use of subjective probabilities for risky events without restricting choice among ambiguous acts.

Axiom P4*($\mathcal{R}$). *For all risky partitions $E = A \cup A' = B \cup B'$, for all outcomes $x \succ x'$ and $z \succ z'$, and for all acts $h, h' \in \mathcal{G} \Cap \{\neg E\}$,*

$$(xAx')Eh \succeq (xBx')Eh \quad \Rightarrow \quad (zAz')Eh' \succeq (zBz')Eh'. \tag{4.6}$$

P4*($\mathcal{R}$) states that the preference to bet on the risky event A rather than on the risky event B is independent of the stakes that are involved in such bets and of the ($\mathcal{R}$-measurable) subact that is obtained on the risky event $\neg E$ when neither A nor B occurs.[17] Accordingly, P4*($\mathcal{R}$) does not restrict preference to bet on

[16]See Grant [10] for a model of probabilistic sophistication where even monotonicity of risk preference is relaxed.

[17]Note that unlike Machina–Schmeidler's P4*, P4*($\mathcal{R}$) does not require that A and B are disjoint. If $\mathcal{R}$ is an algebra, then the two formulations are equivalent. In general, the requirement

ambiguous events. Neither does P4*($\mathcal{R}$) rule out non-linear risk preferences over lotteries induced by risky acts.

In order to state our second main representation, we need a few preliminaries. Given a finite subset $Y \subset X$, denote by $\mathcal{L}(Y)$ the set of lotteries in $\mathcal{L}$ that have support in Y. Denote by $\mathcal{L}_p$ the set of lotteries that are induced by risky acts via a probability measure $p : \mathcal{R} \to [0, 1]$:

$$\mathcal{L}_p = \{l \in \mathcal{L} \,:\, l = [f]_p \quad \text{for some } f \in \mathcal{G}\}.$$

Consider a binary relation $\succeq_1$ on $\mathcal{L}_p$. Call $\succeq_1$ *strictly monotonic* if for all lotteries $l, l' \in \mathcal{L}_p$, $l \geqslant l'$ and $l \gg l'$ imply $l \succeq_1 l'$ and $l \succ_1 l'$ respectively. Call $\succeq_1$ *continuous* if for all finite $Y \subset X$ and for all lotteries $l \in \mathcal{L}(Y) \cap \mathcal{L}_p$, the sets $\{l' \in \mathcal{L}(Y) \cap \mathcal{L}_p \,:\, l' \succeq_1 l\}$ and $\{l' \in \mathcal{L}(Y) \cap \mathcal{L}_p \,:\, l' \preceq_1 l\}$ are closed in $\mathcal{L}(Y) \cap \mathcal{L}_p$.

Theorem 4.1 (Part I). *Let Σ be an algebra, and let $\mathcal{R} \subset \Sigma$ be a mosaic. Then the following two statements are equivalent.*

1. *$\succeq$ satisfies P1($\mathcal{R}$), P3($\mathcal{R}$), P4*($\mathcal{R}$), P5($\mathcal{R}$), P6($\mathcal{R}$).*

2. *$\succeq$ on the set $\mathcal{G}$ is represented by*

$$f \succeq g \quad \Leftrightarrow \quad [f]_p \succeq_1 [g]_p \qquad \text{for } f, g \in \mathcal{G}, \tag{4.7}$$

where $p : \mathcal{R} \to [0, 1]$ is a finely ranged probability measure, and the binary relation $\succeq_1$ on $\mathcal{L}_p$ is non-degenerate, complete, transitive, continuous, and strictly monotonic.

---

that A and B are disjoint is not required in order to motivate the axiom, but leads to unnecessary technical complications. Note also that P4*($\mathcal{R}$) preserves all the natural logical connections to other axioms. P4*($\mathcal{R}$) is implied by the combination of P2($\mathcal{R}$) and P4($\mathcal{R}$). Conversely, P4*($\mathcal{R}$) implies P4($\mathcal{R}$) as a special case for $E = S$, but does not imply P2($\mathcal{R}$) (providing that X has at least three elements). Finally, P4* implies P4*($\mathcal{R}$) and is equivalent to P4*(Σ). In order to prove this, let $E' = E \cup (A \cap B)$, $h'' = x(A \cap B)h$ and $h''' = x'(A \cap B)h'$; then

$$(xAx')Eh \succeq (xBx')Eh \quad \Rightarrow \quad (x(A \setminus B)x')E'h'' \succeq (x(B \setminus A)x')E'h'' \quad \Rightarrow \{\text{P4}^*\}$$
$$(z(A \setminus B)z')E'h''' \succeq (z(B \setminus A)z')E'h''' \quad \Rightarrow \quad (zAz')Eh' \succeq (zBz')Eh'.$$

On the other hand, P4* is a special case of P4*(Σ) where $A = B'$ and $B = A'$.

In this representation, the probability measure p and the strictly monotonic binary relation $\succeq_1$ are unique.

Similar to Theorem 3.1, Theorem 4.1(I) models behavior only for risky acts, and derives subjective probabilities only for risky events. In representation (4.7), the probabilities are used to translate risky acts $f \in \mathcal{G}$ into lotteries $[f]_p \in \mathcal{L}_p$, and the induced lotteries are ranked via the continuous and strictly monotonic weak order $\succeq_1$ rather than via expected utility. Therefore, P1($\mathcal{R}$), P3($\mathcal{R}$), P4*($\mathcal{R}$), P5($\mathcal{R}$), P6($\mathcal{R}$) provide foundations for probabilistic sophistication on risky acts.

If $\mathcal{R} = \Sigma$ and Σ is a σ-algebra, then Theorem 4.1(I) becomes a reduced version of the Machina–Schmeidler Theorem. In this case, both results use equivalent axioms, and the function V in (4.5) represents the risk preference $\succeq_1$ in (4.7). Moreover, Theorem 4.1(I) shows that Machina–Schmeidler's axioms *taken "as is"* characterize a version of probabilistic sophistication over Σ-measurable acts even if the algebra Σ does not satisfy (σ).

In general, the axioms in Theorem 4.1(I) are parallel to Machina–Schmeidler's counterparts but apply only to risky acts and events. Most importantly, P4*($\mathcal{R}$) does not restrict the preference to bet on ambiguous events. Accordingly, only risky events $A \in \mathcal{R}$ are assigned subjective probabilities, only risky acts $f \in \mathcal{G}$ are translated into lotteries, and only lotteries $[f]_p \in \mathcal{L}_p$ that are induced by risky acts are ranked by the risk preference $\succeq_1$. Finally, in (4.7), the probability measure p need not be convex-ranged, and no utility representation V is specified for the risk preference $\succeq_1$. Thus, representation (4.7) constitutes a notion of probabilistic sophistication different from, albeit closely related to the one due to Machina–Schmeidler. In (4.7), the risk preference is retained as a binary relation $\succeq_1$ on a subdomain $\mathcal{L}_p \subset \mathcal{L}$, while in (4.5) the risk preference on the set $\mathcal{L}$ of all lotteries is represented by a utility function V. It seems intuitive that risk preference in the general case might be described by a binary relation rather than by a utility function. In fact, any utility representation for the risk preference might restrict the use of probabilities and hence, be viewed as excess baggage for a general theory of probabilistic sophistication.

On the other hand, a utility function for the ranking $\succeq_1$ may be desirable for analytical tractability. It is an open question whether conditions of Theorem 4.1(I) imply that such a utility function exists. In particular, one cannot use Machina–Schmeidler's construction of V (see an example in Section 5.2).

One way to obtain a utility representation for the risk preference $\succeq_1$ is to impose a stronger continuity axiom on the underlying preference over risky acts.

Axiom P6*($\mathcal{R}$) (Small Event Uniform Continuity). *For any outcome x, for any finite collection $\mathcal{E} \subset \mathcal{R}$, for any $\mathcal{E}$-measurable acts $f \succ g$, there exists a risky partition $\{S_1, \ldots, S_m\} \subset \mathcal{R} \Cap \mathcal{E}$ of S such that for all $i = 1 \ldots m$ and for all risky acts $h \in \mathcal{G} \Cap \{\neg S_i\}$,*

$$\begin{aligned} h(S) \subset f(S) \text{ and } h \succeq f \quad &\Rightarrow \quad xS_ih \succ g, \\ h(S) \subset g(S) \text{ and } h \preceq g \quad &\Rightarrow \quad xS_ih \prec f. \end{aligned}$$

P6($\mathcal{R}$) is a special case of P6*($\mathcal{R}$) where $h = f$ or $h = g$. Both axioms require that S can be partitioned into (sufficiently) small risky events $S_1, \ldots, S_m$ that belong to the subclass $\mathcal{R} \Cap \mathcal{E} \subset \mathcal{R}$. However, the two axioms use slightly different notions of small events. In P6*($\mathcal{R}$), S_i is small if

(i) not only for $h = f$ (as in P6($\mathcal{R}$)) but for any risky act $h \succeq f$ that has range $h(S) \subset f(S)$ and is $\mathcal{R}$-measurable on $\neg S_i$, replacing $h(s)$ by x on S_i does not reverse the strict preference $h \succ g$;[18]

(ii) not only for $h = g$ (as in P6($\mathcal{R}$)) but for any risky act $h \preceq g$ that has range $h(S) \subset g(S)$ and is $\mathcal{R}$-measurable on $\neg S_i$, replacing $h(s)$ by x on S_i does not reverse the strict preference $f \succ h$.

In P6*($\mathcal{R}$), the act h is allowed to vary while the partition $S = \cup_{i=1}^m S_i$ is fixed. In this way, P6*($\mathcal{R}$) reflects an intuitive notion of uniform continuity of the preference $\succeq$ with respect to small risky events. Note that both P6($\mathcal{R}$) and P6*($\mathcal{R}$) require that small events belong to the subclass $\mathcal{R} \Cap \mathcal{E}$. This requirement is intuitive if

[18]The range of h is restricted to a finite set in order to avoid situations when the fixed outcome x can replace unboundedly good outcomes $h(s)$ on $s \in S_i$.

the description of the world includes infinitely many coin tosses that are risky and independent of all events in $\mathcal{E}$.

Call a function $V : \mathcal{L}_p \to \mathbb{R}$ *uniformly continuous* if for each finite $Y \subset X$, V is uniformly continuous on $\mathcal{L}_p \cap \mathcal{L}(Y)$.

Theorem 4.1 (Part II). *Let Σ be an algebra and let $\mathcal{R} \subset \Sigma$ be a mosaic. Then the following two statements are equivalent.*

1. *$\succeq$ satisfies P1($\mathcal{R}$), P3($\mathcal{R}$), P4***($\mathcal{R}$), P5($\mathcal{R}$), P6***($\mathcal{R}$).*

2. *$\succeq$ on the set $\mathcal{G}$ is represented by*

$$U(f) = V([f]_p) \qquad \text{for } f \in \mathcal{G}, \tag{4.8}$$

where $p : \mathcal{R} \to [0,1]$ is a finely ranged probability measure, and the utility function $V : \mathcal{L}_p \to \mathbb{R}$ is non-constant, strictly monotonic, and uniformly continuous.

In this representation, the probability measure p is unique.

Therefore, strengthening P6($\mathcal{R}$) to P6*($\mathcal{R}$) is sufficient for the risk preference $\succeq_1$ to have a utility representation V. Note that the expected utility functional form is strictly monotonic and uniformly continuous; therefore, the expected utility representation (3.5) is a special case of (4.8). Another special case of (4.8) is Machina–Schmeidler's (4.5).[19] Similar to (4.5), the utility representation (4.8) on $\mathcal{G}$ need not have any particular parametric form.

4.2 Fully Subjective Probabilistic Sophistication

Theorem 4.1 derives subjective probabilities on the class of risky events $\mathcal{R}$ that is exogenous to the model. The exogenous formulation may seem problematic if decision makers may disagree about the identity of events to which they assign

[19] If the function $V : \mathcal{L} \to [0,1]$ is strictly monotonic, then the properties of mixture continuity and uniform continuity are equivalent.

probabilities. This motivates defining risky events subjectively, that is, in terms of preference.

Section 3.2 provides one such definition, due to Zhang, that takes an event $E \in \Sigma$ to be subjectively risky if for all outcomes $x, y \in X$ and for all acts $f, g \in \mathcal{F}$,

$$\begin{aligned} xEf \succeq xEg \quad &\Rightarrow \quad yEf \succeq yEg; \quad \text{and} \\ fEx \succeq gEx \quad &\Rightarrow \quad fEy \succeq gEy. \end{aligned}$$

In general, these requirements may be too demanding. For example, the decision maker may be probabilistically sophisticated on the domain $\mathcal{F}$ of all acts but has a non-linear risk preference. In this case, subjective probabilities for all events can be derived from preference in Σ but some events will not satisfy Zhang's definition.

Consider instead the following definition that is motivated by Strong Comparative Probability rather than by the Sure-Thing Principle. Call acts $f, g \in \mathcal{F}$ *complementary bets* if there exists a partition $E = A \cup B$, outcomes $z \succ z'$, and an act $h \in \mathcal{F}$ such that $f = (zAz')Eh$ and $g = (zBz')Eh$.

Definition (Epstein–Zhang). *An event $E \in \Sigma$ is called subjectively risky if for all outcomes $x, y \in X$, and for all acts $f, g \in \mathcal{F}$ that are complementary bets,*

(i) $xEf \succeq xEg \quad \Rightarrow \quad yEf \succeq yEg;$ *and*

(ii) $fEx \succeq gEx \quad \Rightarrow \quad fEy \succeq gEy.$

Otherwise, E is called subjectively ambiguous.

In other words, this definition takes events E and $\neg E$ to be subjectively risky if for all events $A, B \in \Sigma$,

(i) the preference to bet on $A \cap \neg E$ rather than on $B \cap \neg E$ is independent of the outcome obtained on E;

(ii) the preference to bet on $A \cap E$ rather than on $B \cap E$ is independent of the outcome obtained on $\neg E$.

This definition does not require separability across E and $\neg E$ for preference over acts other than complementary bets because such a stronger form of separability may reflect a non-linear risk preference rather than the decision maker's unwillingness to assign probabilities.[20] Note that Epstein–Zhang's definition uses only acts that are constant on E or alternatively, on $\neg E$. Therefore, subsets of a subjectively risky event may be subjectively ambiguous.

Denote by $\mathcal{R}_{EZ} \subset \Sigma$ and $\mathcal{G}_{EZ} \subset \mathcal{F}$ the class of subjectively risky events defined by Epstein–Zhang and the associated set of subjectively risky ($\mathcal{R}_{EZ}$-measurable) acts. The domains $\mathcal{R}_{EZ}$ and $\mathcal{G}_{EZ}$ are uniquely determined by preference. By definition, $\mathcal{R}_{EZ}$ is a *mosaic* (the proof is the same as in Zhang's case, the only difference being that f and g are required everywhere to be complementary bets), and P4($\mathcal{R}_{EZ}$) implies P4*($\mathcal{R}_{EZ}$).[21] The four-color example in Section 3.2 illustrates that $\mathcal{R}_{EZ}$ need not be an algebra.

Thus, in order to characterize probabilistic sophistication on subjectively risky acts in $\mathcal{G}_{EZ}$, one can use Theorem 4.1.

Corollary 4.2. *Let $\succeq$ be a reflexive binary relation on $\mathcal{F}$, and let $\mathcal{R}_{EZ}$ be the mosaic of subjectively risky events defined by Epstein–Zhang. Then the following two statements are equivalent.*

1. *$\succeq$ satisfies P1($\mathcal{R}_{EZ}$), P3($\mathcal{R}_{EZ}$), P4($\mathcal{R}_{EZ}$), P5($\mathcal{R}_{EZ}$), P6($\mathcal{R}_{EZ}$);*

2. *$\succeq$ is represented on the set $\mathcal{G}_{EZ}$ by*

$$f \succeq g \quad \Leftrightarrow \quad [f]_p \succeq_1 [g]_p \qquad \text{for } f, g \in \mathcal{G}_{EZ}, \tag{4.9}$$

[20]Epstein–Zhang' definition relaxes Zhang's; however, the two definitions are equivalent if X has only two elements $x \succ x'$, in which case any acts $f, g \in \mathcal{F}$ are complementary bets.

[21]For all events $E \in \mathcal{R}_{EZ}$, for all subjectively risky partitions $E = A \cup A'$ and $E = B \cup B'$, for all outcomes $x \succ x'$ and $z \succ z'$, and for all acts $h, h' \in \mathcal{G}_{EZ} \Cap \{E\}$, use Epstein–Zhang's definition and P4($\mathcal{R}_{EZ}$) to obtain:

$$(xAx')Eh \succeq (xBx')Eh \quad \Rightarrow \quad (xAx')Ex' \succeq (xBx')Ex' \quad \Rightarrow \{\text{P4}(\mathcal{R}_{EZ})\}$$
$$(zAz')Ez' \succeq (zBz')Ez' \quad \Rightarrow \quad (zAz')Eh' \succeq (zBz')Ez'.$$

where $p : \mathcal{R}_{EZ} \to [0,1]$ is a finely ranged probability measure, and the binary relation $\succeq_1$ on $\mathcal{L}_p$ is non-degenerate, complete, transitive, continuous, and strictly monotonic.

In this representation, the probability measure p and the strictly monotonic binary relation $\succeq_1$ are unique.

Moreover, $\succeq$ satisfies the additional axiom P6($\mathcal{R}_{EZ}$) if and only if the risk preference $\succeq_1$ is represented by a utility function $V : \mathcal{L}_p \to \mathbb{R}$ that is strictly monotonic and uniformly continuous.*

Note that all components of representation (4.9)—the domains $\mathcal{R}_{EZ}$ and $\mathcal{G}_{EZ}$, the probability measure p, and the risk preference $\succeq_1$—are derived from preference. Therefore, Corollary 4.2 provides a *fully subjective* theory of probabilistic sophistication. One special case of this theory is Machina–Schmeidler's result; in this case, P4* implies that $\mathcal{R}_{EZ} = \Sigma$ and that representation (4.9) holds on the set $\mathcal{F}$ of all acts, where it is equivalent to (4.5). More generally, the theory applies when P4* does not hold. In these situations, subjectively risky events in $\mathcal{R}_{EZ}$ are assigned probabilities, and the associated subjectively risky acts in $\mathcal{G}_{EZ}$ are translated into lotteries that are ranked by a continuous and monotonic risk preference. Thus, axioms P1($\mathcal{R}_{EZ}$), P3($\mathcal{R}_{EZ}$), P4($\mathcal{R}_{EZ}$), P5($\mathcal{R}_{EZ}$), and P6($\mathcal{R}_{EZ}$) provide foundations for the distinction between subjectively risky events in $\mathcal{R}_{EZ}$ and all other (subjectively ambiguous) events. This distinction, unlike the one characterized in Section 3.2, does not exclude Allais-type behavior.

Besides $\mathcal{R}_{EZ}$, there may be other mosaics $\mathcal{R}$ such that preference $\succeq$ satisfies P1($\mathcal{R}$), P3($\mathcal{R}$), P4*($\mathcal{R}$), P5($\mathcal{R}$), and P6($\mathcal{R}$). The selection of $\mathcal{R}_{EZ}$ rather than of any such $\mathcal{R}$ as *the* domain of subjectively risky events is motivated by Epstein–Zhang's explicit and intuitive behavioral definition.

To relate Corollary 4.2 to the main representation result in Epstein–Zhang, consider a special case of mosaics. Call $\mathcal{R}_{EZ}$ a (finitely additive) λ-*system* if it satisfies

(λ) $\quad A_1 \in \mathcal{R}_{EZ}$, $A_2 \in \mathcal{R}_{EZ}$ and $A_1 \cap A_2 = \emptyset \quad \Rightarrow \; A \cup B \in \mathcal{R}_{EZ}$.

The property (λ) strengthens (μ) and requires that $\mathcal{R}_{EZ}$ is closed under arbitrary *disjoint* unions. Without additional assumptions, the class $\mathcal{R}_{EZ}$ (or similarly, $\mathcal{R}_Z$) may not be a λ-system. For example, let $S = \{B, G, R\}$, $X = \{x, x'\}$, and let the decision maker have the following ranking:

$$x \succ x\{G, R\}x' \succ x\{B, R\}x' \succ x\{B, G\}x' \succ x\{G\}x' \succ x\{R\}x' \succ x\{B\}x' \succ x'.$$

Then the complementary events $\{B\}$ and $\{G, R\}$ are subjectively ambiguous because of the preference reversal from $x\{G\}x' \succ x\{R\}x'$ to $x\{B, R\}x' \succ x\{B, G\}x'$. However, the events $\{G\}$ and $\{B, R\}$, $\{R\}$ and $\{B, G\}$ are subjectively risky. It follows that

$$\mathcal{R}_Z = \mathcal{R}_{EZ} = \{\emptyset, \{B, R\}, \{B, G\}, \{R\}, \{G\}, \{B, G, R\}\}$$

is not closed under disjoint unions. In this example, the class $\mathcal{R}_{EZ}$ is a mosaic but not a λ-system.

It is arguably intuitive that the decision maker, who assigns probabilities $p(A)$ and $p(B)$ to disjoint events A and B, also assigns the sum $p(A) + p(B)$ to the union $A \cup B$. Motivated by this intuition, Epstein and Zhang impose additional conditions on preference that guarantee that $\mathcal{R}_{EZ}$ is a λ-system.[22] It is, however, an open question whether the simple intuition underlying the λ-system structure for risky events has an equally simple behavioral foundation.

Besides the difference in structure for the class of risky events, Corollary 4.2 differs substantially from Epstein–Zhang in the axioms adopted. Roughly, the corollary dispenses with two of their arguably least appealing axioms (Monotone Continuity and Strong Partition Neutrality).

5 Sketch of Proofs

In this section we discuss the main aspects that differentiate the proofs of Theorems 3.1 and 4.1 from those of Savage and Machina–Schmeidler. First, we focus

[22]In fact, their conditions are too weak and their theorem is not valid as stated. To correct Epstein–Zhang's theorem, strengthen their Axiom 4 so that, in their notation, A_n and B_m can vary over all of Σ but not necessarily over $\mathcal{A}$.

on the construction of subjective probabilities for risky events.

5.1 Construction of Subjective Probability

Analogously to de Finetti [3] and Savage [20], we seek a probability measure p that represents the *comparative likelihood* relation $\succeq_0$. For arbitrary risky events A and B, the relation $A \succeq_0 B$ is defined in terms of preference by:

$$A \succeq_0 B \quad \Leftrightarrow \quad xAx' \succeq xBx' \text{ for all outcomes } x \succ x',$$

and is interpreted as A being subjectively at least as likely as B. The list of axioms imposed in Theorem 3.1, or alternatively, in Theorem 4.1, implies that the relation $\succeq_0$ has properties analogous to those of de Finetti's *qualitative probability* and to Savage's *fineness* and *tightness*. However, the fact that the mosaic $\mathcal{R}$ is not closed under intersections complicates the formulation and the formal proof of these properties in Appendix.

The construction of subjective probability by Savage (or by Fishburn [7]) relies on partitioning the state space S into an arbitrarily large number N of equiprobable events. Such equipartitions need not exist if the domain of risky events is a mosaic or even a finitely additive algebra (see an example in Section 5.3). In the absence of equipartitions, we propose the following construction of the unique finely ranged probability measure $p : \mathcal{R} \to [0, 1]$ that represents $\succeq_0$ on $\mathcal{R}$. Like Savage's, our construction is explicit and captures a simple intuition.

Fix an arbitrary risky event $A \in \mathcal{R}$. Say that a risky partition $S = \cup_{i=1}^m S_i$ is *finer* than A if $A \succ_0 S_i$ for all $i = 1 \ldots m$, that is, if every S_i is subjectively less likely than A. Among all risky partitions of S finer than A, take one that has a minimal number of elements. Let $\nu(A)$ be this minimal number; let $\nu(A) = +\infty$ if there are no risky partitions of S finer than A. The subjective probability $p(A)$ is constructed in terms of the function ν as follows:

$$p(A) = \sup_{\substack{A = A_1 \cup \cdots \cup A_n \\ A_i \in \mathcal{R} \text{ and } A_i \cap A_j = \emptyset}} \left\{ \sum_{i=1}^{n} \frac{1}{\nu(A_i)} \right\}. \tag{5.1}$$

In other words, the least upper bound is taken over all risky partitions of A.

To motivate formula (5.1), suppose that some finely ranged probability measure p does represent $\succeq_0$. Fix an arbitrary risky event $A \in \mathcal{R}$ with $p(A) > 0$. Let N be the unique integer such that

$$\tfrac{1}{N} < p(A) \leq \tfrac{1}{N-1}.$$

Then any risky partition of S finer than A must have at least N elements because all events in such a partition have probabilities that are strictly smaller than $\frac{1}{N-1}$ but add up to unity. On the other hand, fine-rangedness of p implies that there exists an N-element risky partition finer than A.[23] It follows that $\nu(A) = N$. Therefore,

$$\frac{1}{\nu(A)} < p(A) \leq \frac{1}{\nu(A) - 1}. \tag{5.2}$$

For the fixed risky event A, vary a number n and a risky partition $A = \cup_{i=1}^n A_i$. The sums $\sum_{i=1}^n \frac{1}{\nu(A_i)}$ are lower bounds for $p(A) = \sum_{i=1}^n p(A_i)$. Moreover, these sums become arbitrarily close to $p(A)$ for sufficiently fine risky partitions because (5.2) implies that the ratios $\frac{1}{\nu(A_i)}/p(A_i)$ approach 1 as the probabilities $p(A_i)$ approach zero. This argument suggests constructing $p(A)$ via formula (5.1).

Even though the above motivation for formula (5.1) is relatively transparent, it takes a lot of work in Appendix to show that the function $p : \mathcal{R} \to [0, 1]$ given by (5.1) is indeed a finely ranged probability measure that represents $\succeq_0$.

[23]By fine-rangedness, it is possible to partition S into risky events $\{S_1, \ldots, S_m\}$ such that $p(S_i) < p(A) - \frac{1}{N}$ for all $i = 1 \ldots m$. Construct indices $0 < k_1 < \cdots < k_{N-1} < m$ and risky events $E_1, \ldots, E_{N-1}, E_N$ by

$$S = \underbrace{S_1 \cup \cdots \cup S_{k_1}}_{E_1} \cup \cdots \cup \underbrace{S_{k_{N-2}+1} \cup \cdots \cup S_{k_{N-1}}}_{E_{N-1}} \cup \underbrace{S_{k_{N-1}+1} \cup \cdots \cup S_m}_{E_N}$$

so that for all $i = 1 \ldots N-1$, $\frac{1}{N} < p(E_i) < p(A)$. Then $p(E_N) < 1 - (N-1) \cdot \frac{1}{N} = \frac{1}{N} < p(A)$, and hence, the risky partition $S = \cup_{i=1}^N E_i$ is finer than A. Note that this argument relies on partitioning S into *approximately* equiprobable risky events E_i's. In this sense, it might reflect approximate symmetry considerations (see Savage [pp. 63–67] for discussion of the role of symmetry in the foundations of probability).

5.2 Utility Representations for Risk Preference

The next step after the construction of a subjective probability measure p is to translate the preference $\succeq$ over risky acts into a ranking $\succeq_1$ of lotteries and to obtain a suitable utility representation for $\succeq_1$. The natural definition of the risk preference $\succeq_1$ used by Savage and by Machina–Schmeidler can be adapted to our setting as follows:

$$l \succeq_1 l' \quad \Leftrightarrow \quad l = [f]_p \text{ and } l' = [g]_p \text{ for some risky acts } f \succeq g. \tag{5.3}$$

In order to find a representation for the ranking $\succeq_1$, one could attempt to use two well-known results, due to von Neumann–Morgenstern and Machina–Schmeidler respectively.

Say that $\succeq_1$ is *mixture continuous* if for all lotteries $l, l', l'' \in \mathcal{L}$, the sets

$$\{\tau \in [0,1] : \tau \cdot l + (1-\tau) \cdot l' \succeq_1 l''\} \text{ and } \{\tau \in [0,1] : \tau \cdot l + (1-\tau) \cdot l' \preceq_1 l''\}$$

are closed in $[0,1]$. Say that $\succeq_1$ is *mixture separable* if for all lotteries $l, l', l'' \in \mathcal{L}$,

$$l \succeq_1 l' \quad \Rightarrow \quad \tfrac{1}{2}l + \tfrac{1}{2}l'' \succeq_1 \tfrac{1}{2}l' + \tfrac{1}{2}l''. \tag{5.4}$$

Invariance (5.4) is a special case of von Neumann–Morgenstern's Independence Axiom.

1. A binary relation $\succeq_1$ on $\mathcal{L}$ is complete, transitive, mixture continuous, and mixture separable if and only if $\succeq_1$ is represented by expected utility

   $$U(l) = \sum_{x \in X} u(x) \cdot l(x) \qquad \text{for } l \in \mathcal{L},$$

   where $u : X \to \mathbb{R}$ is a utility index.

   Proof. See Herstein–Milnor [12]. □

2. A binary relation $\succeq_1$ on $\mathcal{L}$ is complete, transitive, mixture continuous, and strictly monotonic if and only if $\succeq_1$ is represented by $V : \mathcal{L} \to \mathbb{R}$ that is mixture continuous and strictly monotonic.

Proof. See Steps 4 and 5 in Machina–Schmeidler's proof of their Theorem 2. $\square$

Note that while the latter representation is more general than the former, both require that the relation $\succeq_1$ be complete on the set $\mathcal{L}$ of all lotteries. In the proofs of Savage and Machina–Schmeidler, where $\mathcal{R} = \Sigma$ is a σ-algebra and $p : \Sigma \to [0,1]$ is convex-ranged, the risk preference $\succeq_1$ defined by (5.3) is complete on $\mathcal{L}$ because for any $l \in \mathcal{L}$ there exists $f \in \mathcal{F}$ such that $l = [f]_p$. However, this argument fails in our model because the set $\mathcal{L}_p \subset \mathcal{L}$ of lotteries induced by risky acts may be smaller than $\mathcal{L}$ in which case the risk preference $\succeq_1$ defined by (5.3) is complete on $\mathcal{L}_p$ but not on all of $\mathcal{L}$.

In order to apply the above representations results, we extend $\succeq_1$ from $\mathcal{L}_p$, which is dense in $\mathcal{L}$, to all of $\mathcal{L}$ by continuity. We show that the list of axioms imposed in Theorem 3.1 guarantees that the risk preference $\succeq_1$ defined by (5.3) has a unique complete, transitive, mixture continuous, and mixture separable extension to all of $\mathcal{L}$. The von Neumann–Morgenstern theorem applied to this extension delivers the expected utility representation (3.5).

However, the weaker list of axioms imposed in Theorem 4.1(I) does not guarantee that the risk preference $\succeq_1$ defined by (5.3) has a complete, transitive, and mixture continuous extension to all of $\mathcal{L}$. For example, let $S = \prod_{k=1}^{\infty}\{H_k, T_k\}$. Identify every $A \subset \prod_{k=1}^{n}\{H_k, T_k\}$ with a natural event in S, and let Σ be the algebra of all such cylinders of finite ranks $n \geq 0$ (if $n = 0$, then $A = S$ or $A = \emptyset$). Let $\mathcal{R} = \Sigma$, and define a finely ranged probability measure $p : \mathcal{R} \to [0,1]$ by

$$p(A) = \frac{|A|}{2^n} \qquad \text{for } A \subset \prod_{k=1}^{n}\{H_k, T_k\}. \tag{5.5}$$

In other words, suppose that coin tosses are viewed as fair and independent. Then the range of $p : \mathcal{R} \to [0,1]$ is the set of all dyadic rational values, which is not convex. In particular, there is no event $A \in \mathcal{R}$ such that $p(A) = \frac{1}{3}$. Let $X = \{x, y, z\}$. Suppose that the risk preference $\succeq_1$ on $\mathcal{L}_p$ is represented by the

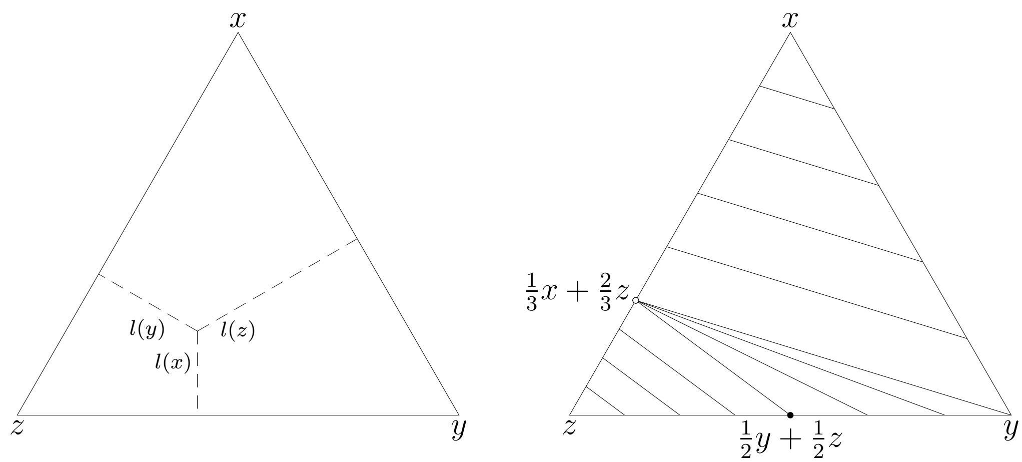

Figure 1: Non-extendable risk preference

utility function

$$V(l) = \begin{cases} 3l(x) + 2l(y) - 1 & \text{if } 1 > 3l(x) + 2l(y) \geq 3l(x) + l(y) \\ \frac{3l(x)+2l(y)-1}{1-3l(x)} = 1 + 2 \cdot \frac{3l(x)+l(y)-1}{1-3l(x)} & \text{if } 3l(x) + 2l(y) \geq 1 \geq 3l(x) + l(y) \\ 3l(x) + l(y) & \text{if } 3l(x) + 2l(y) \geq 3l(x) + l(y) > 1. \end{cases}$$

This function is well-defined on $\mathcal{L}_p$ because $l(x) \neq \frac{1}{3}$ for any $l \in \mathcal{L}_p$. Figure 1 illustrates the indifference curves of $\succeq_1$. It is clear from this illustration that the risk preference $\succeq_1$ on $\mathcal{L}_p$ is non-degenerate, complete, transitive, continuous and strictly monotonic, that is, $\succeq_1$ satisfies all conditions of Theorem 4.1(I). Accordingly, the ranking $\succeq$ of $\mathcal{G}$ represented by

$$f \succeq g \quad \Leftrightarrow \quad [f]_p \succeq_1 [g]_p,$$

satisfies P1($\mathcal{R}$), P3($\mathcal{R}$), P4*($\mathcal{R}$), P5($\mathcal{R}$), and P6($\mathcal{R}$). However, $\succeq_1$ cannot be extended to a mixture continuous weak order on the set $\mathcal{L}$ of all lotteries. Suppose on the contrary that such an extension exists. Then by mixture continuity,

$$\tfrac{1}{2}y + \tfrac{1}{2}z \succeq_1 \tfrac{1}{3}x + \tfrac{2}{3}z \succeq_1 y$$

because $\frac{1}{2}y + \frac{1}{2}z \succeq_1 \tau \cdot x + (1-\tau) \cdot z$ for all $\tau < \frac{1}{3}$ and $\tau \cdot x + (1-\tau) \cdot z \succeq_1 y$ for all $\tau > \frac{1}{3}$. By transitivity, $\frac{1}{2}y + \frac{1}{2}z \succeq_1 y$, which contradicts $V(\frac{1}{2}y + \frac{1}{2}z) < V(y)$.

In the above example, Machina–Schmeidler's analysis does not apply, and the risk preference is retained as a binary relation $\succeq_1$ on $\mathcal{L}_p$.[24] Note that the example of a non-extendable risk preference is analogous to well-known examples in calculus where a function continuous on a dense subset of a metric space has no continuous extension to the whole space. These examples motivate the use of a notion of *uniform continuity* to ensure existence of continuous extensions. P6*($\mathcal{R}$) provides an appropriate form of uniform continuity for rankings of risky acts. When this axiom holds together with the other conditions of Theorem 4.1, the induced risk preference has a unique complete, transitive, and mixture continuous extension from $\mathcal{L}_p$ to $\mathcal{L}$. One can obtain a utility representation for this extension via Machina–Schmeidler's result.

5.3 Subjective Probability without Equipartitions

Finally, we provide an example where no equiprobable events exist. Adopt the coin-tossing setting, where $S = \prod_{k=1}^{\infty}\{H_k, T_k\}$, Σ is the set of all cylinders, and $\mathcal{R} = \Sigma$. Construct a finely ranged probability measure $p^* : \mathcal{R} \to [0, 1]$ that assigns different probabilities to any two different events in $\mathcal{R}$.

Define p^* in several steps. For each sequence $(s_1, \ldots, s_n) \in \prod_{k=1}^{n}\{H_k, T_k\}$, let

$$p^*(s_1, \ldots, s_n) = \pi_1(s_1) \cdot \pi_2(s_2) \cdot \cdots \cdot \pi_n(s_n),$$

where π_i is a probability measure on the two-element set $\{H_i, T_i\}$. In other words, suppose that coin tosses are viewed as independent though not as identically distributed. Construct $\pi_1, \pi_2, \ldots$ inductively so that for all $n = 0, 1, 2, \ldots$ and for all cylinders A and B of rank n,

$$A \neq B \quad \Rightarrow \quad p^*(A) \neq p^*(B). \tag{5.6}$$

For $n = 0$, $p^*(S) \neq p^*(\emptyset)$, and (5.6) holds. Next, suppose that (5.6) holds for all cylinders of rank n. Fix arbitrary cylinders A and B of rank $n + 1$ such that

[24] Our intuition is that $\succeq_1$ may have no utility representation on $\mathcal{L}_p$, but we have no formal counterexample.

$A \neq B$. They can be written as

$$A = (A_H \cap H_{n+1}) \cup (A_T \cap T_{n+1})$$
$$B = (B_H \cap H_{n+1}) \cup (B_T \cap T_{n+1}),$$

where A_H, A_T, B_H, B_T are cylinders of rank n. The inequality $A \neq B$ implies that either $A_H \neq B_H$, or $A_T \neq B_T$, or both. Note that $p^*(A) = p^*(B)$ holds only if

$$(p^*(A_H) - p^*(B_H)) \cdot \pi_{n+1}(H_{n+1}) + (p^*(A_T) - p^*(B_T)) \cdot \pi_{n+1}(T_{n+1}) = 0. \quad (5.7)$$

This equation is non-trivial because by (5.6), either $p^*(A_H) \neq p^*(B_H)$, or $p^*(A_T) \neq p^*(B_T)$, or both. Hence, there exists at most one solution to the system of two linear equations (5.7) and

$$\pi_{n+1}(H_{n+1}) + \pi_{n+1}(T_{n+1}) = 1. \quad (5.8)$$

As the number of cylinders of rank $n+1$ is finite, there is only a finite number of probability measures π_{n+1}'s that satisfy (5.7) for some cylinders $A \neq B$ of rank $n+1$. Thus it is possible to choose π_{n+1} so that (5.6) holds for all cylinders of rank $n+1$. To ensure that $p^* : \mathcal{R} \to [0,1]$ is finely ranged, the values $\pi_{n+1}(H_{n+1})$ and $\pi_{n+1}(T_{n+1})$ can be taken in the interval $[0.4, 0.6]$ so that $p^*(s_1, \ldots, s_n) \leq 0.6^n$. By induction, there exists a finely ranged measure $p^* : \mathcal{R} \to [0,1]$ that satisfies (5.6) for all cylinders A and B.[25]

It is obvious that (5.6) rules out equipartitions. Therefore, Savage's construction of subjective probability fails for mosaics and even for finitely additive algebras. There is another lesson to be drawn from this example. Note that if (5.6) holds, then $f \leftrightarrow [f]_{p^*}$ is a bijection between the set $\mathcal{G}$ of risky acts and the set $\mathcal{L}_p$ of induced lotteries because $f \neq g$ implies $[f]_{p^*} \neq [g]_{p^*}$. Therefore, *every* complete and transitive preference $\succeq$ on $\mathcal{G}$ is represented by

$$f \succeq g \quad \Leftrightarrow \quad [f]_{p^*} \succeq_1 [g]_{p^*} \qquad \text{for } f, g \in \mathcal{G},$$

[25]Formally, the induction relies on the axiom of choice. It is possible to avoid this reliance, and construct p^* explicitly, but then the argument becomes less transparent.

where $\succeq_1$ is a complete and transitive binary relation on $\mathcal{L}_p$. This observation shows that in our setting, any meaningful notion of probabilistic sophistication must impose additional conditions on the risk preference besides completeness and transitivity. In Machina–Schmeidler's paradigm, these conditions are monotonicity and continuity.

A APPENDIX: Proofs of Theorems

In the proofs we use a briefer terminology. *Events* are elements of the mosaic $\mathcal{R}$ and are denoted by A, B, C, D, E, F, H, and S. *Partitions*, written as $\{A = A_1^n\}$, are collections of disjoint events $\{A_1, \ldots, A_n\} \subset \mathcal{R}$ such that $\cup_{i=1}^n A_i = A \in \mathcal{R}$. *Acts* are elements of $\mathcal{G}$ and are denoted by f, g, and h. The adjective *risky* is skipped as redundant hereafter.

If events A and B are disjoint, write $A \oplus B$ instead of $A \cup B$. If B is a subset of A, write $A \ominus B$ instead of $A \setminus B$. Given a partition $\{A = A_1^n\}$ and indices $i, j \geq 0$, let

$$A_i^j = \bigcup_{k \in [1,n]\,:\, i \leq k \leq j} A_k.$$

Then $A_i^j \in \mathcal{R}$ because $\mathcal{R}$ satisfies (μ) and $S = A_1 \oplus A_2 \oplus \cdots \oplus A_n \oplus \neg A$ is a partition of the universal event.

Given a collection $\mathcal{E} \subset \mathcal{R}$ of events, denote by $\alpha(\mathcal{E})$ the smallest algebra that contains $\mathcal{E}$. If $\mathcal{E}$ is finite, then $\alpha(\mathcal{E})$ is also finite. With a slight abuse of notation, write $\alpha(A)$ instead of $\alpha(\{A\})$, $\alpha(A, B)$ instead of $\alpha(\{A, B\})$ etc. As $\mathcal{R}$ satisfies $(*)$ and $(**)$, then for all $A \in \mathcal{R}$, $\alpha(A) = \{\emptyset, A, \neg A, S\} \subset \mathcal{R}$. As $\mathcal{R}$ satisfies (μ), then for all partitions $\{S = A_1^m\}$ and $\{S = B_1^n\}$,

$$\begin{gathered} \alpha(A_1, A_2, \ldots, A_m, B_1, B_2, \ldots, B_n) \subset \mathcal{R} \quad \Leftrightarrow \\ A_i \cap B_j \in \mathcal{R} \quad \text{for all } i = 1 \ldots m \text{ and } j = 1 \ldots n \quad \Leftrightarrow \\ \{A_1, A_2, \ldots, A_m\} \subset \mathcal{R} \Cap \{B_1, B_2, \ldots, B_n\}. \end{gathered} \tag{A.1}$$

A.1 Construction of Subjective Probability

Suppose that $\succeq$ satisfies P1($\mathcal{R}$), P3($\mathcal{R}$), P4*($\mathcal{R}$), P5($\mathcal{R}$), P6($\mathcal{R}$). Fix outcomes $x \succ x'$. For arbitrary events $A, B \in \mathcal{R}$, define a *comparative likelihood* relation $A \succeq_0 B$ as follows:

$$A \succeq_0 B \quad \Leftrightarrow \quad xAx' \succeq xBx'.$$

Note that by P4*($\mathcal{R}$), $A \succeq_0 B$ if and only if $zAz' \succeq zBz'$ *for all* $z \succ z'$. We seek a subjective probability measure $p : \mathcal{R} \to [0, 1]$ that represents $\succeq_0$:

$$p(A) \geq p(B) \quad \Leftrightarrow \quad A \succeq_0 B \qquad \text{for } A, B \in \mathcal{R}. \tag{A.2}$$

Such p is called a *quantitative probability.*

Call an event $A \in \mathcal{R}$ *non-null* if $A \succ_0 \emptyset$; call A *null* otherwise. Say that a partition $\{A = A_1^n\}$ is *non-null* if $A_i \succ_0 \emptyset$ for all $i = 1 \dots n$. Say that $\{A = A_1^n\}$ is *finer* than an event B if $A_i \prec_0 B$ for all $i = 1 \dots n$.

Axioms P1($\mathcal{R}$), P3($\mathcal{R}$), P4*($\mathcal{R}$), P5($\mathcal{R}$), P6($\mathcal{R}$) imply the following properties for the comparative likelihood relation $\succeq_0$.

Q1. $\succeq_0$ is complete and transitive. (Q1 follows from P1($\mathcal{R}$)).

Q2. $S \succ_0 \emptyset$.

Q3. For all events $A \in \mathcal{R}$, $S \succeq_0 A \succeq_0 \emptyset$. (Q3 follows from P3($\mathcal{R}$)).

Q4. For all partitions $\{A = A_1^2\}$ and $\{B = B_1^2\}$ such that $\alpha(A_1, A_2, B_1, B_2) \subset \mathcal{R}$,

$$A_1 \succeq_0 (\succ_0) B_1 \text{ and } A_2 \succeq_0 B_2 \quad \Rightarrow \quad A \succeq_0 (\succ_0)\, B.$$

Proof. Let $A_3 = \neg A$, $B_3 = \neg B$, and $E_{ij} = A_i \cap B_j$ for $i, j = 1 \dots 3$. All of these sets belong to the algebra $\alpha(A_1, A_2, B_1, B_2)$ and hence, belong to $\mathcal{R}$. Denote by $f = \begin{pmatrix} x_{11} & x_{12} & x_{13} \\ x_{21} & x_{22} & x_{23} \\ x_{31} & x_{32} & x_{33} \end{pmatrix}$ the act that yields outcomes x_{ij} on E_{ij}. Then

$$\underbrace{A_1 \succeq_0 (\succ_0) B_1}_{\text{def.}} \qquad\qquad \underbrace{A_2 \succeq_0 B_2}_{\text{def.}}$$

$$\underbrace{\begin{pmatrix} x & x & x \\ x' & x' & x' \\ x' & x' & x' \end{pmatrix} \succeq (\succ) \begin{pmatrix} x & x' & x' \\ x & x' & x' \\ x & x' & x' \end{pmatrix}}_{\text{P4}^*(\mathcal{R})} \qquad \underbrace{\begin{pmatrix} x' & x' & x' \\ x & x & x \\ x' & x' & x' \end{pmatrix} \succeq \begin{pmatrix} x' & x & x' \\ x' & x & x' \\ x' & x & x' \end{pmatrix}}_{\text{P4}^*(\mathcal{R})}$$

$$\underbrace{\begin{pmatrix} x' & x & x \\ x' & x' & x \\ x' & x' & x' \end{pmatrix} \succeq (\succ) \begin{pmatrix} x' & x' & x' \\ x & x' & x \\ x & x' & x' \end{pmatrix} = \begin{pmatrix} x' & x' & x' \\ x & x' & x \\ x & x' & x' \end{pmatrix} \succeq \begin{pmatrix} x' & x & x' \\ x' & x' & x' \\ x & x & x' \end{pmatrix}}_{\text{P1}(\mathcal{R})}$$

$$\underbrace{\begin{pmatrix} x' & x & x \\ x' & x' & x \\ x' & x' & x' \end{pmatrix} \succeq (\succ) \begin{pmatrix} x' & x & x' \\ x' & x' & x' \\ x & x & x' \end{pmatrix}}_{\text{P4}^*(\mathcal{R})}$$

$$\underbrace{\begin{pmatrix} x & x & x \\ x & x & x \\ x' & x' & x' \end{pmatrix} \succeq (\succ) \begin{pmatrix} x & x & x' \\ x & x & x' \\ x & x & x' \end{pmatrix}}_{\text{def.}}$$

$$A \succeq_0 (\succ_0) B.$$

$\square$

By induction, Q4 can be extended to n-element partitions as follows: for all partitions $\{A = A_1^n\}$ and $\{B = B_1^n\}$ such that $\alpha(A_1, \dots, A_n, B_1, \dots, B_n) \subset \mathcal{R}$,

$$A_1 \succeq_0 (\succ_0) B_1 \text{ and } A_i \succeq_0 B_i \text{ for all } i > 1 \quad \Rightarrow \quad A \succeq_0 (\succ_0)\, B.$$

Roughly, Q4 asserts that comparative likelihood is additive over any subalgebra of the mosaic $\mathcal{R}$.

QF (Fineness). For any non-null event A and for any finite collection $\mathcal{E} \subset \mathcal{R}$, there exists a partition $\{S = S_1^m\} \subset \mathcal{R} \Cap \mathcal{E}$ finer than A.

Proof. Fix $A \succ_0 \emptyset$ and finite $\mathcal{E} \subset \mathcal{R}$. The collection $\mathcal{E}' = \alpha(A) \cup \mathcal{E}$ is finite, and the acts $f = xAx'$ and $g = x'$ are $\mathcal{E}'$-measurable. $A \succ_0 \emptyset$ implies that $f \succ g$. By P6($\mathcal{R}$), there exists a partition $\{S = S_1^m\} \subset \mathcal{R} \Cap \mathcal{E}' \subset \mathcal{R} \Cap \mathcal{E}$ such that for all $i = 1 \dots m$, $f \succ xS_i g$, that is, $A \succ_0 S_i$. □

QT (Tightness). For any events $A \succ_0 B$, there exists a non-null partition $A = C \oplus D$ such that $C = A \ominus D \succ_0 B$ and a non-null partition $\neg B = C' \oplus D'$ such that $A \succ_0 B \oplus C'$.

Proof. Fix $A \succ_0 B$. Let $\mathcal{E} = \alpha(A) \cup \alpha(B)$. The acts $f = xAx'$ and $g = xBx'$ are $\mathcal{E}$-measurable, and $f \succ g$ because $A \succ_0 B$. By P6($\mathcal{R}$), there exists a partition $\{S = S_1^m\} \subset \mathcal{R} \Cap \mathcal{E}$ such that for all $i = 1 \dots m$, $x'S_i f \succ g$, that is, $A \ominus (A \cap S_i) \succ_0 B$. Let $E_i = S_i \cap A$; then $E_i \in \mathcal{R}$ because $A \in \mathcal{E}$ and $S_i \in \mathcal{R} \Cap \mathcal{E}$. Let $D = E_k$ and $C = A \ominus D$, where k is such that $E_k \succ_0 \emptyset$; if there is no such k, then by Q4, $\emptyset \succeq E_1^m = A$, which contradicts $A \succ_0 B \succeq_0 \emptyset$. Thus, the partition $A = C \oplus D$ is non-null, and $C \succ_0 B$.

By P6($\mathcal{R}$), there exists (another) partition $\{S = S_1^m\} \subset \mathcal{R} \Cap \mathcal{E}$ such that for all $i = 1 \dots m$, $f \succ xS_i g$, that is, $A \succ_0 B \oplus (S_i \cap \neg B)$. Take $E_i = S_i \cap (\neg B)$; then $E_i \in \mathcal{R}$ because $\neg B \in \mathcal{E}$ and $S_i \in \mathcal{R} \Cap \mathcal{E}$. Let $C' = E_k$ and $D' = \neg B \ominus C'$, where k is such that $E_k \succ_0 \emptyset$; if there is no such k, then by Q4, $B \succeq_0 B \oplus E_1^m = S$, which contradicts $S \succeq_0 A \succ_0 B$. Then $A \succ_0 B \oplus C'$, and D' is non-null because otherwise by Q4, $B \oplus C' \succeq_0 (B \oplus C') \oplus D' = S \succeq_0 A$. Thus, the partition $\neg B = C' \oplus D'$ is non-null, and $A \succ_0 B \oplus C'$. □

Note that Q1–Q4 generalize de Finetti's [3] definition of a *qualitative probability*, while QF and QT generalize Savage's conditions of *fineness* and *tightness*. Savage shows that any fine and tight qualitative probability $\succeq_0$ on a σ-algebra can be represented by quantitative probability.[26] Niiniluoto [18] and Wakker [22] extend this result to algebras. We extend it further to mosaics. Therefore, our construction of subjective probability does not rely directly on properties of preference $\succeq$ but relies instead on properties of the comparative likelihood relation $\succeq_0$. This construction applies in the proofs of both Theorems 3.1 and 4.1; it applies also in a de Finetti-type setting where $\succeq_0$ is taken as a primitive and Q1–Q4, QF, QT are imposed as axioms.

Properties Q1–Q4, QF, and QT have the following implications.

Q5. For all partitions $\{A = A_1^2\}$ and $\{B = B_1^2\}$,

$$A_1 \succeq_0 (\succ_0) B_1 \text{ and } A_2 \succeq_0 B_2 \quad \Rightarrow \quad A \succeq_0 (\succ_0)\, B.$$

Proof. Fix partitions $\{A = A_1^2\}$ and $\{B = B_1^2\}$ and consider three cases.

Case I. $A_1 \succ_0 B_1$ and $A_2 \succ_0 B_2$. Construct a partition $\{E = E_1^2\}$ such that

$$\begin{aligned} \alpha(A_1, A_2, E_1, E_2) \subset \mathcal{R} \quad &\text{and} \quad \alpha(B_1, B_2, E_1, E_2) \subset \mathcal{R}, \\ A_1 \succ_0 E_1 \succ_0 B_1 \quad &\text{and} \quad A_2 \succ_0 E_2 \succ_0 B_2. \end{aligned} \tag{A.3}$$

[26] It is well-known that Q1–Q4 alone are not sufficient for a quantitative probability representation (Kraft et al. [14]).

Then by Q4, $A \succ_0 E \succ_0 B$ and by transitivity, $A \succ_0 B$. To construct E_1 and E_2, use QT and fix non-null partitions $A_1 = C_1 \oplus D_1$ and $A_2 = C_2 \oplus D_2$ such that $C_1 \succ_0 B_1$ and $C_2 \succ_0 B_2$. By QF, there exists a partition

$$\{S = S_1^m\} \subset \mathcal{R} \Cap \{C_1, D_1, C_2, D_2, \neg A, B_1, B_2, \neg B\}$$

that is finer than D_1 and also finer than D_2. By (A.1),

$$\begin{aligned}&\alpha(S_1, \ldots, S_m, C_1, D_1, C_2, D_2) \subset \mathcal{R} \quad \text{and} \\ &\alpha(S_1, \ldots, S_m, B_1, B_2) \subset \mathcal{R}.\end{aligned}$$

As $S_1^m \succ_0 C_1 \succ_0 S_1^0$, there exists $k_1 \in [0, m]$ such that $S_1^{k_1} \succ_0 C_1 \succeq_0 S_1^{k_1-1}$. Then

$$A_1 = C_1 \oplus D_1 \underset{\text{Q4}}{\succ_0} S_1^{k_1-1} \oplus S_{k_1} = S_1^{k_1} \succ_0 C_1 \succ_0 B_1.$$

Note that $S_{k_1+1}^m \succ_0 C_2$ because $C_2 \succeq_0 S_{k_1+1}^m$ implies a contradiction:

$$A_1 \oplus C_2 \underset{\text{Q4}}{\succ_0} S_1^{k_1} \oplus S_{k_1+1}^m = S.$$

As $S_{k_1+1}^m \succ_0 C_2 \succ_0 S_{k_1+1}^{k_1}$, there exists $k_2 \in [k_1, m]$ such that $S_{k_1+1}^{k_2} \succ_0 C_2 \succeq_0 S_{k_1+1}^{k_2-1}$. Then

$$A_2 = C_2 \oplus D_2 \underset{\text{Q4}}{\succ_0} S_{k_1+1}^{k_2-1} \oplus S_{k_2} = S_{k_1+1}^{k_2} \succ_0 C_2 \succ_0 B_2.$$

Let $E_1 = S_1^{k_1}$ and $E_2 = S_{k_1+1}^{k_2}$. Then E_1 and E_2 satisfy conditions (A.3).

Case II. $A_1 \succ_0 B_1$ and $A_2 \succeq_0 B_2$. Use QT and fix a non-null partition $A_1 = C_1 \oplus D_1$ such that $C_1 \succ_0 B_1$. Then $A_2 \oplus D_1 \succ_0 A_2 \succeq_0 B_2$. By Case I,

$$A = C_1 \oplus (A_2 \oplus D_1) \succ_0 B_1 \oplus B_2 = B$$

and $A \succ_0 B$.

Case III. $A_1 \succeq_0 B_1$ and $A_2 \succeq_0 B_2$. Suppose that $B \succ_0 A$. Use QT and fix a non-null partition $\neg A = C' \oplus D'$ such that $B \succ_0 A \oplus C'$. Then $A_1 \oplus C' \succ_0 A_1 \succeq_0 B_1$. By Case II,

$$A \oplus C' = (A_1 \oplus C') \oplus A_2 \succ_0 B_1 \oplus B_2 = B$$

and $A \oplus C' \succ_0 B$, which is a contradiction. Thus, $A \succeq_0 B$. $\square$

By induction, Q5 can be extended to n-element partitions as follows: for all partitions $\{A = A_1^n\}$ and $\{B = B_1^n\}$,

$$A_1 \succeq_0 (\succ_0) B_1 \text{ and } A_i \succeq_0 B_i \text{ for all } i > 1 \quad \Rightarrow \quad A \succeq_0 (\succ_0)\, B.$$

Roughly, Q5 asserts that comparative likelihood is additive over the entire mosaic $\mathcal{R}$.

Q6. For any non-null event A and for any event B, there exists a partition $\{B = B_1^m\}$ finer than A.

Proof. By QF, there exists a partition $\{S = S_1^m\} \subset \mathcal{R} \Cap \{B, \neg B\}$ finer than A. For all $i = 1 \dots m$, the sets $B_i = B \cap S_i$ and $\neg B \cap S_i$ belong to $\mathcal{R}$ and partition S_i. Therefore, $B_i \preceq_0 S_i \prec_0 A$, and the partition $\{B = B_1^m\}$ is finer than A. $\square$

Q7. For any events $A \succ_0 B$ and for any non-null partition $\{A = A_1^n\}$, there exists an n-element partition $\{B = B_1^n\}$ such that $A_i \succ_0 B_i$ for all $i = 1 \dots n$.

Proof. First, suppose that $n = 2$. By QT, there exists a non-null partition $\neg B = C' \oplus D'$ such that $A \succ_0 B \oplus C'$. Use Q6 and fix a partition $\{B = E_1^m\}$ finer than C'. Let k be the maximal index in $[0, m]$ such that $A_1 \succ_0 E_1^k$. If $A_2 \succ_0 E_{k+1}^m$, then $B = E_1^k \oplus E_{k+1}^m$ is the required partition. Suppose on the contrary that $E_{k+1}^m \succeq_0 A_2$. Then $k < m$, and

$$E_1^k \oplus C' \underset{\text{Q5}}{\succ_0} E_1^k \oplus E_{k+1} = E_1^{k+1} \underset{\text{def. of } k}{\succeq_0} A_1.$$

This implies a contradiction

$$B \oplus C' = (E_1^k \oplus C') \oplus E_{k+1}^m \underset{\text{Q5}}{\succ_0} A_1 \oplus A_2 = A.$$

Complete the proof by induction with respect to n. $\square$

Q8. For any events $A \succ_0 B$ and for any partition $\{B = B_1^n\}$, there exists an n-element partition $\{A = A_1^n\}$ such that $A_i \succ_0 B_i$ for all $i = 1 \dots n$.

Proof. First, suppose that $n = 2$. By QT, there exists a non-null partition $A = C \oplus D$ such that $C \succ_0 B$. Use Q6 and fix a partition $\{C = E_1^m\}$ finer than D. Let k be the maximal index in $[0, m]$ such that $B_1 \succ_0 E_1^k$. Then $E_1^k \oplus D \succ_0 E_1^k \oplus E_{k+1} \succeq_0 B_1$. Also $E_{k+1}^m \succ_0 B_2$ because there is a contradiction otherwise:

$$B_1 \oplus B_2 \underset{\text{Q5}}{\succ_0} E_1^k \oplus E_{k+1}^m = C.$$

Thus, $A = (E_1^k \oplus D) \oplus E_{k+1}^m$ is the required partition of A.

Complete the proof by induction with respect to n. $\square$

Next, we define approximate numerical likelihoods of events. For every $A \in \mathcal{R}$, let

$$\nu(A) = \min_{\{S = S_1^m\} \prec_0 A} m. \tag{A.4}$$

In other words, let $\nu(A)$ be the minimal number of elements that a partition of S finer than A may have. If A is non-null, then by QF, $\nu(A)$ is finite; if A is null, then $\nu(A)$ is equal to $+\infty$. Note that if $\{S = S_1^{\nu(A)}\}$ is finer than A, then this partition is non-null because otherwise S can be partitioned into $\nu(A) - 1$ events that are finer than A.

Define the *approximate likelihood* of an event $A \in \mathcal{R}$ as

$$r(A) = \frac{1}{\nu(A)}. \tag{A.5}$$

The range of the function r belongs to the interval $[0, 1]$ or more precisely, to the set $\left\{\frac{1}{2}, \frac{1}{3}, \dots, 0\right\}$. By QF, $r(A) = 0$ if and only if A is null.

The notion of approximate likelihood satisfies the following properties R1–R6.

R1. The function r almost agrees with $\succeq_0$, that is, for all events A and B,

$$\begin{aligned} A \succeq_0 B \quad &\Rightarrow \quad r(A) \geq r(B)\ ; \\ r(A) > r(B) \quad &\Rightarrow \quad A \succ_0 B. \end{aligned}$$

Proof. Any partition finer than $B \preceq_0 A$ is also finer than A. Therefore, $A \succeq_0 B$ implies $\nu(A) \leq \nu(B)$ and $r(A) \geq r(B)$. Conversely, $r(A) > r(B)$ implies $A \succ_0 B$. □

R2. For all partitions $\{A = A_1^n\}$, $\max\limits_{i=1\dots n} \nu(A_i) > n$, or equivalently, $\min\limits_{i=1\dots n} r(A_i) < \frac{1}{n}$.

Proof. Fix a partition $\{A = A_1^n\}$, and let A_k be such that $A_k \preceq_0 A_i$ for all $i = 1 \dots n$. By R1, $\max\limits_{i=1\dots n} \nu(A_i) = \nu(A_k)$. Suppose that $\nu(A_k) < n$. Take a partition $\{S = S_1^{\nu(A_k)}\}$ finer than A_k. For all $i = 1 \dots \nu(A_k)$, $A_i \succeq_0 A_k \succ_0 S_i$. By Q5, $A_1^{\nu(A_k)} \succ_0 S_1^{\nu(A_k)} = S$, which is impossible. □

R3. For any n, there exists $A \succ_0 \emptyset$ such that $\nu(A) > n$, that is, $r(A) < \frac{1}{n}$. In other words, the range of the function r includes arbitrarily small positive values.

Proof. Fix $A \succ_0 \emptyset$ and a non-null partition $\{S = S_1^{\nu(A)}\}$ finer than A. By R2, $\nu(A) < \max\limits_{i=1\dots\nu(A)} \nu(S_i)$. Therefore, the function ν is unbounded on non-null events. □

R4. For any $A \succ_0 \emptyset$, there exists a non-null partition $A = C \oplus D$ such that $r(C) = r(A)$.

Proof. Fix an event $A \succ_0 \emptyset$ and a partition $\{S = S_1^{\nu(A)}\}$ finer than A. Let S_k be such that $S_k \succeq_0 S_i$ for all $i = 1 \dots \nu(A)$. By QT, there exists a non-null partition $A = C \oplus D \succ_0 S_k$ such that $C \succ_0 S_k$. Then $\{S = S_1^{\nu(A)}\}$ is finer than C. Therefore, $\nu(C) \leq \nu(A)$. On the other hand, $\nu(C) \geq \nu(A)$ because $C \prec_0 A$. Thus, $r(C) = r(A)$. □

Given a non-null event B, say that $\{A = B_1^n\}$ is a *B-partition* if this partition is finer than B, and $r(B_i) = r(B)$ for all $i = 1 \dots n-1$.

R5. For any $A \in \mathcal{R}$ and for any non-null $B \in \mathcal{R}$, there exists a B-partition $\{A = B_1^n\}$.

Proof. Say that a partition $\{A = A_1^n\}$ is B-acceptable, if for all $i = 1 \dots n-1$, $A_i \prec_0 B$ and $r(A_i) = r(B)$. A trivial B-acceptable partition is $\{A = A\}$. If $\{A = A_1^n\}$ is B-acceptable, then by R2, $n - 1 < \max\limits_{i=1\dots n-1} \nu(A_i) = \nu(B)$. Therefore, there exists a B-acceptable partition $\{A = B_1^n\}$ that has the maximal number of elements among all B-acceptable partitions. Suppose that $B_n \succeq_0 B$. By R4 and Q6, there exists a non-null partition $B = C \oplus D$ such that $r(C) = r(B)$ and a partition $\{B_n = E_1^m\}$ finer than D. Let k be the maximal index in $[0, m]$ such that $C \succ_0 E_1^k$. Then $B = C \oplus D \succ_0 E_1^k \oplus E_{k+1} = E_1^{k+1} \succeq_0 C$. It follows that $r(B) \geq r(E_1^{k+1}) \geq r(C)$, that is, $r(B) = r(E_1^{k+1})$. Therefore, the $(n+1)$-element partition

$$A = B_1 \oplus \cdots \oplus B_{n-1} \oplus E_1^{k+1} \oplus E_{k+2}^m$$

is is B-acceptable. This contradiction implies that $B \succ_0 B_n$. Thus, $\{A = B_1^n\}$ is a B-partition. □

R6. For any event A, there exists $\varepsilon > 0$ such that for all non-null events B and for all B-partitions $\{A = B_1^n\}$,

$$r(B) < \varepsilon \quad \Rightarrow \quad r(A) \leq (n-1) \cdot r(B).$$

Proof. If A is null, then $r(A) = 0$. Suppose that A is non-null. By R4, there exists a non-null partition $A = C \oplus D$ such that $r(C) = r(A)$. Let $\varepsilon = r(D)$. Fix a non-null event B such that $r(B) < \varepsilon$ and a B-partition $\{A = B_1^n\}$. Then $B_n \prec_0 B \prec_0 D$ and $A \succeq_0 B_1^{n-1} \succ_0 C$. Therefore, $r(B_1^{n-1}) = r(A)$. Take a partition $\{S = S_1^{\nu(A)}\}$ finer than B_1^{n-1}. By Q7, each event $S_j \prec_0 B_1^{n-1}$ can be subpartitioned into $n-1$ events,

$$S_j = S_{j,1} \oplus S_{j,2} \oplus \cdots \oplus S_{j,n-1}$$

such that $S_{j,i} \prec_0 B_i \prec_0 B$ for all $i = 1 \ldots n-1$. The partition of S into $(n-1) \cdot \nu(A)$ elements $S_{j,i}$ is finer than B. Thus, $(n-1) \cdot \nu(A) \geq \nu(B)$, and $r(A) \leq (n-1) \cdot r(B)$. $\square$

The following theorem delivers a quantitative probability representation for $\succeq_0$.

Theorem A.1. *A binary relation $\succeq_0$ on a mosaic $\mathcal{R}$ satisfies Q1–Q4, QF, and QT if and only if $\succeq_0$ is represented by a finely ranged probability measure $p : \mathcal{R} \to [0,1]$. The probability measure p that represents $\succeq_0$ is unique and for all $A \in \mathcal{R}$,*

$$p(A) = \sup_{\{A = A_1^n\}} \sum_{i=1}^{n} r(A_i). \tag{A.6}$$

Proof. Suppose that $\succeq_0$ satisfies Q1–Q4, QF, QT. Let $p : \mathcal{R} \to [0,1]$ be the function given by (A.6). Show that p is a probability measure. For all $A \in \mathcal{R}$, $p(A) \geq r(A) \geq 0$. Fix a partition $\{S = S_1^m\}$, a non-null event B, and B-partitions $\{S_i = B(i)_1^{n_i}\}$. Then the partition of S into $\sum_{i=1}^m n_i$ events $B(i)_j$ is finer than B, and hence, $\sum_{i=1}^m n_i \geq \nu(B)$. It follows that

$$\sum_{i=1}^{m} p(S_i) \geq \sum_{i=1}^{m} (n_i - 1) \cdot r(B) = \left(\sum_{i=1}^{m} n_i \right) \cdot r(B) - m \cdot r(B) \geq 1 - m \cdot r(B).$$

As $r(B)$ can be arbitrarily small, $\sum_{i=1}^m p(S_i) \geq 1$. On the other hand, by R2,

$$\sum_{i=1}^{m} (n_i - 1) < \max_{\substack{i=1\ldots m \\ j=1\ldots(n_i-1)}} \nu(B(i)_j) = \frac{1}{r(B)}.$$

By R6, there exists $\varepsilon > 0$ such that

$$r(B) < \varepsilon \quad \Rightarrow \quad r(S_i) < (n_i - 1) \cdot r(B) \quad \text{for all } i = 1 \ldots m.$$

As $r(B)$ can be arbitrarily small,

$$\sum_{i=1}^{m} r(S_i) < \left(\sum_{i=1}^{m} (n_i - 1) \right) \cdot r(B) < \frac{1}{r(B)} \cdot r(B) = 1.$$

This inequality implies that for all partitions $\{S_i = A(i)_1^{m_i}\}$,

$$S = \bigcup_{\substack{i=1\ldots m\\ j=1\ldots m_i}} A(i)_j \quad \Rightarrow \quad \sum_{i=1}^{m}\sum_{j=1}^{m_i} r(A(i)_j) < 1 \quad \Rightarrow$$

$$\sum_{i=1}^{m} p(S_i) = \sum_{i=1}^{m}\left(\sup_{\{S_i=A(i)_1^{m_i}\}} \sum_{j=1}^{m_i} r(A(i)_j)\right) \leq 1.$$

Thus, $\sum_{i=1}^{m} p(S_i) = 1$, and p is a probability measure.

Show that p represents $\succeq_0$. Fix arbitrary events A and B. Suppose that $A \succ_0 B$. Then by QT, there exists a non-null partition $A = C \oplus D$ such that $C \succ_0 B$. By Q8, for any partition $\{B = B_1^n\}$, there exists a partition $\{C = C_1^n\}$ such that $C_i \succ_0 B_i$ for all $i = 1 \ldots n$. It follows that

$$p(A) \geq r(D) + \sup_{\{C=C_1^n\}} \sum_{i=1}^{n} r(C_i) \geq r(D) + \sup_{\{B=B_1^n\}} \sum_{i=1}^{n} r(B_i) = r(D) + p(B).$$

Thus, $p(A) > p(B)$. On the other hand, suppose that $p(A) > p(B)$. Then, there exists a partition $\{A = A_1^n\}$ such that $\sum_{i=1}^{n} r(A_i) > p(B)$. Without loss of generality, $r(A_1) > 0$. By R4, there exists a non-null partition $A_1 = C \oplus D$ such that $r(A_1) = r(C)$. Then $p(C) \geq \sum_{i=1}^{n} r(A_i) > p(B)$, and $A \succ_0 C \succeq_0 B$. Thus, p represents $\succeq_0$.

Show that p is finely ranged. Fix an arbitrary finite collection $\mathcal{E} \subset \mathcal{R}$ and an arbitrary $\varepsilon > 0$. Use R3 and fix a non-null event B such that $r(B) < \varepsilon$. Take a non-null partition $\{S = B_1^{\nu(B)}\}$ finer than B. Then $\sum_{i=1}^{\nu(B)} p(B_i) = 1$. Therefore, there exists $B_i \succ_0 \emptyset$ such that $p(B_i) \leq \frac{1}{\nu(B)} < \varepsilon$. By QF, there exists a partition $\{S = S_1^m\} \subset \mathcal{R} \Cap \mathcal{E}$ finer than B_i. It follows that $p(S_i) < p(B_i) < \varepsilon$ for all $i = 1 \ldots m$.

Show that p is the unique probability measure that represents $\succeq_0$. Suppose that another probability measure $p^* : \mathcal{R} \to [0,1]$ represents $\succeq_0$. Fix an event $A \in \mathcal{R}$. If A is null, then $p^*(A) = p^*(\emptyset) = 0 = p(A)$. If A is non-null, then there exists a partition $\{S = S_1^{\nu(A)}\}$ finer than A. As p^* is additive and represents $\succeq_0$, then

$$1 = p^*(S) = \sum_{i=1}^{\nu(A)} p^*(S_i) < \nu(A) \cdot p^*(A),$$

that is $p^*(A) > \frac{1}{\nu(A)} = r(A)$. Therefore, for all partitions $\{A = A_1^n\}$,

$$p^*(A) = \sum_{i=1}^{n} p^*(A_i) > \sum_{i=1}^{n} r(A_i)$$

and $p^*(A) \geq p(A)$. Similarly, $p^*(\neg A) \geq p(\neg A)$. Thus, $p(A) = p^*(A)$.

Finally, show that properties Q1–Q4, QF, and QT are necessary for a finely ranged probability measure $p : \mathcal{R} \to [0,1]$ to represent the binary relation $\succeq_0$. This is non-trivial only for QT. Fix arbitrary events $A \succ_0 B$, and let $\varepsilon = p(A) - p(B) > 0$. There exists a partition $\{S = S_1^m\} \subset \mathcal{R} \Cap \{A, \neg A, B, \neg B\}$ such that $p(S_i) < \varepsilon$ for all $i = 1 \ldots m$. Then $\{A = E_1^m\}$, where $E_i = A \cap S_i$, and $p(E_k) > 0$ for some k. Let $D = E_k$. Then $p(A \ominus D) > p(A) - \varepsilon = p(B)$, and $A \ominus D \succ_0 B$. The proof of the second part of QT is analogous. $\square$

A.2 First-Order Stochastic Monotonicity

Suppose that $\succeq$ satisfies P1($\mathcal{R}$), P3($\mathcal{R}$), P4*($\mathcal{R}$), P5($\mathcal{R}$), and P6($\mathcal{R}$). Then the finely ranged probability measure $p : \mathcal{R} \to [0, 1]$ given by (A.6) represents the preference over binary acts that have outcomes $x \succ x'$. As we show next, the preference over the set $\mathcal{G}$ of all acts is first-order stochastically monotonic with respect to the probability measure p, that is, for all acts f and g,

$$[f]_p \geqslant [g]_p \quad \Rightarrow \quad f \succeq g, \tag{A.7}$$

$$[f]_p \gg [g]_p \quad \Rightarrow \quad f \succ g. \tag{A.8}$$

The proof of (A.7) and (A.8) relies on the following lemma.

Lemma A.2. *For all $n > 0$, for all finite collections $\mathcal{E} \subset \mathcal{R}$, and for all events $A \in \mathcal{R} \Cap \mathcal{E}$, the set*

$$p(A, n, \mathcal{E}) = \bigcup_{\{A = A_1^n\} \subset \mathcal{R} \Cap \mathcal{E}} \{(p(A_1), p(A_2), \ldots, p(A_n))\}$$

of n-dimensional vectors $(p(A_1), p(A_2), \ldots, p(A_n))$ is dense in the simplex

$$\Delta(A, n) = \{(v_1, v_2, \ldots, v_n) \in \mathbb{R}^n_+ \,:\, \sum_{i=1}^n v_i = p(A)\}.$$

Proof. Fix a finite collection $\mathcal{E} \subset \mathcal{R}$ and an event $A \in \mathcal{R} \Cap \mathcal{E}$. Fix arbitrary value $v \in [0, p(A)]$ and $\varepsilon > 0$. Show that there exists a partition $\{A = A_1^2\} \subset \mathcal{R} \Cap \mathcal{E}$ such that $v - \varepsilon \leq p(A_1) \leq v$. Let

$$\mathcal{E}' = \alpha(A) \cup \{E' \in \mathcal{R} \,:\, E' = A \cap E \text{ for some } E \in \mathcal{E}\}.$$

As p is finely ranged, there exists a partition $\{S = S_1^m\} \subset \mathcal{R} \Cap \mathcal{E}'$ such that $p(S_i) < \varepsilon$ for all $i = 1 \ldots m$. Then the sets $A \cap S_i$ and $\neg A \cap S_i$ belong to $\mathcal{R}$ and partition S_i; by (μ), all unions of these events belong to $\mathcal{R}$. For $i = 1 \ldots m$, $p(A \cap S_i) \leq p(S_i) < \varepsilon$. Let $k \in [0, m]$ be the minimal index such that $p(A \cap S_1^k) \geq v - \varepsilon$. Then $p(A \cap S_1^k) \leq v$ because otherwise $p(A \cap S_1^{k-1}) > v - p(A \cap S_k) > v - \varepsilon$. Let $A_1 = A \cap S_1^k$ and $A_2 = A \cap S_{k+1}^m$. For all $E \in \mathcal{E}$ and for all $i = 1 \ldots m$, the events $(A \cap S_i) \cap E = S_i \cap (A \cap E) \in \mathcal{R}$ partition $A \cap E \in \mathcal{R}$; by (μ), $A_1 \cap E \in \mathcal{R}$ and $A_2 \cap E \in \mathcal{R}$. Thus, $A_1 \in \mathcal{R} \Cap \mathcal{E}$ and $A_2 \in \mathcal{R} \Cap \mathcal{E}$.

Fix an arbitrary vector $(v_1, v_2, \ldots, v_n) \in \mathbb{R}^n_+$ such that $\sum_{i=1}^n v_i = p(A)$. By induction with respect to n, there exists a partition $\{A = A_1^n\} \subset \mathcal{R} \Cap \mathcal{E}$ such that $v_i - \frac{\varepsilon}{2n} \leq p(A_i) \leq v_i$ for all $i = 1 \ldots n - 1$. Then

$$\sum_{i=1}^n |p(A_i) - v_i| \leq 2 \cdot (n-1) \cdot \frac{\varepsilon}{2n} < \varepsilon.$$

Thus, $p(A, n, \mathcal{E})$ is dense in $\Delta(A, n)$. $\square$

Show that $\succeq$ satisfies (A.7) and (A.8). Say that acts f and g differ through n-outcome subacts, if f and g can be written as

$$f = \begin{bmatrix} z_1 & \text{if } s \in F_1 \\ z_2 & \text{if } s \in F_2 \\ \ldots\ldots & \ldots\ldots \\ z_n & \text{if } s \in F_n \\ h(s) & \text{if } s \in H \end{bmatrix} \quad \text{and} \quad g = \begin{bmatrix} z_1 & \text{if } s \in G_1 \\ z_2 & \text{if } s \in G_2 \\ \ldots\ldots & \ldots\ldots \\ z_n & \text{if } s \in G_n \\ h(s) & \text{if } s \in H \end{bmatrix} \tag{A.9}$$

for some outcomes $z_1 \succeq z_2 \succeq \cdots \succeq z_n$, for some event H (that can be empty), for some act $h \in \mathcal{G} \Cap H$, and for some partitions $S = F_1 \oplus F_2 \oplus \cdots \oplus F_n \oplus H$ and $S = G_1 \oplus G_2 \oplus \cdots \oplus G_n \oplus H$.

Prove (A.7) and (A.8) by induction with respect to n. If $n = 1$, then $f = g$ and $f \sim g$. Fix $n > 1$ and suppose that (A.7) and (A.8) hold for all acts that differ through $(n-1)$-outcome subacts. Fix arbitrary acts f and g that have the form (A.9).

If $z_i \sim z_j$ for some $i \neq j$, then by P3($\mathcal{R}$), $z_i F_j f \sim f$ and $z_i G_j g \sim g$; the acts $z_i F_j f$ and $z_i G_j g$ differ through $(n-1)$-outcome subacts. Therefore,

$$[f]_p \geqslant (\gg)[g]_p \quad \Rightarrow \quad [z_i F_j f]_p \geqslant [f]_p \geqslant (\gg)[g]_p \geqslant [z_i G_j g]_p \quad \Rightarrow$$
$$z_i F_j f \succeq (\succ) z_i G_j g \quad \Rightarrow \quad f \succeq (\succ) g.$$

Without loss of generality, suppose that $z_1 \succ z_2 \succ \cdots \succ z_n$. Then $[f]_p \geqslant [g]_p$ if and only if $p(F_1^k) \geq p(G_1^k)$ for all $k = 1 \ldots n$; $[f]_p \gg [g]_p$ if and only if $p(F_1^k) \geq p(G_1^k)$ for all $k = 1 \ldots n$ and $p(F_1^k) > p(G_1^k)$ for some $k \in [1, n]$. Consider several cases.

Case I. $p(F_1^k) > p(G_1^k)$ for all $k = 1 \ldots n-1$, and

$$\alpha(F_1, F_2, \ldots, F_n, H, G_1, G_2, \ldots, G_n) \subset \mathcal{R}.$$

Take $\varepsilon > 0$ such that $p(F_1^k) > p(G_1^k) + \varepsilon$ for all $k = 1 \ldots n-1$. Then

$$\begin{aligned}
&p(F_1 \setminus G_1) = \sum_{i=i}^{n} v_i, \quad \text{where} \\
&v_i = p(F_i \cap G_1) \quad \text{for } i = 2, \ldots, n, \quad \text{and} \\
&v_1 = p(F_1 \setminus G_1) - \sum_{i=2}^{n} v_i = p(F_1 \setminus G_1) - p(G_1 \setminus F_1) = p(F_1) - p(G_1) > \varepsilon.
\end{aligned}$$

By Lemma A.2, there exists a partition $\{F_1 \setminus G_1 = A_1^n\}$ such that $p(A_1) > v_1 - \varepsilon$ and $p(A_i) > v_i$ for $i = 2, \ldots, n$. Let

$$f_1 = \begin{bmatrix} z_2 & \text{if } s \in A_1 \\ f(s) & \text{if } s \notin A_1 \end{bmatrix}.$$

By P3($\mathcal{R}$), $f_1 \succeq f$. For all $i = 2 \ldots n$, let

$$f_i = \begin{bmatrix} z_1 & \text{if } s \in F_i \cap G_1 \\ z_i & \text{if } s \in A_i \\ f_{i-1}(s) & \text{if } s \notin A_i \text{ and } s \notin F_i \cap G_1 \end{bmatrix}.$$

As $A_i \succ_0 F_i \cap G_1$, then by P4*($\mathcal{R}$),

$$f_i \prec \begin{bmatrix} z_1 & \text{if } s \in A_i \\ z_i & \text{if } s \in F_i \cap G_1 \\ f_{i-1}(s) & \text{if } s \notin A_i \text{ and } s \notin F_i \cap G_1 \end{bmatrix} = f_{i-1}.$$

The acts f_n and g can be written as

$$f_n = \begin{bmatrix} z_2 & \text{if } s \in E_2 \\ \multicolumn{2}{c}{\cdots\cdots\cdots\cdots\cdots} \\ z_n & \text{if } s \in E_n \\ z_1 & \text{if } s \in G_1 \\ h(s) & \text{if } s \in H \end{bmatrix} \quad \text{and} \quad g = \begin{bmatrix} z_2 & \text{if } s \in G_2 \\ \multicolumn{2}{c}{\cdots\cdots\cdots\cdots\cdots} \\ z_n & \text{if } s \in G_n \\ z_1 & \text{if } s \in G_1 \\ h(s) & \text{if } s \in H \end{bmatrix},$$

where $E_2 = (F_2 \oplus A_1 \oplus A_2) \ominus (F_2 \cap G_1)$, and $E_i = (F_i \oplus A_i) \ominus (F_i \cap G_1)$ for $i = 3 \ldots n$. Thus, f_n and g differ through $(n-1)$-outcome subacts. For all $k = 2 \ldots n$,

$$\sum_{i=2}^{k} p(E_i) = p(A_1) + \sum_{i=2}^{k} (p(F_i) + p(A_i) - p(F_i \cap G_1)) \geq$$
$$(p(F_1) - p(G_1) - \varepsilon) + \sum_{i=2}^{k} p(F_i) = \sum_{i=1}^{k} p(F_i) - p(G_1) - \varepsilon \geq \sum_{i=2}^{k} p(G_i).$$

Thus, $[f_n]_p \geqslant [g]_p$, and $f_n \succeq g$. The preference $f \succeq g$ follows by transitivity from

$$f \succeq f_1 \succ f_2 \succ \cdots \succ f_n \succeq g.$$

Case II. $p(F_1^k) > p(G_1^k)$ for all $k = 1 \ldots n-1$. By Lemma A.2, there exists a partition $\{\neg C = E_1^n\} \subset \mathcal{R} \Cap (\alpha(F_1, \ldots, F_n) \cup \alpha(G_1, \ldots, G_n))$ such that

$$p(F_1^k) > p(E_1^k) > p(G_1^k) \quad \text{for all } k = 1 \ldots n-1.$$

Let

$$f' = \begin{bmatrix} z_1 & \text{if } s \in E_1 \\ \multicolumn{2}{c}{\cdots\cdots\cdots\cdots\cdots} \\ z_n & \text{if } s \in E_n \\ h(s) & \text{if } s \in H \end{bmatrix}.$$

By (A.1), $\alpha(F_1, \ldots, F_n, E_1, \ldots, E_n) \subset \mathcal{R}$ and $\alpha(G_1, \ldots, G_n, E_1, \ldots, E_n) \subset \mathcal{R}$. Case I implies that $f \succeq f' \succeq g$ and hence, that $f \succeq g$.

Case III. $p(F_1^k) \geq p(G_1^k)$ for all $k = 1 \ldots n$. Suppose that, contrary to (A.7), $g \succ f$. If F_n is null, then by P3($\mathcal{R}$), $z_{n-1}F_n f \sim f$ and $z_{n-1}G_n g \succeq g$; the acts $z_{n-1}F_n f$ and $z_{n-1}G_n g$ differ through $(n-1)$-outcome subacts; the first-order dominance

$$[z_{n-1}F_n f]_p \geqslant [f]_p \geqslant [g]_p \geqslant [z_{n-1}G_n g]_p$$

implies that $f \sim z_{n-1}F_n f \succeq z_{n-1}G_n g \succeq g$ and hence, $f \succeq g$. If F_n is non-null, then by P6($\mathcal{R}$), there exists a non-null partition $\{F_n = E_1^m\}$ such that $g \succ x_1 E_i f$ for all $i = 1 \ldots m$. Then $p(F_1^k \oplus E_1) > p(G_1^k)$ for all $k = 1 \ldots n-1$. Case II implies that $x_1 E_1 f \succ g$, which contradicts $g \succ x_1 E_1 f$. Thus, $f \succeq g$, and the proof of (A.7) is complete.

Case IV. $p(F_1^k) \geq p(G_1^k)$ for all $k = 1 \ldots n$ and $p(F_1^k) > p(G_1^k)$ for some $k \in [1, n]$. There exists k such that $p(F_1^k) > p(G_1^k)$ and $p(F_1^{k+1}) = p(G_1^{k+1})$. Note that G_{k+1} is non-null. By Lemma A.2, there exists a non-null partition $G_{k+1} = C \oplus D$ such that $p(F_1^k) \geq p(D) + p(G_1^k)$. Then $l(f) \geqslant l(z_k D g)$, and Case III implies $f \succeq z_k D g$. By P3($\mathcal{R}$), $z_k D g \succ g$. Thus $f \succ g$, and the proof of (A.8) is complete.

A.3 Sufficiency of Axioms

Suppose that $\succeq$ satisfies P1($\mathcal{R}$), P3($\mathcal{R}$), P4*($\mathcal{R}$), P5($\mathcal{R}$), P6($\mathcal{R}$). Let $p : \mathcal{R} \to [0, 1]$ be the probability measure given by (A.6).

Suppose that the set of outcomes X is finite. (The case of infinite X is analyzed later.) Then there exist outcomes x^* and x_* such that $x^* \succeq x \succeq x_*$ for all $x \in X$. Write $l \geqslant l' + o$ if lotteries $l, l' \in \mathcal{L}$ are such that $l \geqslant l''$ for all $l'' \in \mathcal{L}$ in some neighborhood of l'; write $l + o \geqslant l'$ if $l'' \geqslant l'$ for all $l'' \in \mathcal{L}$ in some neighborhood of l.

Lemma A.3. *For any acts $g \succ g'$, there exist acts $h \succ h'$ such that $[g]_p + o \geqslant [h]_p$ and $[h']_p \geqslant [g']_p + o$.*

Proof. Fix acts $g \succ g'$. Let the support of the lottery $l = [g]_p$ consist of outcomes $y_1 \succeq y_2 \succeq \cdots \succeq y_n$. The event $g^{-1}(y_1)$ is non-null. By P6($\mathcal{R}$), there exists a non-null partition $\{g^{-1}(y_1) = E_1^m\}$ such that $x_* E_1 g \succ g'$. Let $h = x_* E_1 g$. Fix an arbitrary lottery $l'' \in \mathcal{L}$ such that $\|[g]_p - l''\| < p(E_1)$. For each $x \in X$, one of the following cases applies.

1. $x \sim x_*$. Then $l''(Y_x) = 1 = [h]_p(Y_x)$.

2. $y_1 \succeq x \succ x_*$. Then $l''(Y_x) \geq [g]_p(Y_x) - p(E_1)$ because $\|[g]_p - l''\| < p(E_1)$, and $[g]_p(Y_x) - p(E_1) = [h]_p(Y_x)$ because $g^{-1}(Y_x) = h^{-1}(Y_x) \oplus E_1$.

3. $x \succ y_1$. Then $l''(Y_x) \geq 0 = [h]_p(Y_x)$.

Therefore, $l''(Y_x) \geq [h]_p(Y_x)$ for all $x \in X$, and $l'' \geqslant [h]_p$. Thus, $[g]_p + o \geqslant [h]_p$. Analogously, $h \succ g'$ implies that there exists a non-null E_1 such that $h \succ h' = x^* E_1 g'$ and $[h']_p \geqslant [g']_p + o$. $\square$

Lemma A.4. *For any lottery $l \in \mathcal{L}$ and for any non-null event E, there exists an act $h \in \mathcal{G} \Cap \{\neg E\}$ such that $\|[h]_p - l\| < p(E)$ and $[x^* E h]_p \geqslant l + o \geqslant [x_* E h]_p$.*

Proof. Fix a lottery l and a non-null event E. Let the support of l consist of outcomes $y_1 \succeq y_2 \succeq \cdots \succeq y_n$. Take $\varepsilon = \frac{1}{2} \cdot p(E) \cdot \min\{l(y_1), l(y_n)\}$. By Lemma A.2, there exist partitions $\{E = A_1^n\}$ and $\{\neg E = B_1^n\}$ such that for all $i = 1 \ldots n$,

$$p(E) \cdot (1 + \varepsilon) \cdot l(y_i) \geq p(A_i) \geq p(E) \cdot (1 - \varepsilon) \cdot l(y_i), \quad \text{and}$$
$$(1 - p(E)) \cdot (1 + \varepsilon) \cdot l(y_i) \geq p(B_i) \geq (1 - p(E)) \cdot (1 - \varepsilon) \cdot l(y_i).$$

Take $h \in \mathcal{G}$ such that $h(s) = y_i$ if $s \in A_i \oplus B_i$ for $i = 1 \ldots n$. Then

$$\|[h]_p - l\| = \sum_{i=1}^{n} |p(A_i) + p(B_i) - l(y_i)| \leq \varepsilon \cdot \sum_{i=1}^{n} l(y_i) = \varepsilon < p(E).$$

Let $l^* = [x^* E h]_p$ and $l_* = [x_* E h]_p$. Fix an arbitrary lottery $l'' \in \mathcal{L}$ such that $\|l - l''\| < \varepsilon$. For each $x \in X$, one of the following cases applies.

1. $x \sim x_*$. Then $l^*(Y_x) = 1 = l''(Y_x) = l_*(Y_x)$.

2. $y_n \succeq x \succ x_*$. Then $l^*(Y_x) = 1 \geq l''(Y_x) \geq l(Y_x) - \varepsilon = 1 - \varepsilon \geq 1 - p(E) = l_*(Y_x)$.

3. $y_1 \succeq x \succ y_n$. Then $p(E) \cdot (1 - l(Y_x)) \geq p(E) \cdot l(y_n) \geq 2 \cdot \varepsilon$ and

$$l^*(Y_x) = p(E) + \sum_{i: y_i \succeq x} p(B_i) \geq p(E) + (1 - p(E)) \cdot (1 - \varepsilon) \cdot l(Y_x) \geq$$
$$l(Y_x) + p(E) \cdot (1 - l(Y_x)) - \varepsilon \geq l(Y_x) + 2 \cdot \varepsilon - \varepsilon = l(Y_x) + \varepsilon \geq l''(Y_x).$$

Also, $p(E) \cdot l(Y_x) \geq p(E) \cdot l(y_1) \geq 2 \cdot \varepsilon$, and

$$l_*(Y_x) = \sum_{i: y_i \succeq x} p(B_i) \leq (1 - p(E)) \cdot (1 + \varepsilon) \cdot l(Y_x) \leq$$
$$l(Y_x) - p(E) \cdot l(Y_x) + \varepsilon \leq l(Y_x) - 2 \cdot \varepsilon + \varepsilon = l(Y_x) - \varepsilon \leq l''(Y_x).$$

4. $x \succ y_1$. Then $l^*(Y_x) = p(E) \geq 0 + \varepsilon = l(Y_x) + \varepsilon \geq l''(Y_x) \geq 0 = l_*(Y_x)$.

Therefore, $l^*(Y_x) \geq l''(Y_x) \geq l_*(Y_x)$ for all $x \in X$, and $l^* \geqslant l'' \geqslant l_*$ for all l'' in the ε-neighborhood of l. Thus, $l^* \geqslant l + o \geqslant l_*$. $\square$

Theorem 4.1(I)

Suppose that the preference $\succeq$ satisfies P1($\mathcal{R}$), P3($\mathcal{R}$), P4*($\mathcal{R}$), P5($\mathcal{R}$), P6($\mathcal{R}$). Let $\mathcal{L}_p$ be the set of lotteries that are induced by acts via the probability p. Show that $\succeq$ is represented by

$$f \succeq f' \quad \Leftrightarrow \quad [f]_p \succeq_1 [f']_p \quad \text{for all acts } f, f' \in \mathcal{G}, \tag{A.10}$$

where the binary relation $\succeq_1$ ranks only $\mathcal{L}_p$ and is non-degenerate, complete, transitive, continuous, and strictly monotonic. Moreover, such $\succeq_1$ is unique.

Proof. For all lotteries $l, l' \in \mathcal{L}_p$, let $l \succeq_1 l$ if $l = [g]_p$ and $l' = [g']_p$ for some acts $g \succeq g'$. Then for all acts f and f',

$$f \succeq f' \quad \Rightarrow \quad [f]_p \succeq_1 [f']_p;$$
$$[f]_p \succeq_1 [f']_p \quad \Rightarrow \quad [f]_p = [g]_p \text{ and } [f']_p = [g']_p \text{ for some } g \succeq g' \quad \Rightarrow \{A.7\}$$
$$f \sim g \text{ and } f' \sim g' \text{ for some } g \succeq g' \quad \Rightarrow \quad f \succeq f'.$$

Thus, representation (A.10) holds.

The relation $\succeq_1$ is non-degenerate, complete and transitive because $\succeq$ is non-degenerate, complete and transitive. By (A.7) and (A.8), $\succeq_1$ is strictly monotonic. Show that $\succeq_1$ is continuous. Fix arbitrary lotteries $l, l' \in \mathcal{L}_p$ such that $l \succ_1 l'$. Then $l = [g]_p$ and $l' = [g']_p$ for some acts $g \succ g'$. By Lemma A.3, there exist acts $h \succ h'$ such that $l + o \geqslant [h]_p$ and $[h']_p \geqslant l' + o$. Therefore, for all $l'' \in \mathcal{L}_p$ in some neighborhood of l, $l'' \geqslant [h]_p \succ_1 [h']_p \geqslant [g']_p = l'$ and $l'' \succ_1 l'$ because $\succeq_1$ is monotonic and transitive. Thus, the set $\{l'' \in \mathcal{L}_p : l'' \succ_1 l'\}$ is open in $\mathcal{L}_p$ for all $l' \in \mathcal{L}_p$. Analogously, the set $\{l'' \in \mathcal{L}_p : l \succ_1 l''\}$ is open in $\mathcal{L}_p$ for all $l \in \mathcal{L}_p$.

The uniqueness of $\succeq_1$ follows from representation (A.10). $\square$

Section 5.2 provides an example, where $\succeq_1$ cannot be extended from $\mathcal{L}_p$ to a continuous weak order on the set $\mathcal{L}$ of all lotteries. To account for such situations, Theorem 4.1(I) retains the risk preference $\succeq_1$ as a binary relation on $\mathcal{L}_p$.

Theorem 3.1

Suppose that $\succeq$ satisfies P1($\mathcal{R}$), P2($\mathcal{R}$), P3($\mathcal{R}$), P4($\mathcal{R}$), P5($\mathcal{R}$), and P6($\mathcal{R}$).[27] Show that $\succeq$ is represented by

$$f \succeq f' \quad \Leftrightarrow \quad [f]_p \succeq_1 [f']_p \quad \text{for all } f, f' \in \mathcal{G}, \tag{A.11}$$

where the *extended risk preference* $\succeq_1$ is non-degenerate, complete, transitive, continuous, and mixture separable on the set $\mathcal{L}$ of *all lotteries*. Moreover, such $\succeq_1$ is unique.

Proof. For all $l, l' \in \mathcal{L}$, let $l \succeq_1 l'$ if there exist sequences of acts $\{g_i\}_{i=1}^{\infty}$ and $\{g_i'\}_{i=1}^{\infty}$ such that

$$\lim_{i\to\infty} [g_i]_p = l, \quad \lim_{i\to\infty} [g_i']_p = l', \text{ and } g_i \succeq g_i' \text{ for all } i = 1, 2, \ldots. \tag{A.12}$$

Analyze properties of $\succeq_1$ in several steps.

Step 1. For all lotteries $l, l' \in \mathcal{L}$,

$$l \succeq_1 l' \quad \Leftrightarrow \quad g \succeq g' \quad \text{for all } g, g' \in \mathcal{G} \,:\, [g]_p \geqslant l + o \text{ and } l' + o \geqslant [g']_p \tag{A.13}$$

$$l' \succ_1 l \quad \Leftrightarrow \quad g' \succ g \quad \text{for some } g, g' \in \mathcal{G} \,:\, [g]_p \geqslant l + o \text{ and } l' + o \geqslant [g']_p. \tag{A.14}$$

Fix arbitrary lotteries $l, l' \in \mathcal{L}$. Two cases are possible.

1. The weak preference $g \succeq g'$ holds for all acts g and g' such that $[g]_p \geqslant l + o$ and $l' + o \geqslant [g']_p$. Fix a sequence of non-null events E_i such that $\lim_{i\to\infty} p(E_i) = 0$. By Lemma A.4, there exist sequences $\{h_i\}_{i=1}^{\infty}$ and $\{h_i'\}_{i=1}^{\infty}$ such that for all i, $\|[h_i]_p - l\| < p(E_i)$, $\|[h_i']_p - l'\| < p(E_i)$, $[x^* E_i h_i]_p \geqslant l + o$, and $l' + o \geqslant [x_* E_i h_i']_p$. It follows that the weak preference $x^* E_i h_i \succeq x_* E_i h_i'$ holds. Note that

   $$\|[x^* E_i h_i]_p - l\| \leq \|[x^* E_i h_i]_p - [h_i]_p\| + \|[h_i]_p - l\| < 3 \cdot p(E_i),$$

   that is, $\lim_{i\to\infty} [x^* E_i h_i]_p = l$. Analogously, $\lim_{i\to\infty} [x_* E_i h_i']_p = l'$. Then $g_i = x^* E_i h_i$ and $g_i' = x_* E_i h_i'$ satisfy (A.12), and hence, $l \succeq l'$.

2. The strict preference $g' \succ g$ holds for some acts g and g' such that $[g]_p \geqslant l + o$ and $l' + o \geqslant [g']_p$. Suppose that some sequences $\{g_i\}_{i=1}^{\infty}$ and $\{g_i'\}_{i=1}^{\infty}$ satisfy (A.12). Then $[g]_p \geqslant \lim_{i\to\infty} [g_i]_p + o$ and $\lim_{i\to\infty} [g_i']_p + o \geqslant [g']_p$ imply that $[g]_p \geqslant [g_i]_p$ and $[g_i']_p \geqslant [g']_p$ for sufficiently large i. By (A.7), $g \succeq g_i \succeq g_i' \succeq g'$ which contradicts $g' \succ g$. Thus, no sequences $\{g_i\}_{i=1}^{\infty}$ and $\{g_i'\}_{i=1}^{\infty}$ satisfy (A.12) and $l \succeq_1 l'$ does not hold. To show that $l' \succeq_1 l$, fix a sequence of non-null events E_i such that $\lim_{i\to\infty} p(E_i) = 0$ and sequences of acts $\{h_i\}_{i=1}^{\infty}$ and $\{h_i'\}_{i=1}^{\infty}$ such that $\lim_{i\to\infty} [x_* E_i h_i]_p = l$, $\lim_{i\to\infty} [x^* E_i h_i']_p = l'$, $l + o \geqslant [x_* E_i h_i]_p$ and $[x^* E_i h_i']_p \geqslant l' + o$. Then for all i, $x^* E_i h_i' \succeq g' \succ g \succeq x_* E_i h_i$ because $[x^* E_i h_i']_p \geqslant l' \geqslant [g']_p$ and $[g]_p \geqslant l \geqslant [x_* E_i h_i]_p$. By definition, $l' \succeq_1 l$. Thus, $l' \succ_1 l$.

[27] This set of axioms is stronger than P1($\mathcal{R}$), P3($\mathcal{R}$), P4*($\mathcal{R}$), P5($\mathcal{R}$), P6($\mathcal{R}$) because P2($\mathcal{R}$) and P4($\mathcal{R}$) imply P4*($\mathcal{R}$).

By (A.13) and (A.14), $\succeq_1$ is complete.

Step 2. Show that $\succeq_1$ is transitive. Fix arbitrary lotteries l, l', l'' such that $l \succeq_1 l' \succeq_1 l''$. Suppose that $l'' \succ_1 l$. By (A.14), there exist acts $g'' \succ g$ such that $[g]_p \geqslant l+o$ and $l''+o \geqslant [g'']_p$. Use P6($\mathcal{R}$) to construct a non-null event E and an act $h \in \mathcal{G} \Cap \{\neg E\}$ such that[28]

$$g'' \succ x^* Eh \succ x_* Eh \succ g.$$

By Lemma A.4, there exists an act $h' \in \mathcal{G} \Cap \{\neg E\}$ such that $[x^* Eh']_p \geqslant l' + o \geqslant [x_* Eh']_p$. Two cases are possible.

1. $x^* Eh \succeq x^* Eh'$. Then $g'' \succ x^* Eh \succeq x^* Eh'$ and by (A.14), $l'' \succ_1 l'$.

2. $x^* Eh' \succeq x^* Eh$. Then by P2($\mathcal{R}$), $x_* Eh' \succeq x_* Eh \succ g$ and by (A.14), $l' \succ_1 l$.

Either case contradicts $l \succeq_1 l' \succeq_1 l''$. Thus, the strict preference $l'' \succ_1 l$ is impossible and $l \succeq_1 l''$ holds.

Step 3. Show that $\succeq_1$ is weakly monotonic. Fix arbitrary lotteries l and l' such that $l \geqslant l'$. For all acts g and g' such that $[g]_p \geqslant l + o$ and $l' + o \geqslant [g']_p$, the weak preference $g \succeq g'$ follows by (A.7) from $[g]_p \geqslant l \geqslant l' \geqslant [g']_p$. By (A.13), $l \succeq_1 l'$.

Step 4. Show that $\succeq_1$ is continuous. Fix arbitrary lotteries l and l' such that $l \succ_1 l'$. By (A.14), there exist acts $g \succ g'$ such that $l + o \geqslant [g]_p$ and $[g']_p \geqslant l' + o$. For all l'' in some neighborhood of l, $l'' + o \geqslant [g]_p$ and by (A.14), $l'' \succ_1 l'$. Thus, the set $\{l'' \in \mathcal{L} : l'' \succ_1 l'\}$ is open in $\mathcal{L}$ for all $l' \in \mathcal{L}$. Analogously, the set $\{l'' \in \mathcal{L} : l \succ_1 l''\}$ is open in $\mathcal{L}$ for all $l \in \mathcal{L}$.

Step 5. Show that representation (A.11) holds. Fix acts f and f'. If $f \succeq f'$, then (A.12) holds for $g_i = f$ and $g'_i = f'$; therefore, $[f]_p \succeq_1 [f']_p$. If $f \succ f'$, then by Lemma A.3, there exist acts $h \succ h'$ such that $[f]_p + o \geqslant [h]_p$ and $[h']_p \geqslant [f']_p + o$. By (A.14), $[f]_p \succ_1 [f']_p$.

Step 6. Show that $\succeq_1$ is mixture separable. Fix lotteries l, l', and l'' such that $l \succeq_1 l'$. Write the supports of the lotteries l, l', and l'' as $y_1 \succeq \cdots \succeq y_n$, $y'_1 \succeq \cdots \succeq y'_{n'}$ and $y''_1 \succeq \cdots \succeq y''_{n''}$ respectively. Fix arbitrary $\varepsilon > 0$ and an event H such that $0 < p(H) < 1$. By Lemma A.2,

[28] More precisely, construct E and h as follows. The event $g^{-1}(x)$ is non-null for some outcome $x \prec x^*$; otherwise by P3($\mathcal{R}$), $g \succeq x^* \succeq g''$. By P6($\mathcal{R}$), there exists a non-null partition $\{g^{-1}(x) = A_1^m\}$ such that $g'' \succ x^* A_1 g$. Let $h = x^* A_1 g$. Then $h \succ x A_1 g = g$. By P6($\mathcal{R}$), there exists a non-null partition $\{A_1 = B_1^n\}$ such that $x_* B_1 h \succ g$. Let $E = B_1$. Then $g'' \succ x^* A_1 h = h = x^* Eh \succ x_* Eh \succ g$.

there exist partitions

$$\{H = A_1^n\} : \begin{cases} 1 \le \frac{p(A_i)}{p(H)\cdot l(y_i)} \le 1+\varepsilon \quad \text{for } i = 1\dots n-1, \text{ and} \\ 1-\varepsilon \le \frac{p(A_n)}{p(H)\cdot l(y_n)} \le 1\ ; \end{cases}$$

$$\{\neg H = B_1^n\} : \begin{cases} 1 \le \frac{p(B_i)}{p(\neg H)\cdot l(y_i)} \le 1+\varepsilon \quad \text{for } i = 1\dots n-1, \text{ and} \\ 1-\varepsilon \le \frac{p(B_n)}{p(\neg H)\cdot l(y_n)} \le 1\ ; \end{cases}$$

$$\{H = C_1^{n'}\} : \begin{cases} 1-\varepsilon \le \frac{p(C_i)}{p(H)\cdot l'(y'_i)} \le 1 \quad \text{for } i = 1\dots n'-1, \text{ and} \\ 1 \le \frac{p(C_{n'})}{p(H)\cdot l'(y'_{n'})} \le 1+\varepsilon\ ; \end{cases}$$

$$\{\neg H = D_1^{n'}\} : \begin{cases} 1-\varepsilon \le \frac{p(D_i)}{p(\neg H)\cdot l'(y'_i)} \le 1 \quad \text{for } i = 1\dots n'-1, \text{ and} \\ 1 \le \frac{p(D_{n'})}{p(\neg H)\cdot l'(y'_{n'})} \le 1+\varepsilon\ ; \end{cases}$$

$$\{H = F_1^{n''}\} :\ 1-\varepsilon \le \frac{p(F_i)}{p(H)\cdot l''(y''_i)} \le 1+\varepsilon \quad \text{for } i = 1\dots n''\ ;$$

$$\{\neg H = G_1^{n''}\} :\ 1-\varepsilon \le \frac{p(G_i)}{p(\neg H)\cdot l''(y''_i)} \le 1+\varepsilon \quad \text{for } i = 1\dots n''.$$

Take acts f, f' and f'' such that $f(s) = y_i$ if $s \in A_i \oplus B_i$ for $i = 1\dots n$, $f'(s) = y'_i$ if $s \in C_i \oplus D_i$ for $i = 1\dots n'$, and $f''(s) = y''_i$ if $s \in F_i \oplus G_i$ for $i = 1\dots n''$. Then $[f]_p \geqslant l \succeq_1 l' \geqslant [f']_p$. It follows that $[f]_p \succeq_1 [f']_p$ and by (A.11), $f \succeq f'$. Two cases are possible.

1. $fHf'' \succeq f'Hf''$. Let $g = fHf''$ and $g' = f'Hf''$. Then

$$\|[g]_p - (\tfrac{1}{2}l + \tfrac{1}{2}l'')\| \le \|[fHf'']_p - (p(H)\cdot l + (1-p(H))\cdot l'')\| + \\ \|(p(H)\cdot l + (1-p(H))\cdot l'') - (\tfrac{1}{2}l + \tfrac{1}{2}l'')\| \le \varepsilon + 2\cdot|p(H) - \tfrac{1}{2}|.$$

   Analogously, $\|[g']_p - (\frac{1}{2}l' + \frac{1}{2}l'')\| \le \varepsilon + 2\cdot|p(H) - \frac{1}{2}|$.

2. $f'Hf'' \succ fHf''$. If $f''Hf' \succ f''Hf$ then by P2($\mathcal{R}$), $f' \succ f'Hf \succ fHf = f$ which contradicts $f \succeq f'$. Hence, the strict preference $f''Hf' \succ f''Hf$ is impossible, and $f''Hf \succeq f''Hf'$. Let $g = f''Hf$ and $g' = f''Hf'$. Then

$$\|[g]_p - (\tfrac{1}{2}l + \tfrac{1}{2}l'')\| \le \varepsilon + 2\cdot|p(H) - \tfrac{1}{2}|;$$
$$\|[g']_p - (\tfrac{1}{2}l' + \tfrac{1}{2}l'')\| \le \varepsilon + 2\cdot|p(H) - \tfrac{1}{2}|.$$

As $\varepsilon > 0$ and $|p(H) - \frac{1}{2}|$ can be arbitrarily small, there exist sequences $\{g_i\}_{i=1}^{\infty}$ and $\{g'_i\}_{i=1}^{\infty}$ such that $\{[g_i]_p\}_{i=1}^{\infty}$ and $\{[g'_i]_p\}_{i=1}^{\infty}$ converge to $\frac{1}{2}l + \frac{1}{2}l''$ and $\frac{1}{2}l' + \frac{1}{2}l''$ respectively, and $g_i \succeq g'_i$ for all i. Thus, $\frac{1}{2}l + \frac{1}{2}l'' \succeq_1 \frac{1}{2}l' + \frac{1}{2}l''$.

Step 7. Suppose that $\succeq$ is represented by

$$f \succeq f' \quad \Leftrightarrow \quad [f]_p \succeq'_1 [f']_p \quad \text{for all } f, f' \in \mathcal{G},$$

where the binary relation $\succeq'_1$ ranks $\mathcal{L}$ and is complete, transitive, continuous, and weakly monotonic. Fix lotteries $l \succeq'_1 l'$. For all acts g and g' such that $[g]_p \geqslant l + o$ and $l' + o \geqslant [g']_p$. $[g]_p \succeq'_1 l \succeq'_1 l' \succeq'_1 [g']_p$ implies $g \succeq g'$. By (A.13), $l \succeq_1 l'$. Fix lotteries $l \succeq_1 l'$ and sequences $\{g_i\}_{i=1}^{\infty}$ and $\{g'_i\}_{i=1}^{\infty}$ that satisfy (A.12). Then $\lim_{i\to\infty}[g_i]_p = l$, $\lim_{i\to\infty}[g'_i]_p = l'$, and for all i, $[g_i]_p \succeq'_1 [g'_i]_p$ because $g_i \succeq g'_i$. By continuity, $l \succeq'_1 l'$. Thus, $\succeq'_1 = \succeq_1$. $\square$

The version of the von Neumann–Morgenstern Theorem due to Herstein–Milnor [12] asserts that $\succeq_1$ is represented by expected utility:

$$U(l) = \sum_{x \in X} u(x) \cdot l(x) \qquad \text{for } l \in \mathcal{L},$$

where $u : X \to \mathbb{R}$ is a non-constant utility index, which is unique up to a positive linear transformation.

Theorem 4.1(II)

Suppose that $\succeq$ satisfies P1($\mathcal{R}$), P3($\mathcal{R}$), P4*($\mathcal{R}$), P5($\mathcal{R}$), and P6*($\mathcal{R}$).[29]

With a slight abuse of notation, let x denote the degenerate lottery that yields the outcome x with probability 1. Call a binary relation $\succeq_1$ on $\mathcal{L}$ *semi-strictly monotonic* if it is weakly monotonic and if for all outcomes $x \succ_1 x'$ and weights $\alpha, \beta \in [0, 1]$,

$$\alpha > \beta \quad \Rightarrow \quad \alpha \cdot x + (1 - \alpha) \cdot x' \succ_1 \beta \cdot x + (1 - \beta) \cdot x'.$$

Show that $\succeq$ is represented by

$$f \succeq f' \quad \Leftrightarrow \quad [f]_p \succeq_1 [f']_p \quad \text{for all } f, f' \in \mathcal{G},$$

where the binary relation $\succeq_1$ ranks the set $\mathcal{L}$ of all lotteries and is non-degenerate, complete, transitive, continuous, and semi-strictly monotonic.

Proof. One can adapt the proof from Theorem 3.1 by making the following changes.

Show that $\succeq_1$ is transitive. Fix arbitrary lotteries l, l', l'' such that $l \succeq_1 l' \succeq_1 l''$. Suppose that $l'' \succ_1 l$. By (A.14), $[g]_p \geqslant l + o$ and $l'' + o \geqslant [g'']_p$ for some acts $g'' \succ g$. There exists an act f such that $f(S) = X$ and $g'' \succ f \succ g$.[30] Take a finite collection $\mathcal{E} \subset \mathcal{R}$ such that the acts h and g are $\mathcal{E}$-measurable. By P6*($\mathcal{R}$), there exists a partition $\{S = S_1^m\} \subset \mathcal{R} \Cap \mathcal{E}$ such that for all i and for all acts $h \in \mathcal{G} \Cap \{\neg S_i\}$,

$$h \succeq f \quad \Rightarrow \quad x_* S_i h \succ g.$$

Let $E = S_k$ be a non-null element of this partition. By Lemma A.4, there exists an act $h' \in \mathcal{G} \Cap \{\neg E\}$ such that $[x^* E h']_p \geqslant l' + o \geqslant [x_* E h']_p$. Two cases are possible.

1. $f \succeq x^* E h'$. Then $g'' \succ f \succeq x^* E h'$ and by (A.14), $l'' \succ_1 l'$.
2. $x^* E h' \succeq f$. Then by P6*($\mathcal{R}$), $x_* E h' \succ g$ and by (A.14), $l' \succ_1 l$.

[29] This set of axioms is stronger than P1($\mathcal{R}$), P3($\mathcal{R}$), P4*($\mathcal{R}$), P5($\mathcal{R}$), P6($\mathcal{R}$) because P6*($\mathcal{R}$) implies P6($\mathcal{R}$).

[30] More precisely, construct f as follows. Find a non-null event E and $h \in \mathcal{G} \Cap \{\neg E\}$ such that $g'' \succ x^* E h \succ x_* E h \succ g$. Partition E into $|X|$ non-null events, and take an act f such that f yields different outcomes on all elements of this partition and $f = h$ on $\neg E$. Then $f(S) = X$, and by P3($\mathcal{R}$), $x^* E h \succeq f \succeq x_* E h$. Thus, $g'' \succ f \succ g$.

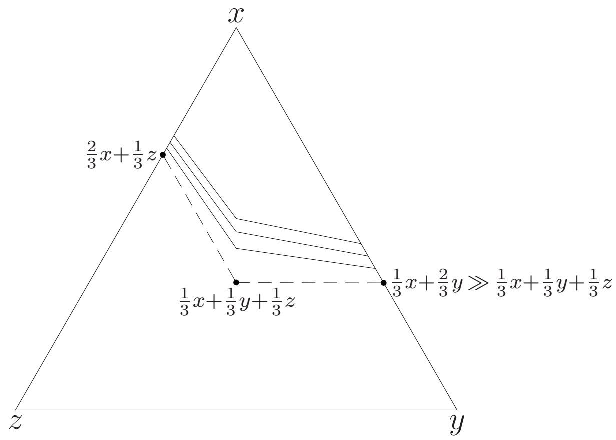

Figure 2: Extended risk preference that is not strictly monotonic

Either case contradicts $l \succeq_1 l' \succeq_1 l''$. Thus, the strict preference $l'' \succ_1 l$ is impossible and $l \succeq_1 l''$ holds.

Show that $\succeq_1$ is semi-strictly monotonic. Fix outcomes $x \succ x'$ and weights $\alpha > \beta$. By Lemma A.2, there exist events A and B such that $\alpha > p(A) > p(B) > \beta$. Then

$$\begin{aligned}
&\alpha \cdot x + (1-\alpha) \cdot x' \geqslant [xAx']_p \succ_1 [xBx']_p \geqslant \beta \cdot x + (1-\beta) \cdot x' \quad \Rightarrow \\
&\alpha \cdot x + (1-\alpha) \cdot x' \succeq_1 [xAx']_p \succ_1 [xBx']_p \succeq_1 \beta \cdot x + (1-\beta) \cdot x'.
\end{aligned}$$

Thus, $\alpha \cdot x + (1-\alpha) \cdot x' \succ_1 \beta \cdot x + (1-\beta) \cdot x'$. $\square$

As the weak order $\succeq_1$ is continuous and semi-strictly monotonic on all of $\mathcal{L}$, then for any $l \in \mathcal{L}$, there exists a unique value $V(l) \in [0,1]$ such that

$$l \sim_1 V(l) \cdot x^* + (1 - V(l)) \cdot x_*.$$

The function $V : \mathcal{L} \to [0,1]$ so defined represents $\succeq_1$ and hence, for all acts $f, g \in \mathcal{G}$,

$$f \succeq g \quad \Leftrightarrow \quad [f]_p \succeq_1 [g]_p \quad \Leftrightarrow \quad V([f]_p) \geq V([g]_p).$$

The function $V : \mathcal{L} \to [0,1]$ is continuous because $\succeq_1$ is continuous and is uniformly continuous because $\mathcal{L}$ is a compact set. When restricted to $\mathcal{L}_p$, the function V is uniformly continuous and strictly monotonic because $\succeq_1$ is strictly monotonic on $\mathcal{L}_p$.

Note that in general, the extended risk preference $\succeq_1$ and the representation V need not be strictly monotonic on all of $\mathcal{L}$. Indeed, it is intuitive that indifference curves that are strictly monotonic over $\mathcal{L}_p$ can converge to curves that are not strictly monotonic. Figure A.3 illustrates this intuition in the coin-tossing framework of Section 5.2 where there is no event $A \in \mathcal{R}$ such that $p(A) = \frac{1}{3}$. We skip a formal example.

Extensions to the General Case

Consider the general case when the set of outcomes X is infinite.

The construction of risk preference $\succeq_1$ in Theorem 4.1(I) remains unchanged. This binary relation is non-degenerate, complete, transitive, and strictly monotonic on $\mathcal{L}_p$. Given a finite $Y \subset X$, $\succeq_1$ is continuous on $\mathcal{L}(Y) \cap \mathcal{L}_p$ and by definition, is continuous on $\mathcal{L}_p$.

Suppose that $\succeq$ satisfies P1($\mathcal{R}$), P2($\mathcal{R}$), P3($\mathcal{R}$), P4($\mathcal{R}$), P5($\mathcal{R}$), and P6($\mathcal{R}$), as in Theorem 3.1. Fix outcomes $x \succ x'$, and let $p : \mathcal{R} \to [0, 1]$ be the probability measure given by (A.6). For all finite $Y \subset X$ such that $x, x' \in Y$, the preference over acts $f \in \mathcal{G}$ that have range in Y, $f(S) \subset Y$, has a unique expected utility representation

$$U(f) = \sum_{y \in Y} u(y) \cdot p(f^{-1}(y))$$

such that $u(x) = 1$ and $u(x') = 0$. As all of these representations are unique, then

$$U(f) = \sum_{x \in X} u(x) \cdot p(f^{-1}(x))$$

represents $\succeq$ over all of $\mathcal{G}$.

Finally, one can construct a utility representation V in the general case of Theorem 4.1 analogously to Step 5 of Machina–Schmeidler's proof of their Theorem 2.

A.4 Necessity of Axioms

Theorem 4.1(I)

Suppose that $\succeq$ on $\mathcal{G}$ is represented by

$$f \succeq g \quad \Leftrightarrow \quad [f]_p \succeq_1 [g]_p \qquad \text{for } f, g \in \mathcal{G},$$

where $p : \mathcal{R} \to [0, 1]$ is a finely ranged probability measure, and the binary relation $\succeq_1$ on $\mathcal{L}_p$ is non-degenerate, complete, transitive, continuous, and strictly monotonic.

Then $\succeq$ satisfies P1($\mathcal{R}$) and P5($\mathcal{R}$) because $\succeq_1$ is non-degenerate, complete and transitive.

Next, for each event $A \in \mathcal{R}$, one of the following cases applies.

1. $p(A) > 0$. Then for all outcomes $x, y \in X$ and for all acts $h \in \mathcal{G} \Cap \{\neg A\}$,

   $$\begin{aligned} &x \succeq (\succ) y \quad \Rightarrow \quad [xAh]_p \geqslant (\gg)[yAh]_p \quad \Rightarrow \\ &[xAh]_p \succeq_1 (\succ_1)[yAh]_p \quad \Rightarrow \quad xAh \succeq (\succ) yAh\,; \end{aligned}$$

   it follows that $x \succeq y$ if and only if $xAh \succeq yAh$.

2. $p(A) = 0$. Then for all outcomes $x, y \in X$ and for all acts $h \in \mathcal{G} \Cap \{\neg A\}$, $[xAh]_p = [yAh]_p$, $[xAh]_p \sim [yAh]_p$, and $xAh \sim yAh$.

Thus, $\succeq$ satisfies P3($\mathcal{R}$).

Next, for all events $E \in \mathcal{R}$, for all partitions $E = A \cup A'$ and $E = B \cup B'$, for all outcomes $x \succ x'$ and $z \succ z'$, and for all acts $h, h' \in \mathcal{G} \Cap \{\neg E\}$,

$$p(A) \geq (>) p(B) \quad \Rightarrow \quad [(xAx')Eh]_p \geqslant (\gg) [(xBx')Eh]_p \quad \Rightarrow$$
$$[(xAx')Eh]_p \succeq_1 (\succ_1) [(xBx')Eh]_p \quad \Rightarrow \quad (xAx')Eh \succeq (\succ) (xBx')Eh.$$

Therefore, $(xAx')Eh \succeq (xBx')Eh \;\Leftrightarrow\; p(A) \geq p(B) \;\Leftrightarrow\; (zAz')Eh' \succeq (zAz')Eh'$. Thus, $\succeq$ satisfies P4*($\mathcal{R}$).

Finally, show that $\succeq$ satisfies P6($\mathcal{R}$). Without loss of generality, X is finite. Fix an outcome $x \in X$, a finite collection $\mathcal{E} \subset \mathcal{R}$, and a pair of $\mathcal{E}$-measurable acts $f \succ g$. By continuity of the risk preference $\succeq_1$, there exists $\delta > 0$ such that $l \succ_1 [g]_p$ for all $l \in \mathcal{L}_p$ in the δ-neighborhood of $[f]_p$, and $[f]_p \succ_1 l'$ for all $l' \in \mathcal{L}_p$ in the δ-neighborhood of $[g]_p$. Partition S into events $\{S_1, \ldots, S_m\} \subset \mathcal{R} \Cap \mathcal{E}$ such that $p(S_i) < \frac{\delta}{2}$ for all $i = 1 \ldots m$. Then $[xS_if]_p$ lies in the δ-neighborhood of $[f]_p$; hence, $[xS_if]_p \succ_1 [g]_p$ and $xS_if \succ g$. Analogously, $f \succ xS_ig$.

Theorem 4.1(II)

Suppose that $\succeq$ is represented by $U(f) = V([f]_p)$ where $V : \mathcal{L} \to \mathbb{R}$ is uniformly continuous. Prove that $\succeq$ satisfies P6*($\mathcal{R}$). Without loss of generality, X is finite. Fix an outcome $x \in X$, a finite collection $\mathcal{E} \subset \mathcal{R}$, and a pair of $\mathcal{E}$-measurable acts $f \succ g$. Let $\varepsilon = V([f]_p) - V([g]_p)$. As V is uniformly continuous, there exists $\delta > 0$ such that for all $l, l' \in \mathcal{L}_p$,

$$\|l - l'\| < \delta \quad \Rightarrow |V(l) - V(l')| < \varepsilon.$$

Partition S into events $\{S_1, \ldots, S_m\} \subset \mathcal{R} \Cap \mathcal{E}$ such that $p(S_i) < \frac{\delta}{2}$ for all $i = 1 \ldots m$. Then for all i and for all acts $h \in \mathcal{G} \Cap \{\neg S_i\}$ such that $h \succeq f$,

$$\|[h]_p - [xS_ih]_p\| \leq 2 \cdot p(S_i) < \delta \quad \Rightarrow \quad |V([h]_p) - V([xS_ih]_p)| < \varepsilon \quad \Rightarrow$$
$$V([xS_ih]_p) > V([h]_p) - \varepsilon \geq V([f]_p) - \varepsilon = V([g]_p) \quad \Rightarrow \quad xS_ih \succ g.$$

Analogously, $f \succ xS_ih$ for all i and for all acts $h \in \mathcal{G} \Cap \{\neg S_i\}$ such that $h \preceq g$.

Theorem 3.1

Suppose that $\succeq$ is represented by expected utility:

$$U(f) = \sum_{x \in X} u(x) \cdot p(f^{-1}(x)) \qquad \text{for } f \in \mathcal{G}.$$

Prove that $\succeq$ satisfies P2($\mathcal{R}$). For all events $A \in \mathcal{R}$, acts $f, g \in \mathcal{G} \Cap \{A\}$ and outcomes x, y,

$$U(fAx) = u(x) \cdot p(\neg A) + \sum_{z \in X} u(z) \cdot p(f^{-1}(z) \cap A)$$
$$U(gAx) = u(x) \cdot p(\neg A) + \sum_{z \in X} u(z) \cdot p(g^{-1}(z) \cap A)$$
$$U(fAy) = u(y) \cdot p(\neg A) + \sum_{z \in X} u(z) \cdot p(f^{-1}(z) \cap A)$$
$$U(gAy) = u(y) \cdot p(\neg A) + \sum_{z \in X} u(z) \cdot p(g^{-1}(z) \cap A).$$

It follows that $U(fAx) \geq U(gAx)$ if and only if $U(fAy) \geq U(gAy)$, that is, $fAx \succeq gAx$ if and only if $fAy \succeq gAy$.

References

[1] M. Allais. La psychologie de l'home rationnel devant le risque: critique des postulats et axiomes de l'école Américaine. *Econometrica*, 21:503–46, 1953. Translated and reprinted in Allais and Hagen, 1979.

[2] P. Billingsley. *Probability and Measure*. Wiley, New York, third edition, 1995.

[3] B. de Finetti. La prévision: ses lois logiques, ses sources subjectives. *Annales de l'Institute Henri Poincare*, 7:1–68, 1937. Translated and reprinted in Kyburg and Smokler, 1964.

[4] D. Ellsberg. Risk, ambiguity, and the Savage axioms. *Quaterly Journal of Economics*, 75:643–669, 1961.

[5] L. Epstein. A definition of uncertainty aversion. *Review of Economic Studies*, 66:579–608, 1999.

[6] L. Epstein and J. Zhang. Subjective probabilities on subjectively unambiguous events. *Econometrica*, 69:265–306, 2001.

[7] P. C. Fishburn. *Utility Theory for Decision-Making*. Wiley, New York, 1970.

[8] P. Ghirardato and M. Marinacci. Ambiguity mase precise: A comparative foundation. *Journal of Economic Theory*, 102:251–289, 2002.

[9] I. Gilboa and D. Schmeidler. Maxmin expected utility with non-unique prior. *Journal of Mathematical Economics*, 18:141–153, 1989.

[10] S. Grant. Subjective probability without monotonicity: or how Machina's Mom may also be probabilistically sophisticated. *Econometrica*, 63:159–180, 1995.

[11] F. Gul. Savage's Theorem with a finite number of states. *Journal of Economic Theory*, 61:99–110, 1993.

[12] I. Herstein and J. Milnor. An axiomatic approach to measurable utility. *Econometrica*, 21:291–297, 1953.

[13] F. Knight. *Risk, Uncertainty and Profit*. Houghton Mifflin, Boston, 1921.

[14] C. Kraft, J. Pratt, and A. Seidenberg. Intuitive probability on finite sets. *Annals of Mathematical Statistics*, 30:408–430, 1959.

[15] D. M. Kreps. *Notes on the Theory of Choice*. Westview Press, Boulder and London, 1988.

[16] M. Machina and D. Schmeidler. A more robust definition of subjective probability. *Econometrica*, 57:571–58, 1992.

[17] K. Nehring. Capacities and probabilistic beliefs: A precarious coexistence. *Mathematical Social Sciences*, 38:197–213, 1999.

[18] I. Niiniluoto. A note on fine and tight qualitative probabilities. *The Annals of Mathematical Statistics*, 43:1581–1591, 1972.

[19] R. Sarin and P. Wakker. A simple axiomatization of non-additive expected utility. *Econometrica*, 60:1255–1272, 1992.

[20] L. J. Savage. *The Foundations of Statistics*. Dover Publications Inc., New York, second revised edition, 1972.

[21] D. Schmeidler. Subjective probability and expected utility without additivity. *Econometrica*, 57:571–585, 1989.

[22] P. Wakker. Agreeing probability measures for comparative probability structures. *The Annals of Statistics*, 9:658–662, 1981.

[23] J. Zhang. Qualitative probabilities on λ-systems. *Mathematical Social Sciences*, 38:11–20, 1999.

[24] J. Zhang. Subjective ambiguity, expected utility and Choquet expected utility. *Economic Theory*, 20:159–181, 2002.