Subjective probabilities on “small” domains

Igor Kopylov\textsuperscript{a, b, *, 1}

\textsuperscript{a}Institute for Mathematical Behavioral Sciences, University of California, Irvine, CA 92697, USA
\textsuperscript{b}Department of Economics, University of California, Irvine, CA 92697, USA

Received 16 March 2005
Available online 18 January 2006

Abstract

The Savagian choice-theoretic construction of subjective probability does not apply to preferences, like those in the Ellsberg Paradox, that reflect a distinction between risk and ambiguity. We formulate 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. Risky events can be either specified exogenously or 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 intuitive for both exogenous and behavioral specifications of risky events. This structure is weaker than an algebra or even a $\lambda$-system.

$©$ 2005 Elsevier Inc. All rights reserved.

\textit{JEL classification:} D80; D81

\textit{Keywords:} Subjective probability; Expected utility; Probabilistic sophistication; Ellsberg Paradox; Mosaic

1. Introduction

Savage [19] provides foundations for the use of \textit{subjective probability} in decision making. More precisely, he axiomatizes an \textit{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 randomizing device. One component of this representation is a probabilistic belief assigned subjectively on a universal $\sigma$-\textit{algebra} of events $\Sigma$, such as the power set of all subsets of $S$. 

\* Corresponding author at: Department of Economics, University of California, Irvine, CA 92697, USA.
\textit{E-mail address:} ikopylov@uci.edu.

\* This paper is based on a chapter of my doctoral thesis (University of Rochester, 2003) supervised by Larry Epstein. I owe special thanks to him for guidance and generous advice. I am grateful to Peter Wakker, Eddie Dekel, Itzhak Gilboa, Klaus Nehring, Mark Machina, Peter Klibanoff, Arthur Robson, and anonymous referees for helpful comments.

0022-0531/$ - see front matter © 2005 Elsevier Inc. All rights reserved.
Savage’s theory has some well-known limitations. In a paradox due to Allais [2], all events have explicit numerical probabilities, but preferences are not represented by expected utility. Motivated by the Allais paradox, Machina and Schmeidler [17] derive subjective probabilities separately from expected utility. 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 prizes), and then she ranks the induced lotteries but not necessarily via expected utility. Note that both Savage’s theory and Machina–Schmeidler’s extension derive subjective probabilities for all events in the universal \( \sigma \)-algebra \( \Sigma \).

In another paradox, due to Ellsberg [4], the decision maker is unwilling to assign probabilities to some events in \( \Sigma \). For example, consider an urn that contains balls of three possible colors—blue \( (B) \), green \( (G) \), and red \( (R) \). The decision maker is told only that the total number of balls is 90 and \( R = 30 \). Then the typical preference is to bet on the event \( \{R\} \) rather than on the event \( \{B\} \) because the frequency of \( \{R\} \) is known to be \( \frac{1}{3} \), while the frequency of \( \{B\} \) is not known precisely and lies between 0 and \( \frac{2}{3} \). Analogously, it is typical to bet on \( \{B, G\} \) rather than on \( \{R, G\} \). This preference reversal shows that the decision maker does not rely on any belief \( p \) because otherwise \( p(\{R\}) > p(\{B\}) \) but also \( p(\{B\}) + p(\{G\}) > p(\{R\}) + p(\{G\}) \). Thus, the Ellsberg Paradox demonstrates the behavioral significance of the distinction between risk, which can be represented by numerical probabilities, and ambiguity, which cannot.\(^2\)

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”. For instance, one might restrict Savage’s (or Machina–Schmeidler’s) axioms to acts measurable with respect to a \( \sigma \)-algebra \( \Sigma_* \subseteq \Sigma \) generated by an extraneous randomizing device, such as a fair coin or a roulette wheel. Then one could still obtain subjective probabilities, but only on the “small” domain \( \Sigma_* \subseteq \Sigma \).

Our main objective is to adapt Savage’s and Machina–Schmeidler’s axioms and characterize the use of subjective probability on a suitable domain \( \mathcal{R} \subseteq \Sigma \) of risky events that need not be a \( \sigma \)-algebra. Weakening the mathematical structure of the domain \( \mathcal{R} \) (which may seem a technical issue at first) is well-motivated in the literature on decision making under ambiguity. In this literature, \( \mathcal{R} \) is often understood as an exogenous class of events for which objective probabilities (frequencies) are given to the decision maker. As pointed out by Zhang [22], such \( \mathcal{R} \) is only a \( \lambda \)-system, that is, closed under complements and disjoint unions, but not necessarily under intersections. For example, consider an urn containing balls of four possible colors—blue \( (B) \), green \( (G) \), red \( (R) \), and yellow \( (Y) \)—and suppose that the decision maker is told only that \( B + G + R + Y = 100 \) and \( B + G = B + Y = 50 \). Then

\[
\mathcal{R} = \{\emptyset, \{B, G\}, \{R, Y\}, \{B, Y\}, \{G, R\}, S\}
\]

is a \( \lambda \)-system, but not an algebra.

Alternatively, risky events can be defined endogenously. In other words, one can define whether an event \( A \in \Sigma \) is subjectively risky or ambiguous in terms of the decision maker’s preference over acts. Zhang [23] and Epstein–Zhang [5] formulate two leading definitions of subjectively risky events.\(^3\) However, neither of these definitions implies that the domain of such events, \( \mathcal{R}_Z \subseteq \Sigma \)

---

2 This distinction is often attributed to Knight [14]. He refers to uncertainty rather than ambiguity. We adopt Ellsberg’s terminology, in which uncertainty is comprehensive and includes both risk and ambiguity.

3 These authors use the term unambiguous rather than risky. The difference between the two definitions in Zhang [23] and Epstein–Zhang [5] reflects the distinction between expected utility and probabilistic sophistication. Ghirardato [8] reviews other possible approaches.
or $R_{EZ} \subseteq \Sigma$, respectively, is a $\sigma$-algebra or even a $\lambda$-system. In general, both $R_Z$ and $R_{EZ}$ have a weaker structure that is called mosaic and is defined later.

In a more abstract vein, assuming countable unions or intersections can be a problem, even for the universal class of events $\Sigma$. Indeed, the decision maker must conceive of all events in $\Sigma$ prior to ranking all $\Sigma$-measurable acts. However, a $\sigma$-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, let $\Sigma$ be the algebra of cylinders—events that rely on finitely many coin tosses—and let $\sigma(\Sigma)$ be the smallest $\sigma$-algebra that contains $\Sigma$. It is well-known that many events in $\sigma(\Sigma)$ are not readily obtained via combinations of cylinders. More precisely, for any countable ordinal $\gamma$, there are sets in $\sigma(\Sigma)$ that cannot be arrived at from $\Sigma$ by a $\gamma$-sequence of set-theoretic operations, each operation being a complement, a countable union or a countable intersection [3, pp. 31–32]. In this sense, the class $\sigma(\Sigma)$ is substantially more complex than $\Sigma$. Probability theorists avoid this complexity and do not construct measures directly on a $\sigma$-algebra; rather, the typical procedure is to describe the measure explicitly on “simple” events and then to apply a measure extension theorem (recall the construction of the Lebesgue measure on the Borel $\sigma$-algebra). Similarly, the decision maker who conceives of all events in an algebra $\Sigma$ may be unable or unwilling to conceive of some sets in $\sigma(\Sigma)$. Therefore, reliance on all events in a $\sigma$-algebra may be problematic for a normative theory of subjective probability. Our model alleviates this problematic aspect by taking both $\Sigma$ and $R$ to be finitely additive domains (an algebra and a mosaic, respectively).

We formulate two main representation results—one for expected utility (Theorem 3.1), the other for probabilistic sophistication (Theorem 4.1)—that derive subjective probabilities on a given mosaic $R \subseteq \Sigma$ from preference over $R$-measurable acts. In particular, one can apply these results to $R = R_Z$ and to $R = R_{EZ}$, respectively, and thus, obtain fully subjective models of expected utility and probabilistic sophistication. Both models are constructive and derive the values and the domain of the subjective probability measure from preference. However, the two models use different definitions of risk and obtain different representations for preference. The former delivers expected utility over $R_Z$-measurable acts, while the latter delivers probabilistic sophistication over $R_{EZ}$-measurable acts.

The noted Theorems 3.1 and 4.1 show that the content of Savage’s and Machina–Schmeidler’s axioms can be preserved almost entirely when restricted to a “small” domain of risky events. An essential modification is required only for Savage’s P6; all the other axioms are restricted in a straightforward fashion. (Other extensions of Savage’s result, such as Gul [12] and Abdellaoui and Wakker [1], adopt arguably less transparent axioms for preference and additional structure for outcomes of acts.) However, the construction of subjective probability, which is a key part of Savage’s proof and is used also by Machina–Schmeidler, does not carry through to mosaics—or even to algebras—and needs to be replaced by an alternative strategy. We propose a compact formula that computes subjective probabilities for all risky events from the ranking of bets on such events. This formula is intuitive and reveals a general connection between the betting behavior and the underlying numerical probability.

Our results allow a lot of flexibility for applications. In particular, choice outside the domain of risky acts is not restricted by any conditions, or a fortiori by any parametric model, such as Choquet expected utility [20] or the multiple priors model [10]. Also, there is freedom in the

---

4 Savage acknowledges that the assumption that $\Sigma$ is a $\sigma$-algebra is essential for his proof, but notes that it is peculiar that “one should use countable unions of events in order to derive a finitely additive probability measure.” Machina–Schmeidler and Epstein–Zhang also take $\Sigma$ to be a $\sigma$-algebra.
identity of the mosaic $\mathcal{R}$ where subjective probabilities are derived. We focus on applications where $\mathcal{R} = \mathcal{R}_{Z}$ and $\mathcal{R} = \mathcal{R}_{EZ}$, but in general, $\mathcal{R}$ need not equal either of these domains.

This paper proceeds as follows. Next we present our formal framework and define the notion of a mosaic. In Sections 3 and 4, we formulate the two main representation results—one for expected utility (Theorem 3.1), the other for probabilistic sophistication (Theorem 4.1). In Section 5, we apply these results to $\mathcal{R} = \mathcal{R}_{Z}$ and to $\mathcal{R} = \mathcal{R}_{EZ}$, respectively. Proofs are sketched in Section 6 and are presented in detail in the Appendix.

2. Preliminaries

We use a version of Savage’s framework. Given are

- a set $X$ of deterministic outcomes,
- a set $S$ of states of the world,
- a class $\Sigma$ of subsets of $S$ that are called events.

No formal structure is imposed on $X$ and $S$. The class $\Sigma$ is assumed to be an algebra rather than a $\sigma$-algebra.

Call a function $f : S \rightarrow X$ an act if it is $\Sigma$-measurable and has finite range. Identify outcomes $x \in X$ with the corresponding constant acts. Let $\mathcal{F}$ be the set of all acts. Interpret each act $f \in \mathcal{F}$ as a physical action that yields the outcome $f(s)$ contingent on the state $s \in S$. The decision maker’s preference over such actions is given as a binary relation $\succeq$ on $\mathcal{F}$. 5

In addition to the primitives $S$, $X$, $\Sigma$, and $\succeq$, fix a class $\mathcal{R} \subseteq \Sigma$ of risky events; call all other events in $\Sigma$ ambiguous. Say that $S = \cup_{i=1}^{m} S_i$ is a grand partition if the events $S_1, \ldots, S_m$ are disjoint and risky. Assume that $\mathcal{R}$ is a mosaic, that is, satisfies the following conditions:

\begin{itemize}
  \item[(*)] $S \in \mathcal{R}$;
  \item[(**)] $A \in \mathcal{R} \Rightarrow A^c \in \mathcal{R}$; and
  \item[(\mu)] $S = \cup_{i=1}^{m} S_i$ is a grand partition $\Rightarrow S_i \cup S_j \in \mathcal{R}$ for all $i, j = 1 \ldots m$.
\end{itemize}

Equivalently, (\mu) requires that any partition coarser than a grand partition is also a grand partition. Note that a mosaic need not be closed under intersections, or unions, or even disjoint unions. For example, if $S = \{R, G, B\}$, then the mosaic $\{\emptyset, \{R\}, \{G\}, \{G, B\}, \{R, B\}, S\}$ does not contain the disjoint union of its elements $\{R\}$ and $\{G\}$. The mosaic structure—unlike the more restrictive $\lambda$-system—is implied by the behavioral definitions of risky events due to Zhang [23] and Epstein–Zhang [5] (see Section 5 for further discussion).

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

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

$$\mathcal{R} \cap \mathcal{E} = \{A \in \mathcal{R} : A \cap E \in \mathcal{R} \text{ for all } E \in \mathcal{E}\}.$$
and let \( G \cap E \) be the set of \((\mathcal{R} \cap E)\)-measurable acts. In other words, \( G \cap E \) consists of risky acts that remain \( \mathcal{R} \)-measurable when restricted to any event \( E \) from the given collection \( \mathcal{E} \). Note that if \( \mathcal{R} \) is an algebra, then the above notation is redundant because \( \mathcal{R} = \mathcal{R} \cap E \) and \( G = G \cap E \).

For any acts \( f, h \in \mathcal{F} \) and any event \( A \in \Sigma \), define a composite act \( fAh \) that yields \( f(s) \) if \( s \in A \) and \( h(s) \) if \( s \in A^c \). Note that \( fAh \) may be ambiguous even when the event \( A \) and the acts \( f \) and \( h \) are risky. For example, take outcomes \( x \neq y \) and risky events \( A, B \in \mathcal{R} \) such that \( A \cap B \notin \mathcal{R} \); then \( f = xBy \) and \( h = y \) are risky acts, but \( fAh = x(A \cap B)y \) is not. One way to guarantee that \( fAh \in G \) is to require that \( A \in \mathcal{R} \), \( f \in G \cap \{A\} \) and \( h \in G \cap \{A^c\} \).

\begin{align}
A & \in \mathcal{R}, \quad f \in G \cap \{A\} \quad \text{and} \quad h \in G \cap \{A^c\}. \tag{1}
\end{align}

Under these conditions, the events \( f^{-1}(x) \cap A \) and \( h^{-1}(x) \cap A^c \) form a grand partition as \( x \) varies over the finite range of the acts \( f \) and \( h \); by (\( \mu \)),

\[(f^{-1}(x) \cap A) \cup (h^{-1}(x) \cap A^c) = (fAh)^{-1}(x)\]

is a risky event for each \( x \in X \), and the act \( fAh \) is risky.

3. Subjective expected utility

To characterize subjective expected utility, Savage [19] (see also [7,16]) imposes a list of axioms P1–P6 on the preference \( \succeq \) over the set of all acts \( \mathcal{F} \). However, the Ellsberg-type behavior violates one of these axioms, P2.

**Axiom P2 (Sure-Thing Principle).** For all events \( A \in \Sigma \), acts \( f, g \in \mathcal{F} \), and outcomes \( x, y \in X \),

\[ fAx \succeq gAx \Rightarrow fAy \succeq gAy. \tag{3} \]

P2 states that preference is separable across mutually exclusive events and can be conditioned on any event \( A \) independently of the outcome obtained on the complement of \( A \).

To avoid Ellsberg-type preference reversals, rewrite P2 so that all the composite acts \( fAx \), \( fAy \), \( gAx \), and \( gAy \) in the invariance (3) are risky, that is, belong to the class \( G \) of \( \mathcal{R} \)-measurable acts. By (1), it is sufficient to require that \( A \in \mathcal{R} \) and \( f, g \in G \cap \{A\} \).

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

\[ fAx \succeq gAx \Rightarrow fAy \succeq gAy. \tag{4} \]

P2(\( \mathcal{R} \)) asserts separability of preference but only across risky events. More precisely, P2(\( \mathcal{R} \)) requires that preference over risky acts that have \( \mathcal{R} \)-measurable restrictions to a risky event \( A \) can be conditioned on \( A \) independently of the outcome obtained on the complement of \( A \). Note that P2 implies P2(\( \mathcal{R} \)), and the two axioms are equivalent if \( \mathcal{R} = \Sigma \).

---

6 Savage uses an equivalent formulation: for all \( A \in \Sigma \) and \( f, g, h, h' \in \mathcal{F} \),

\[ fAh \succeq gAh \Rightarrow fAh' \succeq gAh'. \tag{2} \]

The invariance (3) is a special case of (2) with constant \( h \) and \( h' \). Conversely, (2) follows from (3) by induction over the number of outcomes in the range of \( h \) and \( h' \).
Even though the Ellsberg-type behavior may comply with Savage’s axioms other than P2, it does not seem intuitive to restrain choices among ambiguous acts in order to derive expected utility for risky acts. Thus, we rewrite the rest of the list P1–P6 so that the rewritten axioms P1(\(\mathcal{R}\))–P6(\(\mathcal{R}\)) involve only risky acts and events.

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

If \(\mathcal{R} = \Sigma\), then P1(\(\mathcal{R}\)) is equivalent to Savage’s P1. In general, P1(\(\mathcal{R}\)) is weaker than P1 and allows an incomplete or intransitive choice among ambiguous acts.

**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 \(h \in \mathcal{G} \cap \{A^c\}\):

(i) \(x \succeq y \iff xAh \succeq yAh\);
(ii) \(xAh \sim yAh\).

P3(\(\mathcal{R}\)) is a version of state-independence for the ranking of deterministic outcomes. This axiom classifies each risky event as either non-null or null so that the ranking of \(X\) remains unchanged when conditioned on any non-null risky event and becomes degenerate when conditioned on any null risky event. This axiom uses only risky acts and does not classify ambiguous events. P3(\(\mathcal{R}\)) follows from Savage’s P3; the two axioms are equivalent if \(\mathcal{R} = \Sigma\).

**Axiom P4(\(\mathcal{R}\)).** For all risky events \(A, B \in \mathcal{R}\), for all outcomes \(x \succ x'\) and \(z \succ z'\), if \(xAx' \succeq xBx'\), then \(zAz' \succeq zBz'\).

P4(\(\mathcal{R}\)) requires that if \(A\) and \(B\) are risky events, then the preference to bet \(A\) rather than on \(B\) should reflect exclusively the subjective relative likelihoods of \(A\) and \(B\) and be independent of the stakes involved in such bets. P4(\(\mathcal{R}\)) follows from Savage’s P4; the two axioms are equivalent if \(\mathcal{R} = \Sigma\).

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

P5(\(\mathcal{R}\)) is identical to Savage’s P5.

Before stating P6(\(\mathcal{R}\)), we briefly discuss Savage’s

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

P6 requires that for any outcome \(x\) and for any acts \(f \succ g\), the universal event \(S\) can be partitioned into small events \(S_i\) so that the strict preference \(f \succ g\) is not reversed when the outcomes \(f(s), \text{ or alternatively } g(s)\), are replaced by \(x\) for all \(s \in S_i\). To motivate P6, consider a coin such that all finite sequences \((s_1, \ldots, s_n)\) of its heads and tails are events in \(\Sigma\). It is intuitive that for a sufficiently large \(n\), each sequence \((s_1, \ldots, s_n)\) is a small event. Note that P6 is a condition on the preference, but also on the other primitives: for instance, P6 implies that both \(S\) and \(\Sigma\) are infinite.

Rewrite P6 in terms of risky events and risky acts as follows.
**Axiom P6(\(\mathcal{R}\)).** For any outcome \(x\), for any finite collection of risky events \(\mathcal{E} \subseteq \mathcal{R}\), and for any \(\mathcal{E}\)-measurable acts \(f > g\), there exists a grand partition \(\{S_1, \ldots, S_m\} \subseteq \mathcal{R} \cap \mathcal{E}\) such that for all \(i = 1 \ldots m\), \(xS_if > g\) and \(f > xS_ig\).\(^7\)

When \(\mathcal{R} = \Sigma\), P6(\(\mathcal{R}\)) is equivalent to P6 because \(\Sigma \cap \mathcal{E} = \Sigma\), but in the general case when \(\mathcal{R} \subset \Sigma\), P6(\(\mathcal{R}\)) does not follow from P6 because the small events in P6(\(\mathcal{R}\)), unlike those in P6, are restricted to a particular subclass \(\mathcal{R} \cap \mathcal{E} \subset \mathcal{R} \subset \Sigma\). To motivate this restriction, fix a finite collection \(\mathcal{E} \subset \mathcal{R}\) and consider a coin such that all finite sequences \((s_1, \ldots, s_n)\) of its heads and tails are risky events in \(\mathcal{R}\) and are independent of all events in \(\mathcal{E}\). In other words, suppose that the decision maker’s perception of the likelihoods of the events in the finite collection \(\mathcal{E}\) is not affected by the results of the coin tosses (this notion of independence is used only to motivate P6(\(\mathcal{R}\)) and is not a formal part of our model). It is intuitive that for a sufficiently large \(n\), every event \((s_1, \ldots, s_n)\) is small, and for each \(E \in \mathcal{E}\), the event \((s_1, \ldots, s_n) \cap E\) is risky because it can be assigned probability \(p(s_1, \ldots, s_n) \cdot p(E)\).

Call a function \(p : \mathcal{R} \to [0, 1]\) a probability measure if \(\sum_{i=1}^m p(S_i) = 1\) for all grand partitions \(S = \bigcup_{i=1}^m S_i\). 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 grand partition \(\{S_1, \ldots, S_m\} \subseteq \mathcal{R} \cap \mathcal{E}\) such that \(p(S_i) < \varepsilon\) for all \(i = 1 \ldots m\).

Our first main result is

**Theorem 3.1.** Let \(\Sigma\) be an algebra, and let \(\mathcal{R} \subseteq \Sigma\) be a mosaic. Then the following two statements are equivalent:

1. \(\geq\) satisfies P1(\(\mathcal{R}\)), P2(\(\mathcal{R}\)), P3(\(\mathcal{R}\)), P4(\(\mathcal{R}\)), P5(\(\mathcal{R}\)), P6(\(\mathcal{R}\));
2. \(\geq\) is represented on the set \(\mathcal{G}\) by expected utility

\[
U(f) = \sum_{x \in X} u(x) \cdot p(f^{-1}(x)) \quad \text{for } f \in \mathcal{G},
\]

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 (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, Theorem 3.1 provides 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 \(\Sigma\) is a \(\sigma\)-algebra and \(\mathcal{R} = \Sigma\). In this case, the two results use equivalent axioms and deliver identical representations (this identity follows from uniqueness). Moreover, Theorem 3.1 implies that even if the algebra \(\Sigma\) is not countably additive, Savage’s axioms taken “as is” are still sufficient for an expected utility

---

\(^7\)All the acts \(xS_if\) and \(xS_ig\) are risky because \(f, g \in \mathcal{G} \cap [S_i]\). To show that \(f \in \mathcal{G} \cap [S_i]\), consider the grand partition of \(S\) into risky events \(f^{-1}(z) \cap S_j\), where \(z \in X\) and \(1 \leq j \leq m\); here, \(f^{-1}(z) \cap S_j \in \mathcal{R}\) because \(f^{-1}(z) \in \mathcal{E}\) and \(S_j \in \mathcal{R} \cap \mathcal{E}\). By (\(\mu\)), \(f^{-1}(y) \cap S_i = \bigcup_{j \neq i} (f^{-1}(y) \cap S_j) \in \mathcal{R}\) for all \(y \in X\).
representation on the set $\mathcal{F}$ of all acts. In other words, the assumption that $\Sigma$ is countably additive 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 (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 ambiguity aversion or by any parametric utility representation, such as Choquet expected utility [20] or maxmin expected utility [10].

The proof of Theorem 3.1—most importantly, the construction of the probability measure $p$—requires a new approach that is sketched in Section 6.

4. Probabilistic sophistication

In order to accommodate Allais-type behavior, Machina and Schmeidler [17] 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 events $A, B, C \in \Sigma$, for all outcomes $x > x'$ and $z > z'$, and for all acts $h, h' \in \mathcal{F}$, 9

\[(xAx')Ch \succeq (xBx')Ch \Rightarrow (zA'z')Ch' \succeq (zBz')Ch'. \tag{6}\]

P4 is a special case for $C = S$. In addition to this special case, P4* requires that the preference to bet on $A \cap C$ rather than on $B \cap C$ is invariant of the outcomes that are obtained when $C$ does not occur. This invariance reflects exclusively the subjective relative likelihoods of the events $A \cap C$ and $B \cap C$, and does not imply that the decision maker’s risk preference is separable (i.e. satisfies Independence over induced lotteries).

The Ellsberg Paradox motivates rewriting P4* for preference over risky acts.

Axiom P4*(\mathcal{R}). For all risky events $A, B \in \mathcal{R}$ and $C \in \mathcal{R} \cap \{A, A^c, B, B^c\}$, for all outcomes $x > x'$ and $z > z'$, and for all acts $h, h' \in \mathcal{G} \cap \{C\}$,

\[(xAx')Ch \succeq (xBx')Ch \Rightarrow (zA'z')Ch' \succeq (zBz')Ch'. \tag{7}\]

8 If $\Sigma$ is a $\sigma$-algebra, then a probability measure $p : \Sigma \rightarrow [0, 1]$ is finely ranged if an only if it is convex-ranged, that is for each event $B \in \Sigma$, the range $\{p(A) : A \in \Sigma, A \subseteq B\}$ equals the interval $[0, p(B)]$. This equivalence does not hold if $\Sigma$ is an algebra, in which case $p$ is finely ranged if and only if for each event $B \in \Sigma$, $\{p(A) : A \in \Sigma, A \subseteq B\}$ is dense in $[0, p(B)]$.

9 Machina–Schmeidler assume that $A$ and $B$ are disjoint and $C = A \cup B$. This assumption is wlog because all the composite acts in (6) are unchanged when $A$ and $B$ are replaced by the disjoint events $A' = (A \setminus B) \cap C$ and $B' = (B \setminus A) \cap C$, $C$ by the union $C' = A' \cup B'$, $h$ by $g = xCh$, and $h'$ by $g' = x'Ch'$.
Theorem 4.1. Let $A \cap C$, $B \cap C$, and $C$ be the set of risky acts and events. Accordingly, $P_4$ is separable over lotteries induced by risky acts, or that her preference to bet is separable across risky acts. In this representation, $P_4$ is weakly dominates $P_4'$ if $R = \Sigma$).

To state our second representation result, we need a few preliminaries. Let $L$ be the set of all risky acts—probability distributions $l : X \rightarrow [0, 1]$ with finite support. Fix a probability measure $p : R \rightarrow [0, 1]$. Every risky act $f \in G$ induces a lottery, written $[f]_p$, such that $[f]_p(x) = p(f^{-1}(x))$ for all $x \in X$. Let

$$L_p = \{l \in L : l = [f]_p \text{ for some } f \in G\}$$

be the set of lotteries induced by risky acts. Given $Y \subseteq X$, let $l(Y) = \sum_{x \in Y} l(x)$ for all $l \in L$, and let $L_p(Y) = \{l \in L_p : l(Y) = 1\}$.

Fix a risk preference relation $\succeq_1$ on $L_p$. Let $\succeq$ be the restriction of $\succeq_1$ to the set $X$ of degenerate lotteries, and let $Y_x = \{y \in X : y \succeq_x x\}$ be the upper contour sets of the ranking $\succeq_x$. Say that a lottery $l$ weakly dominates $l'$, written $l \succeq l'$, if $l(Y_X) \supseteq l'(Y_X)$ for all $x \in X$. Say that $l$ strictly dominates $l'$, written $l \succ l'$, if $l \succeq l'$ and $l(Y_X) > l'(Y_X)$ for some $x \in X$. Say that the relation $\succeq_1$ is (i) strictly monotonic if for all $l, l' \in L_p, l \succeq_1 l'$ implies $l \succeq l'$, and $l \succ l'$ implies $l \succ l'$; (ii) continuous if for all finite $Y \subseteq X$ and $l \in L_p(Y)$, the upper and the lower contour sets $[l \in L_p(Y) : l \succeq_1 l]$ and $[l' \in L_p(Y) : l' \succeq_1 l]$ are closed in the finite-dimensional set $L_p(Y)$.

**Theorem 4.1.** Let $\Sigma$ be an algebra, and let $R \subseteq \Sigma$ be a mosaic. Then the following two statements are equivalent:

(i) $\succeq$ satisfies $P_1(R), P_3(R), P_4^*(R), P_5(R), P_6(R)$;

(ii) $\succeq$ on the set $G$ is represented by

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

where $p : R \rightarrow [0, 1]$ is a finely ranged probability measure, and the binary relation $\succeq_1$ on $L_p$ 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.

Similar to Theorem 3.1, Theorem 4.1 models choice only among risky acts, and derives subjective probabilities only for risky events. In representation (8), the probabilities are used to translate risky acts $f \in G$ into lotteries $[f]_p \in L_p$, and the induced lotteries are ranked via a continuous and strictly monotonic weak order rather than via expected utility. Therefore, Theorem 4.1 provides foundations for probabilistic sophistication on risky acts.

The axioms in Theorem 4.1 are parallel to Machina–Schmeidler’s counterparts but apply only to risky acts and events. Most importantly, $P_4^*(R)$ does not restrict the preference to bet on ambiguous events. Accordingly, only risky events $A \in R$ are assigned subjective probabilities, only risky acts $f \in G$ are translated into lotteries, and only lotteries $[f]_p \in L_p$ that are induced by risky acts are ranked by the risk preference $\succeq_1$. 


Note that representation (8) retains the risk preference as a binary relation \( \succeq_1 \) on the subdomain \( \mathcal{L}_p \subseteq \mathcal{L} \). Thus, (8) constitutes a notion of probabilistic sophistication different from, albeit closely related to the one used by Machina–Schmeidler, who specify a suitable utility representation \( V \) for the risk preference. It is an open question whether conditions of Theorem 4.1 are sufficient for \( \succeq_1 \) to have any utility representation. In particular, one cannot use Machina–Schmeidler’s construction of \( V \) (see the example in Section 6.2). Nevertheless, it seems appropriate to describe the use of probabilities by the strictly monotonic risk preference \( \succeq_1 \) alone.

If \( \Sigma \) is a \( \sigma \)-algebra and \( \mathcal{R} = \Sigma \), then \( p \) is convex-ranged and \( \mathcal{L}_p = \mathcal{L} \). In this case, Theorem 4.1 becomes a version of the Machina–Schmeidler Theorem: both results use equivalent axioms and deliver the same subjective probability \( p \) and the same risk preference \( \succeq_1 \) on the entire \( \mathcal{L} \). It is straightforward to show that such \( \succeq_1 \) has a continuous utility representation \( V \) (this is the last step of Machina–Schmeidler’s proof). More generally, Theorem 4.1 shows that even if the algebra \( \Sigma \) is not countably additive, axioms P1, P3, P4*, P5, P6 taken "as is" characterize a version of probabilistic sophistication over \( \Sigma \)-measurable acts—a version that does not specify a utility representation for the risk preference.

5. Subjective definitions of risk

Similar to the results of Savage and Machina–Schmeidler, Theorems 3.1 and 4.1 are formulated for a given class of events. Accordingly, subjective probabilities in representations (5) and (8) are derived on a mosaic \( \mathcal{R} \) which is exogenous to the model. The exogenous specification of risky events may be problematic because decision makers may disagree about the identity of events to which they assign probabilities. In order to model the subjective nature of risk, Zhang [23] defines risky events in terms of the preference \( \succeq \). His definition is motivated by the Sure-Thing Principle rather than by any particular attitude towards ambiguity, such as ambiguity aversion.10

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

\[
x Ef \succeq x Eg \implies y Ef \succeq y Eg
\]

and

\[
f Ex \succeq g Ex \implies f Ey \succeq g Ey. \quad (9)
\]

Otherwise, call \( E \) subjectively ambiguous.

This definition takes complementary events \( E \) and \( E^c \) to be subjectively risky if the preference \( \succeq \) is separable across \( E \) and \( E^c \) and can be conditioned on either of these events independently of the outcome obtained on the other. The use of acts that are constant on \( E \) (or on \( E^c \)) reflects the intuition that subsets of a subjectively risky event may be subjectively ambiguous.

Zhang’s definition may be problematic because it attributes all violations of Savage’s model to ambiguity. For example, a probabilistically sophisticated decision maker may assign probabilities to all events in \( \Sigma \), but violate conditions (9) for some \( E \in \Sigma \) because of a non-linear risk preference. To address this concern, Epstein–Zhang [5] formulate a less stringent definition of subjectively
risky events that is motivated by Strong Comparative Probability rather than by the Sure-Thing Principle.

Call acts \( f, g \in \mathcal{F} \) joint bets if there exist events \( A, B, C \in \Sigma \), outcomes \( z > z' \), and an act \( h \in \mathcal{F} \) such that \( f = (zA'z')Ch \) and \( g = (zBz')Ch \).

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

\[
x Ef \succeq x Eg \Rightarrow y Ef \succeq y Eg
\]

and

\[
f Ex \succeq g Ex \Rightarrow f Ey \succeq g Ey.
\]

(10)

Otherwise, \( E \) is called subjectively ambiguous.

This definition takes complementary events \( E \) and \( E^c \) to be subjectively risky if the preference to bet on subsets of \( E^c \) (or on subsets of \( E \)) is invariant of the outcome obtained on \( E \) (or on \( E^c \), respectively). The focus on comparisons between joint bets that are constant on \( E \) (or on \( E^c \)), reflects the intuition that (i) the risk preference may be non-linear, and (ii) subsets of a subjectively risky event may be subjectively ambiguous.

Denote by \( \mathcal{R}_Z \) and \( \mathcal{R}_{EZ} \) the classes of subjectively risky events defined by Zhang and by Epstein–Zhang, respectively. Both \( \mathcal{R}_Z \) and \( \mathcal{R}_{EZ} \) are uniquely derived from the preference \( \succeq \). Let \( \mathcal{G}_Z \) and \( \mathcal{G}_{EZ} \) be the associated sets of risky acts. By definition, \( \mathcal{R}_Z \subseteq \mathcal{R}_{EZ} \subseteq \Sigma \) and \( \mathcal{G}_Z \subseteq \mathcal{G}_{EZ} \subseteq \mathcal{F} \), but in the special case when \( X \) has only two elements \( x > x' \), any two acts \( f, g \in \mathcal{F} \) are joint bets and hence, \( \mathcal{R}_Z = \mathcal{R}_{EZ} \) and \( \mathcal{G}_Z = \mathcal{G}_{EZ} \).

Next we apply Theorems 3.1 and 4.1 to the subjective domains \( \mathcal{R}_Z \) and \( \mathcal{R}_{EZ} \), respectively.

**Theorem 5.1.** Let \( \mathcal{F} \) be an algebra, and let \( \succeq \) be a reflexive preference on \( \mathcal{F} \). Then the following statements are true.

I. The classes \( \mathcal{R}_Z \) and \( \mathcal{R}_{EZ} \) are mosaics.

II. The preference \( \succeq \) satisfies \( P1(\mathcal{R}_Z) \), \( P3(\mathcal{R}_Z) \), \( P4(\mathcal{R}_Z) \), \( P5(\mathcal{R}_Z) \), \( P6(\mathcal{R}_Z) \) if and only if \( \succeq \) is represented on the set \( \mathcal{G}_Z \) by expected utility

\[
U(f) = \sum_{x \in X} u(x) \cdot p(f^{-1}(x)) \quad \text{for } f \in \mathcal{G}_Z,
\]

(11)

where \( u : X \rightarrow \mathbb{R} \) is a non-constant utility index, and \( p : \mathcal{R}_Z \rightarrow [0, 1] \) is a finely ranged probability measure. In this representation, \( u \) is unique up to a positive linear transformation, and the probability measure \( p \) is unique.

III. The preference \( \succeq \) satisfies \( P1(\mathcal{R}_{EZ}) \), \( P3(\mathcal{R}_{EZ}) \), \( P4(\mathcal{R}_{EZ}) \), \( P5(\mathcal{R}_{EZ}) \), and \( P6(\mathcal{R}_{EZ}) \) if and only if \( \succeq \) is represented on the set \( \mathcal{G}_{EZ} \) by

\[
f \succeq g \Leftrightarrow [f]_p \succeq [g]_p \quad \text{for } f, g \in \mathcal{G}_{EZ},
\]

(12)

---

11 If \( f \) and \( g \) are joint bets, then both rankings \( xEf \succeq xEg \) and \( yEf \succeq yEg \) reveal a preference to bet on \( A \cap C \cap E^c \subseteq E^c \) rather than on \( B \cap C \cap E^c \subseteq E^c \). Similarly, the rankings \( fEx \succeq gEx \) and \( fEy \succeq gEy \) reveal a preference to bet on \( A \cap C \cap E \subseteq E \) rather than on \( B \cap C \cap E \subseteq E \).
where \( p : \mathcal{R}_{\text{EZ}} \to [0, 1] \) 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.

Note that all components of representations (11) and (12), including the domains \( \mathcal{R}_Z, \mathcal{R}_{\text{EZ}} \) and the values \( p(\cdot) \) of subjective probabilities, are derived from preference. Therefore, Theorem 5.1 provides fully subjective theories of expected utility and probabilistic sophistication.

If \( \succeq \) satisfies P2, then \( \mathcal{R}_Z = \Sigma \) and representation (11) holds on the set \( \mathcal{F} \) of all acts. More generally, when P2 does not hold, preference is still separable across events in \( \mathcal{R}_Z \), and the axioms P1(\( \mathcal{R}_Z \)), P3(\( \mathcal{R}_Z \)), P4(\( \mathcal{R}_Z \)), P5(\( \mathcal{R}_Z \)), P6(\( \mathcal{R}_Z \)) provide foundations for the use of subjective probability on \( \mathcal{R}_Z \) as a component of an expected utility representation on \( \mathcal{G}_Z \).

Alternatively, if \( \succeq \) satisfies P4\( ^* \), then \( \mathcal{R}_{\text{EZ}} = \Sigma \), and representation (12) holds on the set \( \mathcal{F} \) of all acts. If P4\( ^* \) does not hold, then betting preference is separable across events in \( \mathcal{R}_{\text{EZ}} \), and the axioms P1(\( \mathcal{R}_{\text{EZ}} \)), P3(\( \mathcal{R}_{\text{EZ}} \)), P4(\( \mathcal{R}_{\text{EZ}} \)), P5(\( \mathcal{R}_{\text{EZ}} \)), P6(\( \mathcal{R}_{\text{EZ}} \)) provide foundations for probabilistic sophistication on \( \mathcal{G}_{\text{EZ}} \).

In principle, one can apply Theorems 3.1 and 4.1 on domains other than \( \mathcal{R}_Z \) and \( \mathcal{R}_{\text{EZ}} \). The selection of \( \mathcal{R}_Z \) or \( \mathcal{R}_{\text{EZ}} \) has two major benefits. First, this selection is done via explicit and arguably intuitive behavioral definitions. Second, fewer axioms need to be imposed on the preference over \( \mathcal{G}_Z \) and \( \mathcal{G}_{\text{EZ}} \) because the definitions of these domains imply P2(\( \mathcal{R}_Z \)) and P4\( ^* \)(\( \mathcal{R}_{\text{EZ}} \)), respectively. 12

To relate representation (12) to Epstein–Zhang’s main result, consider a special case of mosaics.

Call a mosaic \( \mathcal{R} \) a (finitely additive) \( \lambda \)-system if it is closed under arbitrary disjoint unions rather than only under unions of elements of a grand partition. The class \( \mathcal{R}_{\text{EZ}} \) (or \( \mathcal{R}_Z \)) need not be a \( \lambda \)-system. For example, take \( S = \{R, G, B\} \), \( X = \{x, x'\} \), and the following ranking:

\[
x \succ x\{R, G\}x' \succ x\{R, B\}x' \succ x\{G, B\}x' \succ x\{G\}x' \succ x\{R\}x' \succ x\{B\}x' \succ x'.
\]

Then the events \( \{R\} \) and \( \{G\} \) are subjectively risky, but the event \( \{B\} \) and hence, its complement \( \{R\} \cup \{G\} \) are subjectively ambiguous. It follows that the class \( \mathcal{R}_{\text{EZ}} \) is not a \( \lambda \)-system.

It is arguably intuitive that the decision maker, who assigns probabilities \( p(A) \) and \( p(B) \) to disjoint events \( A \) and \( B \), should also assign the sum \( p(A) + p(B) \) to the union \( A \cup B \). Motivated by this intuition, Epstein–Zhang impose additional conditions on preference that guarantee that \( \mathcal{R}_{\text{EZ}} \) is a \( \lambda \)-system. 13 It is, however, an open question whether the \( \lambda \)-system structure for risky events has a simple behavioral foundation in Savage’s general framework.

Theorem 5.1 assumes a weaker structure for the class of risky events, but also differs substantially from Epstein–Zhang in the axioms adopted. Roughly, it dispenses with two of their arguably least appealing conditions—Monotone Continuity and Strong Partition Neutrality—that are not implied by the definition of \( \mathcal{R}_{\text{EZ}} \). Both axioms are even less appealing if \( \Sigma \) is finitely additive, because the former uses countable intersections, and the latter relies on equipartitions that need not exist (see the example in Section 6.1).

\[\text{12} \text{However, the richness conditions } P6(\mathcal{R}_Z) \text{ and } P6(\mathcal{R}_{\text{EZ}}) \text{ do not follow from the definitions of } \mathcal{R}_Z \text{ and } \mathcal{R}_{\text{EZ}} \text{ and are violated in some settings. This concern can be addressed by strengthening the behavioral definitions of risky events, which is a current research objective.}\]

\[\text{13} \text{In fact, their conditions are too weak and their theorem is not valid as stated (we omit a formal counterexample). To correct Epstein–Zhang’s theorem, strengthen their Axiom 4 so that, in their notation, } A_n \text{ and } B_n \text{ can vary over all of } \Sigma \text{ and not only over } A.\]
6. 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 on the construction of subjective probabilities for risky events.

6.1. Construction of subjective probability

Analogously to de Finetti [6] and Savage [19], 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

\[
A \succeq_0 B \iff xAx' \succeq xBx' \quad \text{for all outcomes } x > x',
\]

and is interpreted as \( A \) being subjectively at least as likely as \( B \). We propose the following construction for the finely ranged probability measure \( p \):

Fix an arbitrary risky event \( A \in \mathcal{R} \). Say that a grand partition \( S = \bigcup_{i=1}^{m} S_i \) is finer than \( A \) if \( A \succeq_0 S_i \) for all \( i = 1 \ldots m \). Among all grand partitions finer than \( A \), take one that has a minimal number of elements. Let \( n(A) \) be this minimal number; let \( n(A) = +\infty \) if there exists no grand partition finer than \( A \). The subjective probability \( p(A) \) can be constructed via the following formula:

\[
p(A) = \sup \left\{ \sum_{i=1}^{n} \frac{1}{n(A_i)} \right\},
\]

(13)

where the least upper bound is taken over all partitions of \( A \) into risky events \( A_1, \ldots, A_n \).

To motivate formula (13), suppose that a finely ranged probability measure \( p \) does represent \( \succeq_0 \). Then \( \frac{1}{n(A)} < p(A) \leq \frac{1}{n(A)-1} \) for all \( A \in A \).\(^{14}\) Fix a risky event \( A \) and vary both the size and the elements of a risky partition \( A = \bigcup_{i=1}^{m} A_i \). The sums \( \sum_{i=1}^{n} \frac{1}{n(A_i)} \) are lower bounds for \( \sum_{i=1}^{n} p(A_i) = p(A) \). Moreover, these sums become arbitrarily close to \( p(A) \) for sufficiently fine risky partitions because the ratios \( \frac{1}{n(A_i)} / p(A_i) \) approach 1 as the probabilities \( p(A_i) \) approach zero. This argument suggests constructing \( p(A) \) via formula (13).

Even though the above motivation for formula (13) is relatively transparent, it takes a lot of work in the Appendix (Theorem A.1) to show that the function \( p : \mathcal{R} \rightarrow [0, 1] \) given by (13) is indeed a finely ranged probability measure that represents \( \geq_0 \).

In contrast with the construction of subjective probability by Savage (or by Fishburn [7]) on a \( \sigma \)-algebra and by Zhang [22] on a countably additive \( \lambda \)-system, formula (13) does not rely on

\[^{14}\text{Take } N \text{ such that } \frac{1}{N} < p(A) \leq \frac{1}{N-1} \text{. Show that } v(A) = N. \text{ First, } v(A) \geq N \text{ because } S \text{ can be partitioned into } v(A) \text{ events that all have probabilities smaller than } \frac{1}{N-1}. \text{ To prove } v(A) \leq N, \text{ construct an } N \text{-element grand partition finer than } A. \text{ As } p \text{ is finely ranged, there is a grand partition } \{S_1, \ldots, S_m\} \text{ such that } p(S_i) < p(A) - \frac{1}{N} \text{ for all } i = 1 \ldots m. \text{ Divide this partition into groups}
\]

\[
S = S_1 \cup \cdots \cup S_{k_1} \cup \cdots \cup S_{k_{N-2}} \cup \cdots \cup S_{k_{N-1}} \cup S_{k_{N-1}+1} \cup \cdots \cup S_m
\]

so that \( \frac{1}{N} < p(B_i) < p(A) \) for all \( i = 1 \ldots N-1 \). Then \( p(B_N) < 1 - \frac{N-1}{N} = \frac{1}{N} < p(A) \), and hence, the grand partition \( \{B_1, \ldots, B_N\} \) is finer than \( A \).
partitioning the state space $S$ into an arbitrarily large number $N$ of equiprobable events. As the following example illustrates, such equipartitions need not exist if the domain of risky events is a mosaic or even a finitely additive algebra.

Adopt the coin-tossing setting, where $S = \prod_{k=1}^{\infty} \{H_k, T_k\}$. For an arbitrary finite $n$, identify every $A \subseteq \prod_{k=1}^{n} \{H_k, T_k\}$ with the obvious subset of $S$, and call $A$ a cylinder of rank $n$. Let $\Sigma$ be the set of all cylinders, and $\mathcal{R} = \Sigma$. Construct a finely ranged probability measure $p^*: \Sigma \to [0,1]$ such that for all $A, B \in \Sigma$,

$$A \neq B \implies p^*(A) \neq p^*(B). \quad (14)$$

Let $p^*(s_1, \ldots, s_n) = \pi_1(s_1) \cdot \pi_2(s_2) \cdot \ldots \cdot \pi_n(s_n)$, where each $\pi_i$ is a probability measure on $\{H_i, T_i\}$. Construct $\pi_1, \pi_2, \ldots$ so that (14) holds. Use induction with respect to the rank $n$ of the cylinders $A$ and $B$. For $n = 0$, $p^*(S) \neq p^*(\emptyset)$. For $n > 0$, assume that (14) holds for all cylinders of rank $n - 1$. Fix arbitrary cylinders $A$ and $B$ of rank $n$ such that $A \neq B$. Write $A = (H_n \cap A_H) \cup (T_n \cap A_T)$ and $B = (H_n \cap B_H) \cup (T_n \cap B_T)$, where $A_H, A_T, B_H, B_T$ are cylinders of rank $n - 1$. Then either $p^*(A_H) \neq p^*(B_H)$ or $p^*(A_T) \neq p^*(B_T)$, and hence, there exists at most one $\pi_n$ such that

$$\pi_n(H_n) \cdot p^*(A_H) + \pi_n(T_n) \cdot p^*(A_T) = \pi_n(H_n) \cdot p^*(B_H) + \pi_n(T_n) \cdot p^*(B_T).$$

As the number of cylinders of rank $n$ is finite, it is possible to take $\pi_n$ such $\pi_n(H_n) \in [0.4, 0.6]$, and (14) holds for all cylinders of rank $n$. By induction, there exists a sequence $\pi_1, \pi_2, \ldots$ such that $p^*: \Sigma \to [0,1]$ satisfies (14) for all $A, B \in \Sigma$ and is finely ranged because $p^*(s_1, \ldots, s_n) \leq 0.6^n$ for all $(s_1, \ldots, s_n)$.

It is obvious that (14) 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 (14) 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

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

for some complete and transitive relation $\succeq_1$ on $\mathcal{L}_p$. This observation shows that if $\Sigma$ is not a $\sigma$-algebra, then 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. (See Grant [11] for a model of probabilistic sophistication where monotonicity of risk preference is relaxed.)

### 6.2. Utility representations for risk preference

The natural step after the construction of a subjective probability measure $p$ is to define the risk preference $\succeq_1$ by taking $l \succeq_1 l'$ whenever $l = [f]_p$ and $l' = [g]_p$ for some risky acts $f \succeq g$. However, the analysis of the relation $\succeq_1$ is complicated by the fact that $\succeq_1$ is complete only on the set $\mathcal{L}_p \subseteq \mathcal{L}$ that need not have the convenient algebraic structure of $\mathcal{L}$. Accordingly, our next step is to 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 in Theorem 3.1 guarantees that such an extension exists, satisfies von Neumann–Morgenstern’s conditions, and hence, is represented by expected utility.

However, the weaker list of axioms in Theorem 4.1 does not guarantee the existence of a continuous extension of the risk preference $\succeq_1$ to the set $\mathcal{L}$ of all lotteries. For example, let
$X = \{x, y, z\}$ and suppose that no risky event is assigned the probability equal to $\frac{1}{3}$, which is the case if risky events are cylinders obtained by tossing a fair coin. Suppose that $\succeq_1$ has indifference curves illustrated by Fig. 1. It is easy to show that such a risk preference cannot be extended by continuity to all of $L$. Therefore, one cannot use Machina–Schmeidler’s construction of a utility representation. Thus, Theorem 4.1 retains the risk preference as a binary relation $\succeq_1$ on $L_p$. Alternatively, one can impose a stronger continuity condition on the underlying preference $\succeq$ that would be sufficient for the risk preference $\succeq_1$ to have a uniformly continuous utility representation (we omit a formal statement of this result).

**Appendix A. Proofs of Theorems 3.1 and 4.1**

**A.1. Preliminaries**

Fix a set of states $S$, a set of outcomes $X$, a mosaic of risky events $R \subseteq 2^S$, and a binary relation $\succeq$ over the class $G$ of risky ($R$-measurable) acts. No other primitives, such as the entire algebra $\Sigma$ or the ranking of acts other than risky, are necessary to formulate and prove Theorems 3.1 and 4.1. Hereafter in Appendix A, the adjective *risky* is omitted because it refers to all relevant objects. Thus, events—written $A$, $B$, $C$, $D$, $E$, $F$, $H$, $S$—are elements of the mosaic $R$, *partitions* belong to $R$, and *acts*—written $f$, $g$, $h$—are elements of $G$.

Write the complement of an event $A$ as $\neg A$ rather than $A^c$ (superscripts are reserved for numbering). Let $A \oplus B$ denote the union $A \cup B$ together with the fact that the sets $A$ and $B$ are disjoint. Let $A \ominus B$ denote the set difference $A \setminus B$ together with the fact that $B$ is a subset of $A$.

Write a partition $A = A_1 \oplus A_2 \oplus \cdots \oplus A_n$ briefly as $\{A = A_i^n\}$. For all indices $i, j \geq 0$, let $A_i^j = \bigcup_{k=i}^j A_k$ (if $i > j$, let $A_i^j = \emptyset$). Then $A_i^j \in R$ because $R$ is a mosaic and $S = A_1 \oplus A_2 \oplus \cdots \oplus A_n \oplus \neg A$ is a grand partition.

For any collection $E \subseteq R$ of events, let $\varepsilon(E)$ be the smallest algebra that contains $E$. If $E$ is finite, then $\varepsilon(E)$ is also finite. For $A, B \in R$, write $\varepsilon(A)$ instead of $\varepsilon([A])$, $\varepsilon(A, B)$ instead of $\varepsilon([A, B])$, etc. Note that $\varepsilon(A) = \{\emptyset, A, \neg A, S\} \subseteq R$ because $R$ is a mosaic, but $\varepsilon(A, B)$ need
not belong to \( \mathcal{R} \). For all grand partitions \( \{ S = A_i^m \} \) and \( \{ S = B_i^n \} \),
\[
\begin{align*}
&\exists (A_1, A_2, \ldots, A_m, B_1, B_2, \ldots, B_n) \subseteq \mathcal{R} \\
\iff & A_i \cap B_j \in \mathcal{R} \quad \text{for all } i = 1 \ldots m \text{ and } j = 1 \ldots n \\
\iff & \{ A_1, A_2, \ldots, A_m \} \subseteq \mathcal{R} \cap \{ B_1, B_2, \ldots, B_n \}.
\end{align*}
\]

(A.1)

A.2. 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 all \( A, B \in \mathcal{R} \), let \( A \succeq_0 B \) if and only if \( Ax' \succeq x Bx' \). Call the binary relation \( \succeq_0 \) comparative likelihood. By P4*(\( \mathcal{R} \)), this relation is invariant if the pair of outcomes \( x \succ x' \) is replaced by another pair \( z \succ z' \). We seek a probability measure \( p : \mathcal{R} \rightarrow [0, 1] \) that represents \( \succeq_0 \). Such \( p \) is called quantitative probability.

Call an event \( A \in \mathcal{R} \) non-null if \( A > 0 \emptyset \); call \( A \) null otherwise. Say that a partition \( \{ A = A_i^n \} \) is non-null if \( A_i > 0 \emptyset \) for all \( i = 1 \ldots n \). Say that a partition \( \{ A = A_i^n \} \) is finer than an event \( B \) if \( A_i > 0 B \) for all \( i = 1 \ldots n \); in this case, write \( \{ A = A_i^n \} <_0 B \).

Axioms P1(\( \mathcal{R} \)), P3(\( \mathcal{R} \)), P4*(\( \mathcal{R} \)), P5(\( \mathcal{R} \)), P6(\( \mathcal{R} \)) imply that \( \succeq_0 \) satisfies the following properties that generalize de Finetti’s [6] definition of qualitative probability and Savage’s conditions of fineness (QF) and tightness (QT).

Q1. \( \succeq_0 \) is complete and transitive.
Q2. \( S > 0 \emptyset \).
Q3. For all events \( A \in \mathcal{R} \), \( S \succeq_0 A \succeq_0 \emptyset \).
Q4. For all partitions \( \{ A = A_i^n \} \) and \( \{ B = B_i^n \} \) such that \( x(A_1, A_2, B_1, B_2) \subseteq \mathcal{R} \),
\[
A_1 \succeq_0 (>0) B_1 \quad \text{and} \quad A_2 \succeq_0 B_2 \quad \Rightarrow \quad A \succeq_0 (>0) B.
\]

By induction, Q4 implies that for all partitions \( \{ A = A_i^n \} \) and \( \{ B = B_i^n \} \) such that \( x(A_1, \ldots, A_n, B_1, \ldots, B_n) \subseteq \mathcal{R} \),
\[
A_1 \succeq_0 (>0) B_1 \quad \text{and} \quad A_i \succeq_0 B_i \quad \text{for all } i > 1 \quad \Rightarrow \quad A \succeq_0 (>0) B.
\]

Thus, Q4 asserts that \( \succeq_0 \) is additive over any subalgebra of the mosaic \( \mathcal{R} \).

**Proof.** Fix partitions \( \{ A = A_i^n \} \) and \( \{ B = B_i^n \} \) such that \( x(A_1, A_2, B_1, B_2) \subseteq \mathcal{R} \). Let \( \Sigma_* = x(A_1, A_2, B_1, B_2) \). Let \( \mathcal{F}_* \) be the class of \( \Sigma_* \)-measurable functions that map \( S \) into \( \{ x, x' \} \). Then \( \mathcal{F}_* \subseteq \mathcal{G} \). When restricted to \( \mathcal{F}_* \), \( \succeq \) satisfies Savage’s axioms P1–P5 (but violates P6 because \( \Sigma_* \) is finite). These axioms imply that the comparative likelihood \( \succeq_0 \) is additive on \( \Sigma_* \) and hence satisfies Q4 (see property C3 in [7]). \( \square \)

QF (Fineness). For any non-null event \( A \) and finite collection \( \mathcal{E} \subseteq \mathcal{R} \), there exists a grand partition \( \{ S = S_i^m \} \subseteq \mathcal{R} \cap \mathcal{E} \) finer than \( A \).

**Proof.** Fix \( A > 0 \emptyset \) and finite \( \mathcal{E} \subseteq \mathcal{R} \). Let \( f = xAx' \), \( g = x' \), and \( \mathcal{E}' = xA \cup \mathcal{E} \). Then \( \mathcal{E}' \) is finite, \( f \) and \( g \) are \( \mathcal{E}' \)-measurable, and \( A > 0 \emptyset \) implies \( f > g \). By P6(\( \mathcal{R} \)), there exists a grand partition \( \{ S = S_i^m \} \subseteq \mathcal{R} \cap \mathcal{E}' \subseteq \mathcal{R} \cap \mathcal{E} \) such that for all \( i \), \( f > xS_ig \), that is, \( A > 0 S_i \). \( \square \)
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 $f = Ax'$, $g = Bx'$, and $\varepsilon = \varepsilon(A) \cup \varepsilon(B)$. Then $\varepsilon$ is finite, $f$ and $g$ are $\mathcal{E}$-measurable, and $A \succ_0 B$ implies $f \succ g$.

By P6($\mathcal{R}$), there exists a grand partition $\{S = S_1^m\} \subseteq \mathcal{R} \cap \mathcal{E}$ such that $x' S_i f \succ g$ for all $i$. Let $C_i = A \setminus S_i$ and $D_i = A \cap S_i$. By (A.1), $\varepsilon(S_1, \ldots, S_m, A, \neg A) \subseteq \mathcal{R}$ and $C_i$, $D_i \in \mathcal{R}$. Take $k$ such that $D_k \succ_0 \emptyset$; if there is no such $k$, then by Q4, $\emptyset \supseteq D_0^m = A$, which contradicts $A \succ_0 B \supseteq \emptyset$. As $x' S_k f \succ g$, then $C_k \succ_0 B$, and $A = C_k \oplus D_k$ is a non-null partition.

By P6($\mathcal{R}$), there exists (another) grand partition $\{S = S_1^m\} \subseteq \mathcal{R} \cap \mathcal{E}$ such that $f \succ x S_i g$ for all $i$. Let $C_i = (\neg B) \cap S_i$ and $D_i = (\neg B) \setminus S_i$. By (A.1), $\varepsilon(S_1, \ldots, S_m, B, \neg B) \subseteq \mathcal{R}$ and $C_i$, $D_i \in \mathcal{R}$. Take $k$ such that $C_k \succ_0 \emptyset$; if there is no such $k$, then by Q4, $B \succ_0 B \oplus C_1^m = S$, which contradicts $S \succ_0 A \succ_0 B$. As $f \succ x S_k g$, then $A \succ_0 B \ominus C_k$. The partition $\neg B = C_k \ominus D_k$ is non-null; otherwise $D_k$ is null, and by Q4 there is a contradiction $B \oplus C_k \succ_0 (B \oplus C_k) \oplus D_k = S \succ_0 A$. □

While a quantitative probability representation need not exist for an arbitrary qualitative probability [15], Savage shows that such a representation exists for any fine and tight qualitative probability on a $\sigma$-algebra. Niiniluoto [18] and Wakker [21] extend this result to algebras. We extend it further to mosaics. Thus, we derive subjective probability from the comparative likelihood relation $\succ_0$ rather than directly from the preference $\succeq$. This construction applies in the proofs of both Theorems 3.1 and 4.1; it applies also in a de Finetti-type setting where $\succ_0$ is taken as a primitive and Q1–Q4, QF, QT are imposed as axioms.

Deduce the following properties Q5–Q8 from Q1 to Q4, QF, QT.

**Q5.** For all partitions $\{A = A_1^2\}$ and $\{B = B_1^2\}$, if $A_1 \succ_0 (\succ_0) B_1$ and $A_2 \succ_0 B_2$, then $A \succ_0 (\succ_0) B$.

In contrast with Q4, Q5 does not require that $\varepsilon(A_1, A_2, B_1, B_2) \subseteq \mathcal{R}$. From Q5, it follows by induction that for all partitions $\{A = A_1^n\}$ and $\{B = B_1^n\}$,

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

**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 $A_2 \succ_0 B_2$. By QT, there are 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 $C_2 \succ_0 B_2$. Let

$$\varepsilon = \varepsilon(C_1, D_1, C_2, D_2, \neg A) \cup \varepsilon(B_1, B_2, \neg B).$$

Take a grand partition $\{S = S_1^m\} \subseteq \mathcal{R} \cap \mathcal{E}$ that is finer than both $D_1$ and $D_2$. Let $k_1$ be the minimal index such that $S_1^{k_1} \succ_0 C_1$, and $k_2$ be the maximal index such that $S_2^{k_2} \succ_0 C_2$. By (A.1), $\varepsilon(S_1, \ldots, S_m, C_1, D_1, C_2, D_2, \neg A) \subseteq \mathcal{R}$. Then

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

and

$$A_2 = C_2 \oplus D_2 \succ_0 S_2^{k_2+1} \oplus S_{k_2} = S_2^{k_2} \succ_0 C_2 \succ_0 B_2.$$
Note that $S_{k_1}^i$ and $S_{k_2}^m$ are disjoint; otherwise $k_2 \leq k_1$, $A_2 \succ 0 \ S_{k_2}^m \succ A_1 \ S_{k_1}^i$ and by Q4, there is a contradiction: $A = A_1 \oslash A_2 \succ 0 \ S_{k_1}^i \oslash S_{k_1+1}^i = S$. Thus,

$$A = A_1 \oslash A_2 \succ 0 \ S_{k_1}^i \oslash S_{k_2}^m \succ 0 \ B_1 \oslash B_2 = B$$

because $x(S_{k_1}^i, S_{k_2}^m, A_1, A_2) \subseteq R$ and $x(S_{k_1}^i, S_{k_2}^m, B_1, B_2) \subseteq R$.

**Case II:** $A_1 \succ 0 \ B_1$ and $A_2 \geq 0 \ B_2$. By QT, there is a non-null partition $A_1 = C_1 \oslash D_1$ such that $C_1 \succ 0 \ B_1$. By Case I, $A = C_1 \oslash (A_2 \oslash D_1) \succ 0 \ B_1 \oslash B_2 = B$.

**Case III:** $A_1 \geq 0 \ B_1$ and $A_2 \geq 0 \ B_2$. Suppose that $B \succ 0 \ A$. Then there is a non-null partition $\neg A = C' \oslash D'$ such that $B \succ 0 \ A \oslash C'$. Yet by Case II, $A \oslash C' = (A_1 \oslash C') \oslash A_2 \succ 0 \ B_1 \oslash B_2 = B$. This contradiction implies $A \geq 0 \ B$. □

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

**Proof.** By QT, there exists a grand partition $\{ S = S_i^m \} \subseteq R \oslash \{ B, \neg B \}$ finer than $A$. For all $i$, the sets $B_i = B \cap S_i$ and $\neg B \cap S_i$ belong to $R$ and partition $S_i$. Therefore, $B_i \geq 0 \ S_i \prec 0 \ A$, and the partition $\{ B = B_i^m \}$ is finer than $A$. □

**Q7.** For any events $A \succ 0 \ B$ and for any non-null partition $\{ A = A_i^1 \}$, there exists an $n$-element partition $\{ B = B_i^n \}$ such that $A_i \succ 0 \ B_i$ for all $i = 1 \ldots n$.

**Proof.** Take $n = 2$. Fix events $A \succ 0 \ B$ and a non-null partition $\{ A = A_i^2 \}$. By QT, there is a non-null partition $\neg B = C' \oslash D'$ such that $A \succ 0 \ B \oslash C'$. By Q6, there is a partition $\{ B = E_i^m \}$ finer than $C'$. Let $k$ be the maximal index in $[0, m]$ such that $A_1 \succ 0 \ E_i^k$. Suppose that $E_{k+1}^m \succ 0 \ A_2$. Then $k < m$,

$$E_i^k \oslash C' \succ 0 \ E_i^k \oslash E_{k+1}^m = E_i^{k+1} \text{ def. of } k \geq 0 \ A_1$$

and

$$B \oslash C' = (E_i^k \oslash C') \oslash E_{k+1}^m \succ 0 \ A_1 \oslash A_2 = A.$$ 

This contradiction implies that $A_2 \succ 0 \ E_{k+1}^m$, and $B = E_i^k \oslash E_{k+1}^m$ is the required two-element partition of $B$. Complete the proof by induction with respect to $n$. □

**Q8.** For any events $A \succ 0 \ B$ and for any partition $\{ B = B_i^n \}$, there exists an $n$-element partition $\{ A = A_i^1 \}$ such that $A_i \succ 0 \ B_i$ for all $i = 1 \ldots n$.

**Proof.** Take $n = 2$. Fix events $A \succ 0 \ B$ and a partition $\{ B = B_i^2 \}$. By QT, there is a non-null partition $A = C \oslash D$ such that $C \succ 0 \ B$. By Q6, there is a partition $\{ C = E_i^m \}$ finer than $D$. Let $k$ be the maximal index in $[0, m]$ such that $B_1 \succ 0 \ E_i^k$. Then $k < m$, and $E_i^k \oslash D \succ 0 \ E_i^k \oslash E_{k+1}^m \geq 0 \ B_1$. Suppose that $B_2 \geq 0 \ E_{k+1}^m$. Then by Q5, $B_1 \oslash B_2 \succ 0 \ E_i^k \oslash E_{k+1}^m = C$. This contradiction implies that $E_{k+1}^m \succ 0 \ B_2$, and $A = (E_i^k \oslash D) \oslash E_{k+1}^m$ is the required partition of $A$. Complete the proof by induction with respect to $n$. □
For every $A \in \mathcal{R}$, let

$$v(A) = \min_{\{S = S^m_1\} < A} m \quad \text{and} \quad r(A) = \frac{1}{v(A)}.$$  

By definition, $v(A)$ is the minimal number of events that may constitute a grand partition finer than $A$. If $A$ is non-null, then by QF, $v(A)$ is finite; if $A$ is null, then $v(A) = +\infty$. Thus, the value $r(A)$ belongs to the set $\left\{\frac{1}{2}, \frac{1}{3}, \ldots, 0\right\} \subseteq [0, 1]$, and $r(A) = 0$ if and only if $A$ is null.

The function $r : \mathcal{R} \to [0, 1]$ satisfies the following properties R1–R7.

**R1.** The function $r$ almost agrees with $\geq 0$, that is, for all events $A$ and $B$,

$$A \geq_0 B \Rightarrow r(A) \geq r(B)$$

and

$$r(A) > r(B) \Rightarrow A \succ_0 B.$$  

**Proof.** Any partition finer than $B \leq_0 A$ is also finer than $A$. Therefore, $A \geq_0 B$ implies $v(A) \leq v(B)$ and $r(A) \geq r(B)$. Conversely, $r(A) > r(B)$ implies $A \succ_0 B$. □

**R2.** For all partitions $\{A = A^1_i\}$, $\max_{i=1\ldots,n} v(A_i) > n$, or equivalently, $\min_{i=1\ldots,n} r(A_i) < \frac{1}{n}$.

**Proof.** Fix a partition $\{A = A^1_i\}$. Pick $A_k$ so that $A_i \geq_0 A_k$ for all $i = 1 \ldots n$. Suppose that $n \geq v(A_k)$. Take a grand partition $\{S = S^v_1(A_k)\}$ finer than $A_k$. Then $A_i \geq_0 A_k \succ_0 S_i$ for all $i = 1 \ldots v(A_k)$. By Q5, $A^v_1(A_k) > 0 S^v_1(A_k) = S$. This contradiction implies $v(A_k) > n$. □

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

**Proof.** For any $A \succ_0 \emptyset$, take a grand partition $\{S = S^v_1(A)\}$ finer than $A$. This partition is non-null; otherwise $S$ could be partitioned into $v(A) - 1$ events finer than $A$. By R2, $\max_{i=1\ldots\nu(A)} v(S_i) > v(A)$. Thus, the function $v$ 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 grand partition $\{S = S^v_1(A)\}$ finer than $A$. Pick $S_k$ so that $S_k \geq_0 S_i$ for all $i$. By QT, there exists a non-null partition $A = C \oplus D \succ_0 S_k$ such that $C \succ_0 S_k$. Then $\{S = S^v_1(A)\}$ is finer than $C$. Therefore, $v(C) \leq v(A)$. On the other hand, $A \geq_0 C$ implies $v(C) \geq v(A)$. Thus, $r(C) = r(A)$. □

Given a non-null event $B$, say that $\{A = A^1_i\}$ is a $B$-partition if

(i) $r(B_i) = r(B)$ for all $i = 1 \ldots n - 1$ (but not necessarily for $i = n$), and

(ii) this partition is finer than $B$.

**R5.** For any $A \in \mathcal{R}$ and for any non-null $B \in \mathcal{R}$, there exists a $B$-partition $\{A = A^1_i\}$.

**Proof.** Say that a partition $\{A = A^1_i\}$ is $B$-acceptable, if $r(A_i) = r(B)$ and $A_i \prec_0 B$ for all $i = 1 \ldots n - 1$ (but not necessarily for $i = n$). For example, the trivial partition $\{A = A\}$ is
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 $A \in \mathcal{R}$,

$$p(A) = \sup_{\{A = A^n_1\}} \sum_{i=1}^n r(A_i). \quad \text{(A.2)}$$

Proof. Suppose first that a finely ranged probability measure $p : \mathcal{R} \to [0, 1]$ represents $\succeq_0$. The proof of Q1–Q4, QF is straightforward. To show QT, fix arbitrary events $A \triangleright_0 B$, and let $\varepsilon = p(A) - p(B) > 0$. Take a grand partition $S = S^{m_1}_1 \subseteq \mathcal{R} \cap \{A, \neg A, B, \neg B\}$ such that...
Let \( D = E_k \). Then \( p(A \ominus D) > p(A) - \varepsilon = p(B) \), and \( A \ominus D >_0 B \). The proof of the second part of QT is analogous.

Suppose instead that \( \succeq_0 \) satisfies Q1–Q4, QF, QT. Define \( p : \mathcal{R} \to [0, 1] \) by (A.2). Show that \( p \) is a probability measure. Fix a grand partition \( \{ S = S_i^{m_i} \} \), a non-null event \( B \), and \( B \)-partitions \( \{ S_j = B(i)_{n_j}^m \} \). Then \( \sum_{i=1}^m n_i \varepsilon(B) = \frac{1}{r(B)} \) because the partition of \( S \) into \( \sum_{i=1}^m n_i \) events \( B(i)_j \) is finer than \( B \). It follows that

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

As \( r(B) \) can be arbitrarily small, \( \sum_{i=1}^m p(S_i) \geq 1 \).

By R7, \( \sum_{i=1}^m \sum_{j=1}^{m_i} r(A(i)_j) < 1 \) for all partitions \( \{ S_i = A(i)_{m_i} \} \). Hence,

\[
\sum_{i=1}^m p(S_i) = \sum_{i=1}^m \left( \sup_{\{ S = A(i)_{m_i} \} \sum_{j=1}^{m_i} r(A(i)_j) \right) \leq 1,
\]

that is, \( \sum_{i=1}^m p(S_i) = 1 \). As \( p \) is non-negative, \( p \) is a probability measure.

Show that \( p \) represents \( \succeq_0 \). Fix events \( A >_0 B \). Then by QT, there exists a non-null partition \( A = C \oplus D \) such that \( C >_0 B \). By Q8, for any partition \( \{ B = B_i \} \), there exists a partition \( \{ C = C_i \} \) such that \( C_i >_0 B_i \) for all \( i = 1 \ldots n \). Thus,

\[
p(A) \geq r(D) + \sup_{\{ C = C_i \}} \sum_{i=1}^n r(C_i) \geq r(D) + \sup_{\{ B = B_i \}} \sum_{i=1}^n r(B_i) = r(D) + p(B) > p(B).
\]

Fix (other) events \( A, B \) such that \( p(A) > p(B) \). Take a partition \( \{ A = A_i \} \) such that \( \sum_{i=1}^n r(A_i) > p(B) \) and \( r(A_1) > 0 \). By R4, there exists a non-null partition \( A_1 = C \oplus D \) such that \( r(A_1) = r(C) \). Then \( p(A \ominus D) \geq \sum_{i=1}^n r(A_i) > p(B) \), and \( A \ominus D \succeq_0 B \). Thus, \( A >_0 B \), and \( p \) represents \( \succeq_0 \).

Show that \( p \) is finely ranged. Fix a finite collection \( \mathcal{E} \subseteq \mathcal{R} \) and \( \varepsilon > 0 \). Take a non-null event \( B \) such that \( r(B) < \varepsilon \) and a (non-null) grand partition \( \{ S = B_i \} \) finer than \( B \). Then \( \sum_{i=1}^n p(B_i) = 1 \). Therefore, there exists \( B_i >_0 \emptyset \) such that \( p(B_i) \leq \frac{1}{r(B)} < \varepsilon \). By QF, there exists a grand partition \( \{ S = S_i^m \} \subseteq \mathcal{R} \cap \mathcal{E} \) 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, take a grand partition \( \{ S = S_i^m \} \) finer than \( A \). As \( p^* \) is additive and represents \( \succeq_0 \), then \( 1 = p^*(S) = \sum_{i=1}^{n(v(A))} p^*(S_i) < r(A) \cdot p^*(A) \), that is \( p^*(A) > \frac{1}{r(A)} = r(A) \). Therefore, for all partitions \( \{ A = A_i^m \} \), \( p^*(A) = \sum_{i=1}^n p^*(A_i) > \sum_{i=1}^n r(A_i) \) and \( p^*(A) \geq p(A) \). Similarly, \( p^*(\neg A) \geq p(\neg A) \). Thus, \( p(A) = p^*(A) \). \( \square \)

Next, we formulate one useful property of finely ranged probability measures.

**Lemma A.2.** If \( p : \mathcal{R} \rightarrow [0, 1] \) is a finely ranged probability measure, then for all \( n > 0 \), for all finite collections \( \mathcal{E} \subseteq \mathcal{R} \), and for all events \( A \in \mathcal{R} \cap \mathcal{E} \), the set

\[
p(A, n, \mathcal{E}) = \{(p(A_1), p(A_2), \ldots, p(A_n)) : \{ A = A_1^n \} \subseteq \mathcal{R} \cap \mathcal{E} \}
\]
of vectors \((p(A_1), p(A_2), \ldots, p(A_n)) \in \mathbb{R}_+^n\) is dense in the simplex

\[
\Delta(A, n) = \left\{ (v_1, v_2, \ldots, v_n) \in \mathbb{R}_+^n : \sum_{i=1}^n v_i = p(A) \right\}.
\]

**Proof.** Fix a finite collection \(E \subseteq R\), an event \(A \in R \cap E\), and numbers \(v \in [0, p(A)]\) and \(\varepsilon > 0\). Show that there exists a partition \(\{A = A_1 \} \subseteq R \cap E\) such that \(v - \varepsilon < p(A_1) \leq v\). Take a finite collection \(E'\) that contains \(x(A)\) and the intersections \(A \cap E\) for all \(E \in E\). As \(p\) is finely ranged, there exists a grand partition \(\{S = S_1 \} \subseteq R \cap E'\) such that \(p(S_i) < \varepsilon\) for all \(i = 1 \ldots m\). For all \(i = 1 \ldots m\), the sets \(A \cap S_i\) and \(\neg A \cap S_i\) belong to \(R\), and \(p(A \cap S_i) < p(S_i) < \varepsilon\). Let \(k \in [0, m]\) be the minimal index such that \(p(A \cap S_k^1) > v - \varepsilon\). Then \(p(A \cap S_k^1) < v\) because otherwise \(p(A \cap S_k^{k-1}) > v - p(A \cap S_k) > v - \varepsilon\). Let \(A_1 = A \cap S_k^1\) and \(A_2 = A \cap S_k^{m+1}\). For each \(E \in E\), the sets \(A \cap S_i) \cap E = S_i \cap (A \cap E)\) belong to \(R\) for all \(i = 1 \ldots m\) and, together with \(\neg(A \cap E)\), form a grand partition. The sets \(A_1 \cap E\) and \(A_2 \cap E\) are unions of elements of this grand partition and hence, are also events in \(R\). Thus, \(A_1 \in R \cap E\) and \(A_2 \in R \cap E\).

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 \(\{A = A_1 \} \subseteq R \cap E\) such that \(v_i - \varepsilon/2n < p(A_i) \leq v_i\) for all \(i = 1 \ldots n - 1\). Then

\[
\sum_{i=1}^n |p(A_i) - v_i| < (n - 1) \cdot \frac{\varepsilon}{2n} + |p(A_n) - v_n| < 2 \cdot (n - 1) \cdot \frac{\varepsilon}{2n} = \varepsilon.
\]

Thus, \(p(A, n, E)\) is dense in \(\Delta(A, n)\). \(\square\)

### A.3. Preference and first-order stochastic dominance

Suppose that \(\succeq\) satisfies P1\(\mathcal{R}\), P3\(\mathcal{R}\), P4\(\mathcal{R}\), P5\(\mathcal{R}\), P6\(\mathcal{R}\). Then the finely ranged probability measure \(p : \mathcal{R} \rightarrow [0, 1]\) defined by (A.2) represents the preference over binary acts that have outcomes \(x > x'\). The following lemma characterizes the preference \(\succeq\) between arbitrary acts \(f\) and \(g\) when the lotteries \([f]_p\) and \([g]_p\) induced by these acts via the measure \(p\) satisfy the first-order stochastic dominance.

**Lemma A.3.** For all acts \(f\) and \(g\),

\[
[f]_p \succeq [g]_p \Rightarrow f \succeq g,
\]

(A.3)

\[
[f]_p \succeq [g]_p \Rightarrow f \geq g.
\]

(A.4)

**Proof.** Write arbitrary acts \(f\) and \(g\)

\[
f = \begin{bmatrix}
z_1 & \text{if } s \in F_1 \\
z_2 & \text{if } s \in F_2 \\
\vdots \\
z_n & \text{if } s \in F_n \\
h(s) & \text{if } s \in H
\end{bmatrix}
\]

and \(g = \begin{bmatrix}
z_1 & \text{if } s \in G_1 \\
z_2 & \text{if } s \in G_2 \\
\vdots \\
z_n & \text{if } s \in G_n \\
h(s) & \text{if } s \in H
\end{bmatrix}\)

for some \(n\), for some outcomes \(z_1 \geq z_2 \geq \cdots \geq z_n\), 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 \(S = G_1 \oplus G_2 \oplus \cdots \oplus G_n \oplus H\). Say that acts \(f\) and \(g\) differ among \(n\) outcomes.
Prove (A.3) and (A.4) by induction with respect to \( n \). If \( n = 1 \), then \( f = g \), and (A.3) and (A.4) hold. Fix \( n > 1 \) and suppose that (A.3) and (A.4) hold for all acts that differ among \( n - 1 \) outcomes. Fix arbitrary acts \( f \) and \( g \) that differ among \( n \) outcomes. If \( z_i \sim z_j \) for some \( i \neq j \), then by P3(\( \mathcal{R} \)), \( z_iF_j f \sim f \) and \( z_iG_j g \sim g \). The dominance \([f]_p \geq (\geq)[g]_p\) implies \([z_iF_j f]_p \geq (\geq)[z_iG_j g]_p\), and hence, the ranking \( z_iF_j f \geq (\geq)z_iG_j g \) because the acts \( z_iF_j f \) and \( z_iG_j g \) differ among \( n - 1 \) outcomes. By transitivity, \( f \geq g \).

Wlog, \( z_1 > z_2 > \cdots > z_n \). Then \([f]_p \geq [g]_p\) if and only if \( p(F^k_1) \geq p(G^k_1) \) for all \( k = 1 \ldots n; [f]_p \geq [g]_p\) if and only if \( p(F^k_1) \geq p(G^k_1) \) for all \( k = 1 \ldots n \) and \( p(F^k_1) > p(G^k_1) \) for some \( k \). Consider several cases.

**Case I:** \( p(F^k_1) > p(G^k_1) \) for all \( k = 1 \ldots n - 1 \), and \( z(F_1, F_2, \ldots, F_n, G_1, G_2, \ldots, G_n) \subseteq \mathcal{R} \). Take \( \epsilon > 0 \) such that \( p(F^k_1) > p(G^k_1) + \epsilon \) for all \( k = 1 \ldots n - 1 \). For all \( i = 2, \ldots, n \), let \( B_i = F_i \cap G_1 \) and \( v_i = p(B_i) \). Then \( p(F_1 \setminus G_1) = \sum_{i=1}^n v_i \), where

\[
v_1 = p(F_1 \setminus G_1) - \sum_{i=2}^n v_i = p(F_1 \setminus G_1) - p(G_1 \setminus F_1) = p(F_1) - p(G_1).
\]

By Lemma A.2, there exists a partition \( \{F_1 \setminus G_1 = A^n_i\} \) such that \( p(A_i) > v_1 - \epsilon \) and \( p(A_i) > v_i \) for \( i = 2, \ldots, n \). Let \( f_i = z_2A_1f \) and \( f_i = (z_1B_i z_i)(A_i \cup B_i) f_{i-1} \) for all \( i = 2 \ldots n \). By P3(\( \mathcal{R} \)), \( f \geq f_1 \). As \( A_i > B_i \), then by P4*(\( \mathcal{R} \)), \( f_{i-1} = (z_1A_1 z_i)(A_i \cup B_i) f_{i-1} > f_i \). By transitivity, \( f \geq f_n \). The acts \( f_n \) and \( g \) can be written as

\[
f_n = \begin{bmatrix} z_2 & \text{if } s \in E_2 \cr \vdots \cr z_1 & \text{if } s \in G_1 \cr h(s) & \text{if } s \in H \end{bmatrix}
\]

and

\[
g = \begin{bmatrix} z_2 & \text{if } s \in G_2 \cr \vdots \cr z_1 & \text{if } s \in G_1 \cr h(s) & \text{if } s \in H \end{bmatrix},
\]

where \( E_2 = (F_2 \oplus A_1 \oplus A_2) \ominus B_2 \), and \( E_i = (F_i \oplus A_i) \ominus B_i \) for \( i = 3 \ldots n \). The dominance \([f_n]_p \geq [g]_p\) holds because for all \( k = 2 \ldots n - 1 \),

\[
p(E^k_2) > p(A_1) + \sum_{i=2}^{k} p(F_i) > p(F_1) - p(G_1) - \epsilon + \sum_{i=2}^{k} p(F_i) = -p(G_1) - \epsilon + \sum_{i=1}^{k} p(F_i) > p(G^k_2).
\]

Then \( f_n \geq g \) because \( f_n \) and \( g \) differ among \( n - 1 \) outcomes. By transitivity, \( f \geq g \).

**Case II:** \( p(F^k_1) > p(G^k_1) \) for all \( k = 1 \ldots n - 1 \). By Lemma A.2, there exists a partition \( \{-H = A^n_i\} \) that belongs to \( \mathcal{R} \cap (z(F_1, \ldots, F_n) \cup z(G_1, \ldots, G_n)) \) and satisfies \( p(F^k_1) > p(A^n_i) > p(G^k_1) \) for all \( k = 1 \ldots n - 1 \). Take an act \( f' \) such that \( f'(s) = z_i \) if \( s \in A_i \) for \( i = 1 \ldots n \), and \( f'(s) = h(s) \) if \( s \in H \). Case I implies that \( f \geq f' \geq g \) and hence, that \( f \geq g \).

**Case III:** \( p(F^k_1) \geq p(G^k_1) \) for all \( k = 1 \ldots n \). Suppose that, contrary to (A.3), \( g \succ f \). If \( F_n \) is null, then \([z_{n-1}F_n f]_p \geq [z_{n-1}G_n g]_p\) and \( z_{n-1}F_n f \geq z_{n-1}G_n g \) because the acts \( z_{n-1}F_n f \) and \( z_{n-1}G_n g \) differ among \( n - 1 \) outcomes. By P3(\( \mathcal{R} \)), \( z_{n-1}F_n f \sim f \) and \( z_{n-1}G_n g \sim g \) and hence, \( f \geq g \). If \( F_n \) is non-null, then by P6(\( \mathcal{R} \)), there exists a non-null partition \( \{F_n = E^n_i\} \) such that \( g > x_1E_1 f \). Then \( p(F^k_1 \oplus E_1) > p(G^k_1) \) for all \( k = 1 \ldots n - 1 \). By Case II, \( x_1E_1 f > g \). This contradiction proves (A.3).
Case IV: \( p(F^k_1) \geq p(G^k_1) \) for all \( k = 1 \ldots n \) and \( p(F^k_1) > p(G^k_1) \) for some \( k \in [1, n] \). There exists \( k \) such that \( p(F^k_1) > p(G^k_1) \) and \( p(F^{k+1}_1) = p(G^{k+1}_1) \). 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 \( p(D) < p(F^k_1) - p(G^k_1) \). Then \( [f]_p \geq [zkDg]_p \), and by Case III, \( f \geq zkDg \). By P3(\( R \)), \( zkDg > g \). Thus \( f > g \), which proves (A.4). \( \square \)

Define a metric on the set \( L \) of all lotteries by
\[
\|l - l'\| = \sum_{x \in X} |l(x) - l'(x)| \quad \text{for } l, l' \in L.
\]
The next two lemmas relate this metric and the first-order stochastic dominance. Assume that the set of outcomes \( X \) is finite. Then there exist \( x^*, x_* \in X \) such that \( x^* \geq x \geq x_* \) for all \( x \in X \). Write \( l \geq l' + o \) if lotteries \( l, l' \in L \) are such that \( l \geq l'' \) for all \( l'' \in L \) in some neighborhood of \( l' \); write \( l + o \geq l' \) if \( l'' \geq l' \) for all \( l'' \in L \) in some neighborhood of \( l \). Note that \( x^* \geq x^* + o \) and \( x_* + o \geq x_* \). Thus, neither \( l \geq l' + o \) nor \( l + o \geq l' \) implies \( l \geq l' \).

**Lemma A.4.** If \( X \) is finite, then for any acts \( g > g' \), there exist acts \( h > h' \) such that \( [g]_p + o \geq [h]_p \) and \( [h']_p \geq [g']_p + o \).

**Proof.** Fix acts \( g > g' \). Let the support of the lottery \( l = [g]_p \) consist of outcomes \( y_1 \geq y_2 \geq \cdots \geq y_n \). By P6(\( R \)), the non-null event \( g^{-1}(y_1) \) has a non-null partition \( \{g^{-1}(y_1) = E^m_1\} \) such that \( x_1E_1g > g' \). Let \( h = x_*E_1g \). For each lottery \( l'' \in L \) such that \( \|l - l''\| < p(E_1) \) and for each \( x \in X \), \( l''(Y_x) \geq [h]_p(Y_x) \). Indeed, if \( x \sim x_* \), then \( l''(Y_x) = 1 = [h]_p(Y_x) \); if \( y_1 \geq x > x_* \), then \( l''(Y_x) \geq [g]_p(Y_x) - p(E_1) = [h]_p(Y_x) \) because \( \|l - l''\| < p(E_1) \) and \( g^{-1}(Y_x) = h^{-1}(Y_x) \oplus E_1 \); if \( x > y_1 \), then \( l''(Y_x) \geq 0 = [h]_p(Y_x) \). Therefore, \( l'' \geq [h]_p \), and hence, \( [g]_p + o \geq [h]_p \). Analogously, \( h > g' \) implies that there exists a non-null \( E_1 \) such that \( h > h' = x_*E_1g' \) and \( [h']_p \geq [g']_p + o \). \( \square \)

**Lemma A.5.** If \( X \) is finite, then for any lottery \( l \in L \) and for any non-null event \( E \), there exists an act \( h \in G \cap \{\neg E\} \) such that \( \|h - l\| < p(E) \) and \( [x^*Eh]_p \geq l + o \geq [x_*Eh]_p \).

**Proof.** Fix a lottery \( l \) and a non-null event \( E \). Let the support of \( l \) consist of outcomes \( y_1 \geq y_2 \geq \cdots \geq y_n \). Take \( \varepsilon = \frac{1}{2} \cdot p(E) \cdot \min(l(y_1), l(y_n)) \). By Lemma A.2, there exist partitions \( \{E = A^n_1\} \) and \( \{\neg E = B^n_1\} \) such that for all \( i = 1 \ldots n \),
\[
p(E) \cdot (1 + \varepsilon) \cdot l(y_i) \geq p(A_i) \geq p(E) \cdot (1 - \varepsilon) \cdot l(y_i)
\]
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 \( h \in G \) 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 \( h^* = [x^*Eh]_p \) and \( l_* = [x_*Eh]_p \). Fix an arbitrary lottery \( l'' \in L \) such that \( \|l - l''\| < \varepsilon \). For each \( x \in X \), one of the following four cases applies.
A.4. Necessity of axioms in Theorem 4.1

Suppose that for all \( f, g \in \mathcal{G} \),

\[
 f \succeq g \iff [f]_p \succeq [g]_p, \tag{A.5}
\]

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.

Show that \( \succeq \) satisfies the list of axioms in Theorem 4.1. P1(\( \mathcal{R} \)) and P5(\( \mathcal{R} \)) are trivial.

For each event \( A \in \mathcal{R} \), either \( p(A) = 0 \) or \( p(A) > 0 \). If \( p(A) = 0 \), then for all outcomes \( x, y \in X \) and for all acts \( h \in \mathcal{G} \cap \{\neg A\} \), \( [xAh]_p = [yAh]_p, [xAh]_p \sim_1 [yAh]_p \), and \( xAh \sim yAh \). If \( p(A) > 0 \), then for all acts \( h \in \mathcal{G} \cap \{\neg A\} \) and for all outcomes \( x \succeq (>) y \), the dominance \( [xAh]_p \succeq (>)[yAh]_p \) holds, which implies \( [xAh]_p \succeq_1 (>)[yAh]_p \) because \( \succeq_1 \) is strictly monotonic, and by (A.5), \( xAh \succeq (>)[yAh] \). In this case, \( x \succeq y \) if and only if \( xAh \succeq yAh \), Thus, \( \succeq \) satisfies P3(\( \mathcal{R} \)).

For all events \( A, B, C \in \mathcal{R} \cap \{A, A^c, B, B^c\} \), for all outcomes \( x \succeq x' \) and \( z \succeq z' \), and for all acts \( h \in \mathcal{G} \cap \{C^c\} \), the inequality \( p(A) \geq (>)[xAb]_p \) implies the dominance \( [(xA')Ch]_p \geq (>)[(xBx')Ch]_p \), and rankings \( [(xAx')Ch]_p \geq (>)[(xAx')Ch]_p \) and \( (xAx')Ch \geq (>)(xBx')Ch \) accordingly. Therefore, for all acts \( h, h' \in \mathcal{G} \cap \{C^c\} \),

\[
(xAx')Ch \geq (xBx')Ch \iff p(A) \geq (>)[xAb]_p \iff (za')Ch' \geq (za)Ch'.
\]

Thus, \( \succeq \) satisfies P4*(\( \mathcal{R} \)).

Show that \( \succeq \) satisfies P6(\( \mathcal{R} \)). Without loss of generality, \( X \) is finite. Fix an outcome \( x \in X \), a finite collection \( \mathcal{E} \subseteq \mathcal{R} \), and a pair of \( \mathcal{E} \)-measurable acts \( f > g \). 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 \( \delta \)-neighborhood of \( [f]_p \), and \( [f]_p \succ_1 l' \) for all \( l' \in \mathcal{L}_p \) in the \( \delta \)-neighborhood of \( [g]_p \). Partition \( S \) into events \( \{S_1, \ldots, S_n\} \subseteq \mathcal{R} \cap \mathcal{E} \) such that \( p(S_i) < \frac{\delta}{2} \) for all \( i = 1 \ldots m \). Then \( xS_if \) lies in the \( \delta \)-neighborhood of \( [f]_p \); hence, \( [xS_if]_p >_1 [g]_p \) and \( xS_if > g \). Analogously, \( f > xS_ig \).
A.7. Sufficiency of axioms in Theorem 3.1

Thus, representation (A.5) holds. The relation 

\[ P_2(\because U(fAx)) \]

for a finely ranged probability measure \( p \) is also represented by (A.5) with the same \( \because \) for some acts \( g \geq g' \).

For all acts \( f \) and \( f' \), \( f \geq f' \) implies \([f]_p \geq [f']_p \). On the other hand, \([f]_p \geq [f']_p \) implies that \([f]_p = [g]_p \) and \([f']_p = [g']_p \) for some \( g \geq g' \), and by (A.3), that \( f \sim g \geq g' \sim f' \).

Thus, representation (A.5) holds. The relation \( \geq_1 \) is non-degenerate, complete and transitive on \( L_p \) because \( \geq \) is non-degenerate, complete and transitive on \( G \). By (A.3) and (A.4), \( \geq_1 \) is strictly monotonic. Show that \( \geq_1 \) is continuous. Suppose first that \( X \) is finite. Fix arbitrary lotteries \( l, l' \in L_p \) such that \( l \succ l' \). Then \( l = [g]_p \) and \( l' = [g']_p \) for some acts \( g \geq g' \). By Lemma A.4, there exist acts \( h > h' \) such that \( l + h' \geq [h]_p \) and \( [h']_p \geq l' + h \). Therefore, for all \( l'' \in L_p \) in some neighborhood of \( l, l' \), \( l'' \geq [h]_p \geq [h']_p \geq [g']_p = l' \) and \( l'' \succ l' \) because \( \geq_1 \) is monotonic and transitive. Thus, the set \( \{l'' \in L_p : l'' \succ l' \} \) is open in \( L_p \) for all \( l' \in L_p \). Analogously, the set \( \{l'' \in L_p : l \succ l'' \} \) is open in \( L_p \) for all \( l \in L_p \). If \( X \) is infinite, then given a finite \( Y \subseteq X \), \( \geq_1 \) is continuous on \( L_p(Y) \) and by definition, is continuous on \( L_p \).

Finally, suppose that

\[ f \geq g \Leftrightarrow [f]_p^* \geq_1 [g]_p^* \]

for a finely ranged probability measure \( p^* : \mathcal{R} \to [0, 1] \) and a strictly monotonic \( \geq_1^* \). Strict monotonicity of \( \geq_1^* \) implies that \( p^* \) represents the comparative likelihood relation \( \geq_0 \) and hence, \( p^* = p \). Then \([f]_p \geq_1 [g]_p \) if and only if \([f]_p \geq_1^* [g]_p \). Thus, \( \geq_1^* \equiv \geq_1 \).

A.6. Necessity of axioms in Theorem 3.1

Suppose that \( \geq \) is represented by expected utility

\[ U(f) = \sum_{x \in X} u(x) \cdot p(f^{-1}(x)) \quad \text{for } f \in G, \quad (A.6) \]

where \( p : \mathcal{R} \to [0, 1] \) is a finely ranged probability measure, and \( u \) maps \( X \) into \( \mathbb{R} \). Then \( \geq \) is also represented by (A.5) with the same \( p \) and with the risk preference \( \geq_1 \) induced on \( L_p \) via the expected utility function \( U \) (all necessary properties of \( \geq_1 \) are easily verified). Thus, \( \geq \) satisfies the list of axioms in Theorem 4.1. Moreover, for all events \( A \subseteq \mathcal{R} \), for all acts \( f, g \in G \otimes [A] \) and for all outcomes \( x, y, U(f Ax) \geq U(g Ax) \) if and only if \( U(f Ay) \geq U(g Ay) \) because \( U(f Ax) - U(f Ay) = U(g Ax) - U(g Ay) = (u(x) - u(y))p(\neg A) \). Thus, \( \geq \) satisfies P2(\( \mathcal{R} \)).

A.7. Sufficiency of axioms in Theorem 3.1

Suppose that \( \geq \) satisfies P1(\( \mathcal{R} \)), P2(\( \mathcal{R} \)), P3(\( \mathcal{R} \)), P4(\( \mathcal{R} \)), P5(\( \mathcal{R} \)), P6(\( \mathcal{R} \)). Then P4*(\( \mathcal{R} \)) is implied by P2(\( \mathcal{R} \)) and P4(\( \mathcal{R} \)), and hence all of our previous results apply. Let \( p : \mathcal{R} \to [0, 1] \) be the probability measure given by (A.2).
Suppose that \( X \) is finite. Construct the extended risk preference \( \succeq_1 \) on the set \( \mathcal{L} \) of all lotteries rather than only on \( \mathcal{L}_p \). For all \( l, l' \in \mathcal{L} \), let

\[
l \succeq_1 l' \iff \exists \text{acts } g, g' \text{ such that } [g]_p \geq l + o \text{ and } l' + o \geq [g']_p.
\]

Fix a sequence of non-null events \( E_i \) such that \( \lim_{i \to \infty} p(E_i) = 0 \). By Lemma A.5, there exist sequences \( \{h_i\}_{i=1}^{\infty} \) and \( \{h'_i\}_{i=1}^{\infty} \) such that for all \( i \), \( \|[h_i]_p - l\| < p(E_i) \), \( \|[h'_i]_p - l'\| < p(E_i) \), \( [x^* E_i h_i]_p \geq l + o \), and \( l' + o \geq [x^* E_i h'_i]_p \). 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.7), and hence, \( l \succeq_1 l' \).

The strict preference \( g' \succ g \) holds for some acts \( g \) and \( g' \) such that \( [g]_p \geq l + o \) and \( l' + o \geq [g']_p \). Suppose that some sequences \( \{g_i\}_{i=1}^{\infty} \) and \( \{g'_i\}_{i=1}^{\infty} \) satisfy (A.7). Then \( [g]_p \geq \lim_{i \to \infty} [g_i]_p + o \) and \( \lim_{i \to \infty} [g'_i]_p + o \geq [g']_p \) imply that \( [g]_p \geq [g_i]_p \) and \( [g'_i]_p \geq [g']_p \) for sufficiently large \( i \). By (A.3), \( g \succeq g_i \succeq g'_i \succeq g' \) which contradicts \( g \succeq g' \). Thus, no sequences \( \{g_i\}_{i=1}^{\infty} \) and \( \{g'_i\}_{i=1}^{\infty} \) satisfy (A.7) and \( l \succeq_1 l' \) does not hold. To show that \( l' \geq_1 l \), fix a sequence of non-null events \( E_i \) such that \( \lim_{i \to \infty} p(E_i) = 0 \) and sequences of acts \( \{h_i\} \) and \( \{h'_i\} \) such that \( \lim_{i \to \infty} [x^* E_i h_i]_p = l \), \( \lim_{i \to \infty} [x^* E_i h'_i]_p = l' \), \( l + o \geq [x^* E_i h_i]_p \), and \( [x^* E_i h'_i]_p \geq l' + o \). Then for all \( i \), \( x^* E_i h'_i \succeq g' \succeq x^* E_i h_i \) because \( [x^* E_i h'_i]_p \geq l' \geq [g']_p \) and \( [g]_p \geq l \geq [x^* E_i h_i]_p \). By definition, \( l' \geq_1 l \). Thus, \( l' \geq_1 l \).

Thus, either \( l \geq l' \) or \( l' \geq l \) must hold. Moreover, for all lotteries \( l, l' \in \mathcal{L} \),

\[
l \geq_1 l' \iff g \geq g' \text{ for all } g, g' \in \mathcal{G} \text{ such that } [g]_p \geq l + o \text{ and } l' + o \geq [g']_p \quad (A.10)
\]

\[
l' \geq_1 l \iff g' \geq g \text{ for some } g, g' \in \mathcal{G} \text{ such that } [g]_p \geq l + o \text{ and } l' + o \geq [g']_p. \quad (A.11)
\]
Step 2: To show that \( \succeq_1 \) is transitive, fix arbitrary lotteries \( l, l', l'' \) such that \( l \succeq_1 l' \succeq_1 l'' \). Suppose that \( l'' \succ l \). By (A.11), there exist acts \( g'' \succ g \) such that \( \{g\}_p \succeq_l l + \epsilon \) and \( l'' + \epsilon \succeq_g \{g''\}_p \). Construct a non-null event \( E \) and an act \( h \in G \setminus \{\bar{E}\} \) such that \( g'' \succ_x x^* Eh > x_* Eh \succ g \). Take a non-null event \( g^{-1}(x) \). By P6(\( R \)), there exists a non-null partition \( \{g^{-1}(x) = A_i^\sharp\} \) such that \( g'' \succ x_i^* A_i^g \). Let \( h = x_i^* A_i^g \). Then \( h \succ x_i A_i^g = g \). By P6(\( R \)), there exists a non-null partition \( \{A_i = B_i'\} \) such that \( x_n B_i h \succ g \). Let \( E = B_1 \). Then \( g'' \succ x_i^* A_i h = x_i^* Eh > x_* Eh \succ g \). By Lemma A.5, there exists an act \( h' \in G \setminus \{\bar{E}\} \) such that \( \{x^* Eh\}_p \succ l + \epsilon \) and \( \{x_* Eh\}_p \). Two cases are possible: either \( x^* Eh \succeq x^* Eh' \) or \( x^* Eh' \succeq x^* Eh \). If \( x^* Eh \succeq x^* Eh' \), then \( g'' \succ x^* Eh \succeq x^* Eh' \) and by (A.11), \( l'' \succ l \). If \( x^* Eh' \succeq x^* Eh \), then by P2(\( R \)), \( x_n Eh' \succeq x_* Eh \succ g \) and by (A.11), \( l'' \succ l \). Either case contradicts \( l \succeq_1 l' \succeq_1 l'' \). Thus, \( l'' \succ l \) is impossible and \( l \succeq_1 l'' \) holds.

Step 3: To show that \( \succeq_1 \) is continuous, fix arbitrary lotteries \( l, l' \) such that \( l \succ l' \). By (A.11), there exist acts \( g \succ g' \) such that \( l + \epsilon \succeq_g \{g\}_p \) and \( l' + \epsilon \succeq_g \{g'\}_p \). For all \( l'' \) in some neighborhood of \( l, l'' + \epsilon \succeq_g \{g\}_p \) and by (A.11), \( l'' \succ l' \). Thus, the set \( \{l'' \in L : l'' \succ l' \} \) is open in \( L \) for all \( l'' \in L \). Analogously, the set \( \{l'' \in L : l'' \succ l'' \} \) is open in \( L \) for all \( l'' \in L \).

Step 4: To show representation (A.9), fix arbitrary acts \( f \) and \( f' \). If \( f \succeq f' \), then (A.7) holds for \( g_i = f \) and \( g_i' = f' \); therefore, \([f]}_p \succeq [f']_p \). If \( f \succ f' \), then by Lemma A.4, there exist acts \( h \succ h' \) such that \([f]_p + \epsilon \succeq [h]_p \) and \([h']_p \succeq [f']_p + \epsilon \). By (A.11), \([f]_p \succ [f']_p \).

Step 5: To show that \( \succeq_1 \) satisfies invariance (A.8), fix arbitrary lotteries \( l, l' \), and \( l'' \) such that \( l \succeq_1 l' \). Write the supports of \( l, l', \) and \( l'' \) as \( y_1 \geq \cdots \geq y_n, y'_1 \geq \cdots \geq y'_n, \) and \( y''_1 \geq \cdots \geq y''_n, \) respectively. Fix arbitrary \( \epsilon > 0 \) and an event \( H \) such that \( \frac{1}{2} - \epsilon < p(H) < \frac{1}{2} + \epsilon \). Let \( \pi = p(H) \).

By Lemma A.2, there exist partitions

(i) \( \{H = A_i^n\} \) and \( \{\bar{H} = B_i^n\} \) such that \( 1 - \epsilon \leq \frac{p(A_i^n)}{p(Y_i)} \leq 1 + \epsilon \) and \( 1 - \epsilon \leq \frac{p(B_i^n)}{p(Y_i)} \leq 1 + \epsilon \) for all \( i = 1 \ldots n - 1 \);

(ii) \( \{H = C_i^n\} \) and \( \{\bar{H} = D_i^n\} \) such that \( 1 - \epsilon \leq \frac{p(C_i^n)}{p(Y_i')} \leq 1 + \epsilon \) and \( 1 - \epsilon \leq \frac{p(D_i^n)}{p(Y_i')} \leq 1 + \epsilon \) for all \( i = 1 \ldots n' - 1 \);

(iii) \( \{H = F_i^n\} \) and \( \{\bar{H} = G_i^n\} \) such that \( 1 - \epsilon \leq \frac{p(F_i^n)}{p(Y_i')} \leq 1 + \epsilon \) and \( 1 - \epsilon \leq \frac{p(G_i^n)}{p(Y_i')} \leq 1 + \epsilon \) for all \( i = 1 \ldots n'' \).

Take acts \( f, f' \), and \( f'' \) such that \( f(s) = y_i \) if \( s \in A_{i} \oplus B_i \) for \( i = 1 \ldots n \), \( f'(s) = y'_i \) if \( s \in C_i \oplus D_i \) for \( i = 1 \ldots n' \), and \( f''(s) = y''_i \) if \( s \in F_i \oplus G_i \) for \( i = 1 \ldots n'' \). Then \([f]_p \succeq_1 l \succeq_1 [f']_p \). It follows that \([f]_p \succeq_1 [f']_p \) and by (A.9), \( f \succeq f' \). By P2(\( R \)), the strict rankings \( f'H f'' > f'H f' \) and \( f'' H f > f'' H f' \) imply \( f' = f'H f'' > f'H f' > f'H f = f \), and hence, cannot hold together. If \( f'H f'' > f'H f' \), let \( g = f'H f'' \) and \( g' = f'' H f' \). Then

\[
\|g\|_p - (\frac{1}{2} l + \frac{1}{2} l'') \leq \epsilon \|f'+Hf''_p\| - (\pi l + (1 - \pi)l'') \leq 3 \epsilon.
\]

Analogously, \( \|g'\|_p - (\frac{1}{2} l' + \frac{1}{2} l'' \| \leq 3 \epsilon \). If \( f'' H f \succeq_1 f'' H f' \), let \( g = f'' H f \) and \( g' = f'' H f' \). Then \( \|g\|_p - (\frac{1}{2} l + \frac{1}{2} l'') \leq 3 \epsilon \) and \( \|g'\|_p - (\frac{1}{2} l' + \frac{1}{2} l'' \| \leq 3 \epsilon \). Take \( \epsilon = \frac{1}{2} \) for \( i = 1, 2, \ldots \) and obtain acts \( g_i \geq g'_i \) such that \( g_i \rightarrow \frac{1}{2} l + \frac{1}{2} l'' \) and \( g'_i \rightarrow \frac{1}{2} l' + \frac{1}{2} l'' \). Thus, \( \frac{1}{2} l + \frac{1}{2} l'' \geq_1 \frac{1}{2} l' + \frac{1}{2} l'' \).
By the main result in Herstein–Milnor [13], which is an extension of the von Neumann–Morgenstern Theorem, \( \succeq_1 \) can be represented by expected utility for all \( l, l' \in \mathcal{L} \)

\[
l \succeq_1 l' \iff \sum_{x \in X} u(x) \cdot l(x) \geq \sum_{x \in X} u(x) \cdot l'(x),
\]

where \( u : X \rightarrow \mathbb{R} \) is unique up to a positive linear transformation. This representation, together with (A.9), implies (A.6). The uniqueness of \( p \) follows from Theorem A.1.

Finally, if \( X \) is infinite, then for all finite \( Y \subseteq X \) such that \( x, x' \in Y \), the preference over acts \( f \in \mathcal{G} \) that have range in \( Y \), \( f(S) \subseteq Y \), has a unique expected utility representation \( U(f) = \sum_{y \in Y} u_Y(y) \cdot p(f^{-1}(y)) \) where \( u_Y : Y \rightarrow \mathbb{R} \) satisfies \( u_Y(x) = 1 \) and \( u_Y(x') = 0 \). As all \( u_Y \)'s are unique, then there exist a function \( u : X \rightarrow \mathbb{R} \) that extends all of these indices simultaneously. Then \( U(f) = \sum_{x \in X} u(x) \cdot p(f^{-1}(x)) \) represents \( \succeq \) on the entire \( \mathcal{G} \).

**Appendix B. Proof of Theorem 5.1**

Fix a set of states \( S \), a set of outcomes \( X \), an algebra of events \( \Sigma \subseteq 2^X \), and a reflexive preference \( \succeq \) over the set \( \mathcal{F} \) of all acts. Derive the domains \( \mathcal{R}_Z \subseteq \Sigma \) and \( \mathcal{R}_{EZ} \subseteq \mathcal{R}_Z \) from \( \succeq \) via Zhang’s and Epstein–Zhang’s definitions, respectively.

Show that \( \mathcal{R}_Z \) is a mosaic. \( S \in \mathcal{R}_Z \) because \( x \succeq x \) for all \( x \in X \) by reflexivity. By definition, \( E \in \mathcal{R}_Z \) is equivalent to \( \neg E \in \mathcal{R}_Z \). Partition \( S \) into \( m \geq 2 \) subjectively risky events \( \{S_1, \ldots, S_m\} \subseteq \mathcal{R}_Z \). Then for all \( x, y \in X \) and \( f, g \in \mathcal{F} \),

\[
x(S_1 \cup S_2)f \succeq x(S_1 \cup S_2)g \quad \Rightarrow \quad xS_1(xS_2f) \succeq xS_1(xS_2g) \\
\quad \Rightarrow \quad yS_1(xS_2f) \succeq yS_1(xS_2g) \quad \Rightarrow \quad xS_2(yS_1f) \succeq xS_2(yS_1g) \\
\quad \Rightarrow \quad yS_2(yS_1f) \succeq yS_2(yS_1g) \quad \Rightarrow \quad y(S_1 \cup S_2)f \succeq y(S_1 \cup S_2)g.
\]

A similar argument repeated inductively for \( S_3, \ldots, S_m \in \mathcal{R}_{EZ} \) shows that for all \( x, y \in X \) and \( f, g \in \mathcal{F}, x(\cup_{i=3}^m S_i)f \succeq x(\cup_{i=3}^m S_i)g \quad \Rightarrow \quad y(\cup_{i=3}^m S_i)f \succeq y(\cup_{i=3}^m S_i)g \), or equivalently,

\[
f(S_1 \cup S_2)x \succeq g(S_1 \cup S_2)x \quad \Rightarrow \quad f(S_1 \cup S_2)y \succeq g(S_1 \cup S_2)y.
\]

Thus, \( S_1 \cup S_2 \in \mathcal{R}_Z \). The axiom \( P2(\mathcal{R}_{EZ}) \) follows immediately from the definition of \( \mathcal{R}_Z \). Theorem 3.1 delivers representation (11).

Analogously, one can show that \( \mathcal{R}_{EZ} \) is a mosaic, the only difference being that \( f \) and \( g \) are required everywhere to be joint bets. \( P4(\mathcal{R}_{EZ}) \) implies \( P4^*(\mathcal{R}_{EZ}) \) because for all \( A, B \in \mathcal{R}_{EZ}, C \in \mathcal{R}_{EZ} \cap \{A, A^c, B, B^c\} \), \( x \succ x', z \succ z' \), and \( h, h' \in \mathcal{G}_{EZ} \cap \{C^c\},

\[
(xAx')Ch \succeq (xBx')Ch \quad \Rightarrow \quad \{\text{induction by the size of the range } h(C)\},
\]

\[
(xAx')Cx' \succeq (xBx')Cx' \quad \Rightarrow \quad \{P4(\mathcal{R}_{EZ})\},
\]

\[
(zAz')Cz' \succeq (zBz')Cz' \quad \Rightarrow \quad \{\text{induction by the size of the range } h'(C)\},
\]

\[
(zAz')Ch' \succeq (zBz')Ch'.
\]

Thus, Theorem 4.1 delivers representation (12).
References