

MOHAMMED ABDELLAOUI and PETER P. WAKKER

## THE LIKELIHOOD METHOD FOR DECISION UNDER UNCERTAINTY\*

**ABSTRACT.** This paper introduces the likelihood method for decision under uncertainty. The method allows the quantitative determination of subjective beliefs or decision weights without invoking additional separability conditions, and generalizes the Savage–de Finetti betting method. It is applied to a number of popular models for decision under uncertainty. In each case, preference foundations result from the requirement that no inconsistencies are to be revealed by the version of the likelihood method appropriate for the model considered. A unified treatment of subjective decision weights results for most of the decision models popular today. Savage’s derivation of subjective expected utility can now be generalized and simplified. In addition to the intuitive and empirical contributions of the likelihood method, we provide a number of technical contributions: We generalize Savage’s nonatomicity condition (“P6”) and his assumption of ( $\sigma$ ) algebras of events, while fully maintaining his flexibility regarding the outcome set. Derivations of Choquet expected utility and probabilistic sophistication are generalized and simplified similarly. The likelihood method also reveals a common intuition underlying many other conditions for uncertainty, such as definitions of ambiguity aversion and pessimism.

**KEY WORDS:** Likelihood method, Subjective expected utility, Probabilistic sophistication, Choquet expected utility, Rank dependence, Ambiguity, Belief measurement.

### 1. INTRODUCTION

Keynes (1921) and Knight (1921) were the first to emphasize the difference between known and unknown probabilities, where Knight assumed that unknown probabilities are unmeasurable. Savage (1954) subsequently developed a full-blown decision model for unknown probabilities that made them measurable through observed choice, building on de Finetti’s

(1937) method of likelihood revelations from bets on events. Savage's analysis was, however, restricted to expected utility. This paper introduces the likelihood method, i.e. a generalization of the Savage–de Finetti method that can also be used if expected utility is violated.

We show how the likelihood method can be applied to several popular decision models, and can reveal subjective beliefs or decision weights in each of these. Preference foundations of each model result from imposing consistency requirements (called likelihood consistency) on those revelations. Thus, a unified treatment of subjective belief results that is applicable to most of the decision models popular today. At the end of Section 5, we explain that the resulting simple and seemingly ordinal consistency conditions can imply cardinal representations by employing the set-theoretic structure that is available on the events anyhow. A similar efficient use of the set-theoretic structure of events underlies the relative likelihood condition in Chew and Sagi (2004).

The basic idea of likelihood consistency can be recognized in axiom P2\* in Gilboa's (1987) preference foundation of Choquet expected utility. This axiom has been criticized for being complex (Fishburn, 1988, p. 202; Sarin and Wakker, 1992, p. 1257), but with the likelihood method its appealing content can be demonstrated: it comprises a direct measurement of capacities and decision weights, reflecting the new dimension of risk attitude introduced by rank-dependence, from observed preference (Sarin and Wakker, 1998, Section 5). We also apply the resulting likelihood measurement to some other models, such as probabilistic sophistication with Bayesian beliefs, where the resulting likelihood consistency generalizes Machina and Schmeidler's (1992) axiom P4\*.

We modify not only the intuitive consistency conditions of Savage (1954) and others, but also generalize the technical richness conditions. Instead of Savage's structural richness condition P6 that excludes the existence of atoms, we use a solvability condition (Luce, 1967) as the nonnecessary richness condition. This condition is easier to understand and yields more general mathematical structures. It allows for

probability measures that are not convex-ranged, and even for particular finite state spaces and atoms. We further relax Savage's requirement, originating from traditions in probability theory and statistics, and followed in most decision models, that the set of events be a sigma-algebra or an algebra. Our likelihood method turns out to work well for the more general and more flexible mosaics of events. Mosaics were introduced by Kopylov (2004) for the study of unambiguous events, and are more general than algebras in that they do not require that all intersections of events be considered. This generalization is desirable for general events, including ambiguous ones, because it facilitates the applicability of decision models. For example, in Ellsberg's two-urn paradox we are interested in drawings from separate urns, and not in joint results of those drawings. In all classical decision models we nevertheless have to assess the likelihoods of all such joint results. In our approach such irrelevant assessments are not needed. Thus, we resolve a criticism by Luce (2000) and others of Savage's (sigma-)algebras of events.

The likelihood method allows for structural generalizations and alternative intuitive interpretations of the probabilistic sophistication examined by Chew and Sagi (2004), Epstein and Zhang (2001), Machina and Schmeidler (1992), and others, and of Gilboa's (1987) Choquet expected utility. The generalizations and re-interpretations are similar to those obtained for Savage's (1954) expected utility. The likelihood method provides simplified characterizations and tests of qualitative properties of decision weights such as convexity ("pessimism" or "uncertainty aversion"). We also consider decision under risk with probabilities given, where our method can be used to measure probability transformations, so that we can also generalize Abdellaoui's (2000) preference foundation of Quiggin's (1981) rank-dependent utility.

The likelihood method reveals a common intuition and new interpretation of several conditions for uncertainty and risk examined in the literature, and unifies the analysis of several popular decision models with subjective beliefs. The

particular versions of separability required in each model, such as Savage's (1954) sure-thing principle for expected utility, need not be added as separate conditions, but follow from the corresponding versions of likelihood consistency plus the other conditions.

Sections 2–4 focus on new intuitive concepts, measurement methods, and testable preference conditions. These sections avoid technical details, and are accessible to readers without mathematical background. Section 2 presents the basic intuition of the likelihood method. The following sections then apply the method to various models. Section 3 applies it to Bayesian beliefs in subjective expected utility and in probabilistic sophistication, and Section 4 applies it to non-Bayesian beliefs (or decision weights) in Choquet expected utility. We incorporate the rank dependence of Choquet expected utility through “ranks,” which capture the intuition of rank dependence in a more tractable manner than the commonly used comonotonicity. Section 5 formally states the generalized preference foundations of Savage's (1954) subjective expected utility, Machina and Schmeidler's (1992) probabilistic sophistication, and Gilboa's (1987) Choquet expected utility, adding a number of technical contributions.

Quantitative measurements and qualitative properties of decision weights are analyzed in Section 6. Section 7 shows how the likelihood method can give new interpretations to several other existing results on uncertainty and risk in the literature. Proofs are in the Appendixes. Lemmas B.11 and B.12 show that the likelihood consistency axioms are the duals of Köbberling and Wakker's (2003) tradeoff consistency axioms, when applied to events instead of outcomes. The underlying duality is the same as that used by Yaari (1987) when he applied de Finetti's bookmaking principle to probabilities, but it is now applied to events instead of probabilities. It will be illustrated in Figure B.1 in Appendix A. Our proofs are mostly based on techniques of Krantz et al. (1971), illustrating once more the power of these techniques.

## 2. THE LIKELIHOOD METHOD IN GENERAL

## 2.1. Preliminaries

$S$  denotes a set of *states of nature*. Exactly one state is true, and it is unknown which one.  $S$  can be finite or infinite. Events are subsets of  $S$ .  $\mathcal{C}$  denotes a set of consequences, or *outcomes*.  $\mathcal{C}$  can also be finite or infinite, and outcomes can be monetary prizes, health states, etc. *Acts* are state-contingent outcomes, i.e. they map states to outcomes; acts are assumed to take only finitely many values in this paper.  $(A_1:c_1, \dots, A_n:c_n)$  denotes the act assigning outcome  $c_j$  to event  $A_j$  (i.e., to each state  $s$  in  $A_j$ ) for each  $j$ . For outcome  $\alpha$ , event  $A$ , and act  $f$ ,  $\alpha_A f$  denotes the act resulting from  $f$  if all outcomes for event  $A$  are replaced by  $\alpha$ . The *preference relation*  $\succsim$  is a binary relation on the acts. The symbols  $\succ$ ,  $\sim$ ,  $\preccurlyeq$ , and  $\prec$  are defined as usual. Preference symbols also designate the preference relations over outcomes that are generated by constant acts in the usual way.

A function  $V$  *represents*  $\succsim$  if  $V$  maps acts to the real numbers and  $V(f) \geq V(g)$  if and only if  $f \succsim g$ . If there exists a representing function for  $\succsim$ , then  $\succsim$  is a *weak order*, which means that it is *complete* ( $f \succsim g$  or  $g \succsim f$  for all  $f, g$ ) and transitive. *Subjective expected utility (SEU)* holds if there exists a probability measure  $P$  on  $S$ , and a utility function  $U: \mathcal{C} \rightarrow \mathbb{R}$ , such that SEU represents preferences; the *SEU* of an act is the expectation under  $P$  of the utility of its outcomes.

## 2.2. Basic likelihood revelations of Savage–de Finetti

To prepare for the likelihood method, we give a detailed explanation of the Savage–de Finetti method for revealing likelihood orderings of events from gambles on events. See Figure 1.  $A$  and  $B$  designate events, such as rain tomorrow or Dow Jones going up tomorrow.  $A^c$  and  $B^c$  denote complementary events.  $\beta$  designates an outcome, i.e. a constant act  $(A:\beta, A^c:\beta) = (B:\beta, B^c:\beta)$ , illustrated by the first row below the lines in the matrix.  $\gamma$  (“good”) denotes an outcome better than  $\beta$  (“bad”), i.e.  $\gamma \succ \beta$ . An agent, whose status quo is

$$\begin{array}{cc|c|cc}
 A & A^c & & B & B^c \\
 \hline
 \beta & \beta & = & \beta & \beta \\
 \gamma & \beta & > & \gamma & \beta
 \end{array}$$

Figure 1. Savage – de Finetti basic likelihood measurement:  $A \succ_b B$ .

$\beta$ , can choose between gambling on  $A$  or on  $B$ . Gambling on  $A$ , depicted in the left half of the last row, means improving  $\beta$  to  $\gamma$  under  $A$ , and it designates the act  $(A:\gamma, A^c:\beta)$ . Gambling on  $B$ , depicted in the right half of the last row, means improving  $\beta$  to  $\gamma$  under  $B$ , and designates the act  $(B:\gamma, B^c:\beta)$ .

**DEFINITION 2.1.** Event  $A$  is *revealed more likely* than event  $B$  in a basic sense ( $A \succ_b B$ ) if there exist outcomes  $\gamma \succ \beta$  such that the configuration of Figure 1 holds. If such a configuration holds with  $\sim$  (or  $\prec$ ) instead of  $\succ$ , then we write  $A \sim_b B$  (or  $A \prec_b B$ ), and call  $A$  *revealed equally likely as (revealed less likely than)  $B$  in the basic sense*.

The subscript  $b$  serves to distinguish this measurement from nonbasic measurements defined later. Under SEU, gambling on  $A$  entails an improvement  $P(A)(U(\gamma) - U(\beta))$ , and gambling on  $B$  an improvement  $P(B)(U(\gamma) - U(\beta))$ . Revealed likelihood then measures probability orderings:

$$\begin{aligned}
 A \succ_b B & \text{ if and only if } P(A) > P(B); \\
 A \sim_b B & \text{ if and only if } P(A) = P(B); \\
 A \prec_b B & \text{ if and only if } P(A) < P(B).
 \end{aligned} \tag{2.1}$$

In general, if SEU does not hold, then likelihood revelations may run into contradictions. For example, some outcomes  $\gamma \succ \beta$  may reveal  $A \sim_b B$ , whereas other outcomes  $\gamma' \succ \beta'$  may reveal  $A \succ_b B$ . Such observed inconsistencies signal that the measurement method does not reveal likelihood as desired, and should be modified. The following condition excludes such contradictions.

**DEFINITION 2.2.** *Basic likelihood consistency* holds if  $[A \succ_b B \text{ and } A \sim_b B]$  for no events  $A, B$ .

The condition holds under SEU, as follows from Equation (2.1). Stated directly in terms of preferences, it requires that no events  $A, B$ , and outcomes  $\gamma > \beta$  and  $\gamma' > \beta'$  exist such that the preferences of Figures 1 and 2 hold simultaneously.

Similarly, contradictions such as  $[A >_b B \text{ and } A <_b B]$ , and contradictory revelations from bets against events ( $\gamma < \beta$ ) are to be avoided. These additional requirements plus basic likelihood consistency together amount to Savage's (1954) P4 consistency condition for likelihood orderings. In our main results, the additional requirements need not be imposed because they will always be implied by the other conditions. Similar observations hold for all other likelihood consistencies presented later.

### 2.3. The likelihood method

Figure 3 depicts the basic idea of the likelihood method introduced in this paper. The method generalizes the Savage–de Finetti likelihood measurements by not taking identical and constant, but instead equivalent and nonconstant, status quos as points of departure. Under the events compared,  $A$  and  $B$ , we maintain the original bad outcome  $\beta$ , to be replaced by a good outcome  $\gamma > \beta$ , as before. Under the complementary events  $A^c$  and  $B^c$  we now allow for general outcomes, described by  $f$  and  $g$ , which need not be constant on  $A^c$  or  $B^c$ . As before, we consider what is preferred more, gam-

$$\begin{array}{cc|cc} A & A^c & & B & B^c \\ \hline \beta' & \beta' & = & \beta' & \beta' \\ \gamma' & \beta' & \sim & \gamma' & \beta' \end{array}$$

Figure 2.  $A \sim_b B$ .

$$\begin{array}{cc|cc} A & A^c & & B & B^c \\ \hline \beta & f & \sim & \beta & g \\ \gamma & f & > & \gamma & g \end{array}$$

Figure 3. The likelihood method:  $A > B$ .

bling on  $A$  (the left improvement) or gambling on  $B$  (the right improvement).

DEFINITION 2.3. Event  $A$  is *revealed more likely* than event  $B$  ( $A \succ B$ ) if there exist outcomes  $\gamma \succ \beta$  and acts  $f, g$  such that the configuration of Figure 3 holds. If such a configuration holds with  $\sim$  (or  $\prec$ ) instead of  $\succ$ , then  $A \sim B$  (or  $A \prec B$ ), i.e.  $A$  is *revealed equally likely as (revealed less likely than)  $B$* .

Using the same symbols for preferences between acts and revealed orderings of events (and preferences over outcomes) will not lead to confusion.<sup>1</sup> We use the term *likelihood method* to refer to the general measurement that takes equivalent non-constant status quos, rather than the identical constant status quos of Savage and de Finetti, as points of departure, i.e. it refers to Figure 3 and its variations introduced later.

The intuition for the revealed likelihood relations is the same as for the basic relations. The former relations are, however, more prone to distorting interactions with what happens off  $A$  and  $B$ . We will examine such interactions in the following sections. As we will then see, for various decision models we can accept the measurements of the likelihood method only under special circumstances that preclude the interactions relevant to the decision model considered. Such acceptable measurements are indicated by a subscript, reflecting the particular decision model considered. Usually, when no confusion is likely to arise, the subscript indicating the relevant model may be dropped. One particular kind of measurement concerns the basic measurements indicated by subscript  $b$ , defined above, which restrict the likelihood method of Figure 3 to the case of  $f = g = \beta$ . Basic likelihood consistency will be satisfied by all decision models considered later.

It will always be easy to demonstrate that particular decision models exclude inconsistencies between the measurements acceptable under these models. It will also be demonstrated in the following sections that excluding such inconsistencies implies that the relevant decision models must hold, but proofs of these reversed implications are not elementary. We thus find that con-

sistencies of the relevant measurements provide preference foundations for the decision models considered.

### 3. THE LIKELIHOOD METHOD FOR ADDITIVE BELIEFS

This section applies the likelihood method to the measurement of additive, “probabilistic,” beliefs, first for SEU, then for probabilistic sophistication – i.e., without commitment to a decision model for given probabilities.

#### 3.1. *Subjective expected utility*

The most restrictive model considered in this paper is SEU. Under SEU, no interactions can occur between disjoint events, and no restrictions need to be imposed upon the measurements  $\succ$ ,  $\sim$ , and  $\prec$  of the likelihood method. Hence, these measurements can be used without subscripts for SEU. Under SEU, gambling on  $A$  in Figure 3 entails an improvement  $P(A)(U(\gamma) - U(\beta))$ , and gambling on  $B$  an improvement  $P(B)(U(\gamma) - U(\beta))$ , exactly as for the basic revealed likelihoods. Equation (2.1) can, therefore, be extended to nonbasic revealed likelihoods. Under SEU:

$$\begin{aligned} A \succ B & \text{ if and only if } P(A) > P(B); \\ A \sim B & \text{ if and only if } P(A) = P(B); \\ A \prec B & \text{ if and only if } P(A) < P(B). \end{aligned} \tag{3.1}$$

It follows that likelihood revelations never run into contradictions under SEU.

**DEFINITION 3.1.** *Likelihood consistency* holds if  $A \succ B$  and  $A \sim B$  for no events  $A, B$ .

Stated directly in terms of preferences, the condition amounts to:

$$\begin{aligned} \beta_A f \sim \beta_B g \quad & \& \quad \beta'_A f' \sim \beta'_B g' \quad & \& \\ \gamma_A f \sim \gamma_B g & & & & \\ \text{imply } \gamma'_A f' \sim \gamma'_B g' & & & & \end{aligned} \tag{3.2}$$

whenever  $\gamma > \beta$  and  $\gamma' > \beta'$ . The left two indifferences correspond to  $A \sim B$ , and the right two exclude  $A > B$  (and, symmetrically,  $A < B$ ).

Under common richness assumptions, likelihood consistency is not only necessary but also sufficient for the existence of probabilities for uncertain events, and, furthermore, utilities for outcomes, such that their SEU governs decisions (Theorem 5.3). Apparently, the condition implies not only enough of Savage's likelihood condition P4 but also enough of his sure-thing principle P2 to imply SEU. Indeed, the condition entails that likelihood revelations concerning  $A$  are to some extent independent ("separable") from what happens off  $A$ , in the spirit of P2. Likelihood consistency when restricted to the special case of  $A = B$  and  $f = g$  immediately implies a part of P2.

### 3.2. Examples

Before proceeding with specifications of the likelihood-method measurements, we give two examples where interactions between disjoint events lead to likelihood inconsistencies, demonstrating that modifications of the measurements are called for in such cases. Our examples concern Allais' and Ellsberg's paradoxes. These paradoxes are well known and, hence, serve well as a first illustration of our concepts. The examples prepare for the generalizations of expected utility introduced later. Both examples concern the special case where we compare the likelihood of an event  $A$  with itself ( $A = B$ ; in formal terms, the following examples violate irreflexivity of the strict revealed-more-likely-than relation  $\succ$ ). For many agents, events  $A$ ,  $C'$ , and  $D'$  can be found such that the preferences in Figure 4(a) hold.

Allais considered the case of known probabilities, with  $P(A) = 0.89$ ,  $P(C') = 0.01$ , and  $P(D') = 0.10$ . The phenomenon also pertains to unknown probabilities (Gonzalez and Wu, 1999; MacCrimmon and Larsson 1979; Tversky and Kahneman, 1992). For agents who exhibit the preferences in Figure 4(a), we can shift an appropriate small part of  $D'$  (e.g. its intersection with the result of some coin-flips) to  $C'$ , leading

to the smaller  $D$  and larger  $C$ , with the first preference turned into an indifference in Figure 4(b) and the second preference intensified.

From Figure 4(b) we obtain the counterintuitive  $A \succ A$ . Together with the trivial  $A \sim A$  (implied by  $\mathbf{0}_A f \sim \mathbf{0}_A f$  and  $\mathbf{M}_A f \sim \mathbf{M}_A f$ ),  $A \succ A$  generates a violation of likelihood consistency (set  $B = A$ ). SEU is violated.

We next consider the three-color Ellsberg paradox. Imagine an urn with 30 red ( $R'$ ) balls, and 60 black ( $B'$ ) and yellow ( $Y$ ) balls in unknown proportion. The preferences in Figure 5(a) are commonly found. For later purposes, we rank-order events from best (left) to worst (right) outcomes. We can move a small subevent from  $R'$  to  $B'$ , leading to a smaller  $R$  and a larger  $B$ , such that the equivalence and preference in Figure 5(b) result.

Figure 5(b) implies the counterintuitive  $Y \succ Y$ . Together with the trivial  $Y \sim Y$ , it generates a violation of likelihood consistency. SEU is violated again.

### 3.3. The likelihood method for probabilistic sophistication

The violations of likelihood consistency in the above examples signal that, to still obtain meaningful likelihood comparisons

$$\begin{array}{c}
 \text{(a)} \\
 \begin{array}{ccccccc}
 \frac{A}{\mathbf{0}} & \frac{C'}{M} & \frac{D'}{M} & < & \frac{A}{\mathbf{0}} & \frac{C'}{0} & \frac{D'}{5M} \\
 \frac{M}{M} & \frac{M}{M} & \frac{M}{M} & > & \frac{M}{M} & \frac{0}{0} & \frac{5M}{5M}
 \end{array} \\
 \text{(b)} \\
 \begin{array}{ccccccc}
 \frac{A}{\mathbf{0}} & \frac{C}{M} & \frac{D}{M} & \sim & \frac{A}{\mathbf{0}} & \frac{C}{0} & \frac{D}{5M} \\
 \frac{M}{M} & \frac{M}{M} & \frac{M}{M} & > & \frac{M}{M} & \frac{0}{0} & \frac{5M}{5M} \\
 & \underbrace{\hspace{2cm}}_{f \text{ on } A^c} & & & \underbrace{\hspace{2cm}}_{g \text{ on } A^c} & & 
 \end{array}
 \end{array}$$

Figure 4. (a) The Allais paradox.  $M$  denotes  $\$10^6$ . (b) Allais paradox modified.  $A \succ A$  by setting  $\beta = 0, \gamma = M$ .

$$\begin{array}{c}
 \text{(a)} \\
 \begin{array}{ccccccc}
 \frac{B'}{M} & \frac{Y}{\mathbf{0}} & \frac{R'}{0} & < & \frac{R'}{M} & \frac{Y}{\mathbf{0}} & \frac{B'}{0} \\
 \frac{M}{M} & \frac{M}{M} & \frac{0}{0} & > & \frac{M}{M} & \frac{M}{M} & \frac{0}{0}
 \end{array} \\
 \text{(b)} \\
 \begin{array}{ccccccc}
 \frac{B}{M} & \frac{Y}{\mathbf{0}} & \frac{R}{0} & \sim & \frac{R}{M} & \frac{Y}{\mathbf{0}} & \frac{B}{0} \\
 \frac{M}{M} & \frac{M}{M} & \frac{0}{0} & > & \frac{M}{M} & \frac{M}{M} & \frac{0}{0}
 \end{array} \\
 \text{Figure 5. (a) The Ellsberg paradox. } M \text{ denotes } \$10^6. \text{ (b) Ellsberg Paradox modified. } Y \succ Y \text{ by setting } \beta = 0, \gamma = M.
 \end{array}$$

and to avoid contradictions for agents as in the examples, we have to impose restrictions on our measurements. Probabilistic sophistication, the model first axiomatized by Machina and Schmeidler (1992), suggests a way to do so. In this model, likelihood is still governed by Bayesian probabilities but otherwise decisions can deviate from expected utility. It leads to the following definition.

**DEFINITION 3.2.** Event  $A$  is *revealed more likely* than event  $B$  in a *probabilistically sophisticated sense* ( $A \succ_{ps} B$ ) if the configuration of Figure 3 holds whenever  $f$  and  $g$  yield only the outcomes  $\gamma$  and  $\beta$ . If such a configuration holds with  $\sim$  (or  $<$ ) instead of  $\succ$ , then  $A \sim_{ps} B$  (or  $A <_{ps} B$ ).

Remember that  $f$  and  $g$  need not be identical or constant on  $A^c$  and  $B^c$ . Under the restriction for  $f$  and  $g$  in the above definition, the decision is driven entirely by beliefs, namely beliefs about the likelihood of receiving  $\gamma$  instead of  $\beta$ . The decision then cannot be affected by an interaction with outcomes under other events such as those occurring in the Allais paradox. This provides the intuition for the following axiom. In Section 5 we will see that  $ps$ -orderings agree with probability orderings for all models based on probabilities.

**DEFINITION 3.3.** *Ps-likelihood consistency* holds if  $A \succ_{ps} B$  and  $A \sim_{ps} B$  for no events  $A, B$ .

Under common richness conditions, this consistency implies the existence of a probability measure  $P$  such that Equation (3.1) holds for  $\succ_{ps}$ ,  $\sim_{ps}$ , and  $<_{ps}$  instead of  $\succ$ ,  $\sim$ , and  $<$  irrespective of the decision model assumed for given probabilities. Probabilistic sophistication further requires that decisions between acts with more than two outcomes are based only on the probability distributions generated over the outcomes. As demonstrated by Sarin and Wakker (2000), this further requirement is captured by *equivalence-dominance* (another implication of likelihood consistency; see Lemma B.4). It requires that

$$f \sim g \text{ whenever } \{s : f(s) \geq \alpha\} \sim_b \{s : g(s) \geq \alpha\} \\ \text{for all outcomes } \alpha. \quad (3.3)$$

The antecedent  $\sim_b$  relationships entail that each “goodnews” event of receiving something not below  $\alpha$  is equally likely in a basic sense for both acts. In such a case, the indifference can again be based exclusively on considerations about beliefs, whatever the outcomes or the decision model are. The condition does not require those beliefs to be Bayesian and is, for instance, satisfied by all Choquet expected utility maximizers, so that it can accommodate all nonadditive measures. It can, accordingly, accommodate all Ellsberg-paradox behavior, such as in Figure 5 and in Example 5.8(ii) hereafter, which further suggests that the condition is not very restrictive. Theorem 5.5 will show that, under common richness assumptions, the two implications of (unrestricted) likelihood consistency just introduced are not only necessary but also sufficient for probabilistic sophistication.

As is well known, probabilistic sophistication can accommodate the violation of SEU revealed by the Allais paradox, but not the violation revealed by the Ellsberg paradox. In Figure 5(b) we do indeed have  $Y \succ_{ps} Y$ . This, together with the trivial  $Y \sim_{ps} Y$ , violates ps-likelihood consistency. The following section considers a modification of likelihood consistency that can accommodate not only the Allais paradox but also the Ellsberg paradox.

#### 4. THE LIKELIHOOD METHOD FOR NONADDITIVE BELIEFS

The modification of SEU in this section incorporates rank-dependence as introduced by Quiggin (1981) and Schmeidler (1989). Rank-dependent theories generalize expected utility through the replacement of probabilities by decision weights, suggesting nonadditive beliefs. These decision weights can depend on the ranking of events as regards the favorability of the outcomes received under the act being considered. section 4.1 introduces “ranks” of events. section 4.2 then presents measurements and preference conditions.

#### 4.1. Ranks of events

Rank-dependence is commonly formalized through comonotonicity, which concerns sets of acts that generate the same rank-ordering of the state space regarding favorability of outcomes. We will use a simpler approach, where only ranks of events are considered, and it is not necessary to specify the rank-ordering of the complete state space. The rank of an event depends on the act being considered. By using ranks, we can avoid some mathematical difficulties of comonotonicity assumptions that complicate their analysis and testability (see Appendix A).

For a given act, we define the *rank*  $R \subset A^c$  of an event  $A$  yielding outcome  $\alpha$  as a “dominating event”: all outcomes under  $R$  are weakly preferred to  $\alpha$ , and all outcomes off  $A$  and  $R$ , denoted  $L = (A \cup R)^c$  henceforth, are weakly less preferred than  $\alpha$ . We leave flexibility regarding outcomes off  $A$  that are equivalent to  $\alpha$ , and that can be allocated to  $R$  or  $L$  as is desirable. This flexibility can be convenient when using the same ranks for different acts. For example, in the Allais paradox in Figure 4(a) (or Figure 4(b)), the rank of  $A$  can be taken to be the same  $C' \cup D' (= C \cup D)$  for the two left acts, and the same  $D'$  (or  $D$ ) for the two right acts. This will simplify the analysis hereafter.

In agreement with linguistic conventions, the smaller (in the sense of set-inclusion) the rank of  $A$ , the more favorable  $A$  is, and the better it is ranked. Under rank-dependence, a pessimist can attach more weight to an event when it is associated with relatively bad outcomes (a bad, large rank) than with relatively good outcomes. Such pessimism can explain the Ellsberg paradox (see later), and the Allais paradox. In Figure 4(a) (and similarly in Figure 4(b)), event  $A$  is associated with the worst outcome for the left acts and with the middle outcomes for the right acts.  $A$  will, consequently, be more important for the left acts, and the improvement from 0 to  $M$  has more effect for the left acts under pessimism.

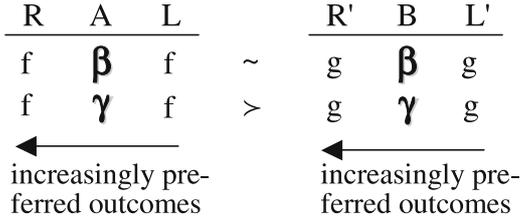


Figure 6.  $\gamma > \beta$ . Rank-dependent revealed likelihood:  $A^R > B^{R'}$ .

#### 4.2. Rank-dependent likelihood measurement

For likelihood revelations, we specify the rank through superscripts, and write  $A^R$  to indicate event  $A$  when its rank is  $R$ . We call  $A^R$  a ranked event. The notation  $\alpha_{A^R}f$  designates the act resulting from  $f$  if all outcomes for event  $A$  are replaced by  $\alpha$ , as did  $\alpha_A f$ , but specifies that  $R$  is the rank of  $A$  in  $\alpha_A f$ . That is, all outcomes under  $R$  are weakly preferred to  $\alpha$ , and those under  $(R \cup A)^c$  are weakly less preferred than  $\alpha$ . Because the likelihood relations below concern ranked events instead of events, no confusion can arise with other likelihood relations, and we do not add a subscript to the relations.

**DEFINITION 4.1.** The ranked event  $A^R$  is *revealed more likely* than the ranked event  $B^{R'}$  ( $A^R > B^{R'}$ ) if the configuration of Figure 3 holds where  $R$  is the rank of event  $A$  for both left acts and  $R'$  the rank of event  $B$  for both right acts (see Figure 6). If such a configuration holds with  $\sim$  (or  $<$ ) instead of  $>$ , then  $A^R \sim B^{R'}$  (or  $A^R < B^{R'}$ ).

The ranks of events  $A$  and  $B$  in the above definition imply that

- all  $f$ -outcomes ( $g$ -outcomes) under event  $R$  ( $R'$ ) are weakly preferred to  $\gamma$  and, hence, to  $\beta$ , and
- all  $f$ -outcomes ( $g$ -outcomes) off  $A$  ( $B$ ) and  $R$  ( $R'$ ), i.e. under event  $L$  ( $L'$ ), are weakly less preferred than  $\beta$  and, hence, than  $\gamma$ .

We think that the rank-dependent models are best understood if the object of revealed likelihood is not taken to be an

event, but a ranked event. For example, many disagreements about the meaning of null events and about updating rules in rank-dependent models can be resolved if ranked events, rather than events, are considered. These disagreements center around different ranks that are assumed for the events considered (Cohen et al., 2000).

There is an ongoing debate about whether capacities and decision weights reflect generalized beliefs, or comprise other decision components besides beliefs. The adjective “revealed” emphasizes that our likelihood relations are choice-based, and that psychological interpretations are yet to be settled.

To see that rank-dependence can explain not only the Allais paradox but also the Ellsberg paradox, consider Figure 5(b). This figure reveals  $Y^B \succ Y^R$ , which entails no contradiction. For rank  $B$ ,  $Y$  receives extra weight because it is exactly what changes an unknown probability of a good outcome into a known probability. For rank  $R$ ,  $Y$  receives less weight because it changes known probability of a good outcome into a larger, but now unknown, probability. The improvement under  $Y$  will, consequently, be more important for the left acts than for the right ones under aversion to unknown probabilities.

Section 5 will give a formal definition of Choquet expected utility, using decision weights instead of subjective probabilities. Observation 5.6 will then show that Figure 6 reveals an ordering of decision weights of the ranked events.

**DEFINITION 4.2.** *Rank-dependent likelihood consistency* holds if  $[A^R \sim B^{R'} \text{ and } A^R \succ B^{R'}]$  for no ranked events  $A^R, B^{R'}$ .

That is, the condition precludes contradictions in the measurements just defined. Stated directly in terms of preferences, the condition amounts to

$$\begin{aligned} \beta_{A^R} f \sim \beta_{B^{R'}} g \ \& \ \beta'_{A^R} f' \sim \beta'_{B^{R'}} g' \ \& \\ \gamma_{A^R} f \sim \gamma_{B^{R'}} g & \\ \text{imply } \gamma'_{A^R} f' \sim \gamma'_{B^{R'}} g' & \end{aligned} \quad (4.1)$$

The left two indifferences correspond to  $A^R \sim B^{R'}$ , and the right two exclude  $A^R \succ B^{R'}$  (and, symmetrically,  $A^R \prec B^{R'}$ ).

From the above informal claim that rank-dependent likelihood revelations reveal orderings of decision weights, it readily follows that Choquet expected utility implies rank-dependent likelihood consistency. We will see later that the implication can be reversed under common richness assumptions, and rank-dependent likelihood consistency is not only necessary, but also sufficient for Choquet expected utility. The existence of utility for CEU apparently follows as a by-product of the consistency of decision weights.

## 5. RICHNESS CONDITIONS AND FORMAL PREFERENCE FOUNDATIONS

This section presents formal results. We use one more intuitive condition, namely, *monotonicity*:  $f \succcurlyeq g$  whenever  $f(s) \succcurlyeq g(s)$  for all  $s$ . Strict versions of monotonicity need not be imposed because they will follow from (weak) monotonicity plus likelihood consistency. The following two subsections present technical richness conditions, the subsequent subsections the preference foundations.

### 5.1. *Mosaics to generalize Savage's ( $\sigma$ -)algebras of events*

A collection  $\mathcal{A}$  is given of subsets of  $S$  called *events*. Acts are *measurable*, which means that each of the finitely many outcomes of an act has an event as its inverse. The extension of our results to infinite-valued acts remains a topic for future research (Gilboa, 1987; Savage, 1954; Wakker, 1993). We assume that  $\mathcal{A}$  is a *mosaic*, i.e. it is closed under complement taking, it contains  $S$  (and, hence,  $\emptyset$ ), and for each finite partition  $(A_1, \dots, A_n)$  of  $S$  consisting of events, all unions of the  $A_j$ 's are contained in  $\mathcal{A}$  too. In decision theory, collections of events are usually assumed to be, more restrictively, an *algebra* (closed under complement taking and finite unions, and containing  $\emptyset$  and  $S$ ) or a  *$\sigma$ -algebra* (also closed under

countable unions), following traditions in probability theory. In the act-notation  $(A_1 : x_1, \dots, A_n : x_n)$ , it is implicitly understood that all  $A_j$ 's are events (including the case where  $x_i = x_j$  for some  $i = j$ ).

There has been an interest in more general collections of events in quantum-mechanics, where we can, for instance, observe the location or the momentum of a particle, but not both. We then do not assume intersections as in algebras and require closedness only with respect to complements and finite disjoint unions (*Dynkin systems*, or QM-algebras in Krantz et al., 1971, Section 5.4). A similar interest recently arose in decision theory in the study of ambiguity, and here again the set of unambiguous events need not be intersection-closed. This led Zhang (1999) and Epstein and Zhang (2001) to consider closedness with respect to complements and countable disjoint unions ( $\lambda$ -systems or  $\sigma$ -*Dynkin systems*). Kopylov (2004) introduced the, more general, mosaics, again for the study of unambiguous events.

Our primary interest in generalized domains of events derives from the greater flexibility that they allow in applications of decision models in general, by resolving a criticism of algebras of events expressed by Luce and others. In Savage's classical approach with ( $\sigma$ -)algebras of events, for two unrelated sources of uncertainty, all joint uncertainties and joint distributions also have to be specified, even if irrelevant to the actual decision problems considered (Krantz et al., 1971, p. 373; Luce, 2000, Section 1.1.6.1). In our generalized approach, we need not consider such joint distributions (Examples 5.4(v) and 5.8(ii), where the mosaic is also useful for the ambiguous events). This makes Dynkin systems and mosaics more suited for applications than the commonly used ( $\sigma$ -)algebras.

There is much interest today in comparing attitudes towards risk and ambiguity for different, unrelated, sources of uncertainty (Ellsberg, 1961; Chew and Sagi, 2004; Ergin and Gul, 2004, who used the term issue instead of source; Fox and Tversky, 1995). Then the flexibility of not having to consider all intersections of the risky and ambiguous uncertainties, as is possible in our approach, is desirable. Luce's

(2000) “chance experiments” generalized Savage’s ( $\sigma$ )-algebras in similar ways. de Finetti (1937), and Gilboa and Schmeidler (2004) in a different decision model (with variable “contexts”), also allowed for event sets that are not intersection-closed.

The required richness of events in our analysis will be implied mainly by solvability. Solvability implies, basically, that the mosaic contains subalgebras that are rich enough to allow for the required constructions of scales. Thus, our mosaics can be thought of as unions of separate algebras. Examples 5.4(v) and 5.8(ii) illustrate unions of two separate algebras.

Event  $A$  is *nonnull* if  $(R : \gamma, A : \gamma, L : \beta) \succ (R : \gamma, A : \gamma, L : \beta)$  for some  $\gamma \succ \beta$  and  $R, L$ , and it is *null* otherwise. *Nondegeneracy* holds if  $(A : \gamma, A^c : \gamma) \succ (A : \gamma, A^c : \beta) \succ (A : \beta, A^c : \beta)$  for some  $A, \beta, \gamma$ . The condition amounts to the existence of two disjoint nonnull events, which also implies the existence of two non-equivalent outcomes. We now summarize the assumptions made so far.

**STRUCTURAL ASSUMPTION 5.1.** The preference relation  $\succsim$  is a binary relation over the set of acts, which consists of all  $\mathcal{A}$ -measurable finite-valued maps from the state space  $S$  to the outcome space  $\mathcal{C}$ , with  $\mathcal{A}$  a mosaic on  $S$ . Nondegeneracy holds.

In general, for an event  $A$  and an act  $f = (E_1 : x_1, \dots, E_n : x_n)$ ,  $f$  with  $f(s)$  replaced by  $\alpha$  on  $A$  need not be an act if  $A$  is not a union of  $E_j$ s. Whenever we use the notation  $\alpha_A f$ , or  $\alpha_{A^R} f$ , it is implicitly understood that  $A$  is an event and  $\alpha_A f$  is an act, which means that it satisfies measurability with respect to the mosaic. Probability measures satisfy the additivity requirement  $P(A \cup B) = P(A) + P(B)$  for disjoint  $A, B$  only if  $A \cup B$  is an event, which on mosaics is a substantially weaker requirement than on algebras. Desirable properties such as monotonicity will nevertheless follow in all our main results, primarily because of the richness provided by solvability defined later. Countable additivity of probability measures, and the corresponding set-continuity of capacities, can be characterized as in Section 4.1 of Wakker (1993) in all results below.

### 5.2. Generalizing Savage's non-atomicity P6

The following condition is the only nonnecessary condition in our theorems. It generalizes Gilboa's nonatomicity condition P6\* by means of techniques from Luce (1967; see also Krantz et al., 1971), so as to allow for atoms.

DEFINITION 5.2. *Solvability* holds if

$$\beta_A f \prec g \prec \gamma_A f \text{ implies } \gamma_G \beta_B f \sim g$$

for some events  $G$  and  $B$  partitioning event  $A$  whenever no outcome of  $f$  or  $g$  is strictly between  $\beta$  and  $\gamma$  in preference.

If improving  $\beta$  to  $\gamma$  on all of  $A$  is too much to give indifference, then we can improve on a properly chosen subpart  $G$  of  $A$  to obtain indifference after all. The restriction of no outcomes of  $f$  and  $g$  between  $\beta$  and  $\gamma$  could be dropped for atomless models, and serves to allow for finite state spaces. Solvability is reminiscent of the intermediate value property of real-valued continuous functions. Solvability entails, roughly, that the mosaic of events is at least as refined as the outcome space. If the outcome space is a continuum, then so must the mosaic of events be, but if the outcome space is finite then the mosaic and state space may be finite too. In the presence of the other conditions, solvability is less restrictive than Savage's (1954) atomlessness condition P6. We prefer it primarily because it is simpler.

Another part of Savage's P6 is captured by the *Archimedean axiom*, which requires that no infinite sequence of disjoint nonnull events  $E_1, E_2, \dots$  exist with  $E_j \sim E_{j-1}$  for all  $j$ . This condition is necessary for all real-valued probability representations because infinitely many disjoint equally likely events with positive probability cannot exist.

### 5.3. Subjective expected utility

The following theorem characterizes SEU for three or more nonequivalent outcomes. Theorem 5.5 will deal with SEU for two nonequivalent outcomes.

**THEOREM 5.3.** *Let the Structural Assumption 5.1 hold, let there exist at least three nonequivalent outcomes, and assume solvability. Then the following two statements are equivalent:*

- (i) *Subjective expected utility holds.*
- (ii) *The following conditions are satisfied:*
  - (a) *weak ordering;*
  - (b) *monotonicity;*
  - (c) *the Archimedean axiom;*
  - (d) *likelihood consistency.*

*In Statement (i), the probability measure is unique and utility is unique up to unit and origin.*

Uniqueness up to unit and origin entails that any real number can be added to utility, and that it can be multiplied by any positive real number. The logical relations between our result and Savage's (1954) are as follows. Savage's axioms require infinite state spaces and a  $\sigma$ -algebra of events (Savage, 1954, p. 43). They imply solvability, the only nonnecessary condition in our theorem. Therefore, a preference relation satisfying Savage's axioms also satisfies ours. In the following Examples 5.4.(ii)–(v) our assumptions are satisfied but those of Savage are not, with his condition P6 and its implication of “convex-rangedness” violated. In this sense, our model is more general than Savage's.

**EXAMPLE 5.4.** Let the outcome set  $\mathcal{C}$  be  $\{0, 1, 2\}$ , let SEU hold, and let the utility  $U$  be the identity. Solvability and, hence, all conditions of Theorem 5.3 are satisfied in the following cases, as will be proved in Appendix D.

- (i) **CONVEX-RANGEDNESS.**  $P$  is any *convex-ranged* probability, i.e. for all  $0 \leq \mu < P(A)$  there exists  $B \subset A$  with  $P(B) = \mu$ .  $P$  is convex-ranged whenever  $P$  is countably additive and atomless and  $\mathcal{A}$  is a  $\sigma$ -algebra, and also whenever  $P$  satisfies Savage's (1954) axioms.
- (ii) **A COUNTABLE STATE SPACE WITH NO ATOMS AND NO CONVEX-RANGEDNESS.**  $S = [0, 1) \cap \mathbb{Q}$ ,  $\mathcal{A}$  contains all finite

unions of intervals  $[a, b)$  with  $a$  and  $b$  rational, and  $P$  is the Lebesgue measure, determined by  $P[a, b) = b - a$  for all  $[a, b)$ . All probabilities are rational-valued, which obviously excludes convex-rangedness.

- (iii) A FINITE STATE SPACE (WITH ATOMS).  $S = \{s_1, \dots, s_n\}$ ,  $\mathcal{A}$  contains all subsets of  $S$ , and  $P(s_j) = 1/n$  for all  $j$ . (More general outcome sets will be discussed following Observation C.1.)
- (iv) NO ATOMS, NO CONVEX-RANGEDNESS, AND NO UNIFORM PARTITIONS OF  $S$ .  $S = \{s \in [0, 1) : s = a + b\sqrt{2} \text{ for integers } a \text{ and } b\}$ . This set is dense in  $[0, 1)$ .  $\mathcal{A}$  contains all finite unions of intervals  $[s, t) \cap S$  with  $s, t \in S$ , and  $P$  is the Lebesgue measure. There exists no event with probability 0.5, or any other rational probability  $k/\ell$  other than 0 or 1. Hence, there exist no uniform partitions.
- (v) DIFFERENT SOURCES OF UNCERTAINTY WITHOUT JOINT DISTRIBUTIONS.  $S = [0, 1] \times [0, 1]$ , where  $(s_1, s_2)$  describes the delay of a train ( $s_1$ ) and the delay of a bus ( $s_2$ ).  $\mathcal{A}$  contains all sets of the form  $A = A_1 \times [0, 1]$  and  $[0, 1] \times A_2$  for Borel-measurable sets  $A_1$  and  $A_2$ . With  $\lambda$  the Lebesgue-measure, we have  $P(A_1 \times [0, 1]) = \lambda(A_1)$  and  $P([0, 1] \times A_2) = \lambda(A_2)$ . Luce (2000, Section 1.1.6.1) used such an example to criticize Savage's (1954) setup, which requires that probabilities be specified for all combinations of bus and train delays. For mosaics of events, such redundant specifications are not needed.

The finite state spaces in Example 5.4 (iii) do not require richness of outcomes, contrary to the finite-state models of Blackorby et al. (1977), Chew and Karni (1994), Grodal (1978), Gul (1992), Köbberling and Wakker (2003), or Wakker (1989, Theorem IV.2.7). As will appear from the proof of Theorem 5.3, our approach can only handle finite state spaces with equally likely states. An alternative preference foundation for such models, with solvability for outcomes, is in Davidson and Suppes (1956).

#### 5.4. Probabilistic sophistication

*Probabilistic sophistication* holds if there exists a probability measure  $P$  on  $S$  such that preferences over acts satisfy *stochastic dominance*:  $f \succcurlyeq g$  whenever  $P\{s : f(s) \succcurlyeq \alpha\} \geq P\{s : g(s) \succcurlyeq \alpha\}$  for all outcomes  $\alpha$ . Stochastic dominance implies that two acts that generate the same probability distribution over outcomes are equivalent, so that preference depends only on the probability distributions generated over the outcomes. Sometimes the further restriction is imposed on probabilistic sophistication that also a general quantitative functional  $V$  should exist that represents preference. This extra assumption can be characterized by the extra requirement in Statement (ii) of Theorem 5.5 that a countable “order-dense” subset of acts exist (Krantz et al., 1971, Theorem 2.2).

The following theorem allows for two outcomes. Then SEU and probabilistic sophistication coincide, and both models amount to the problem of qualitative probability: finding a probability measure that represents a more-probable-than relation on events (Fishburn, 1986). To allow for decision models more general than SEU, we have to weaken the Archimedean axiom. We will impose it only on  $\sim_b$  instead of on  $\sim$ : The *ps-Archimedean axiom* holds if there exists no infinite sequence of disjoint nonnull events  $E_j$  with  $E_j \sim_b E_{j-1}$  for all  $j$ . For the case of only two equivalence classes of outcomes, equivalence-dominance in (e) in Statement (ii) below can be dropped.

**THEOREM 5.5.** *Let the Structural Assumption 5.1 hold, and assume solvability. Then the following two statements are equivalent:*

- (i) *Probabilistic sophistication holds.*
- (ii) *The following conditions are satisfied:*
  - (a) *weak ordering;*
  - (b) *monotonicity;*
  - (c) *the ps-Archimedean axiom;*
  - (d) *ps-likelihood consistency;*
  - (e) *equivalence-dominance.*

*In Statement (i), the probability measure is unique.*

Theorem 5.5 does not require  $\sigma$ -algebras of events and convex-rangedness, as in Chew and Sagi (2004), Epstein and Le Breton (1993), Grant (1995), Machina and Schmeidler (1992), and Sarin and Wakker (2000). Instead, it allows for general mosaics of events and for some atomless state spaces that are not convex-ranged (e.g., Example 5.4). Kopylov (2004) allowed for mosaics, but still required convex-rangedness. Like Chew and Sagi (2004), we further generalize the other works by allowing for particular finite state spaces and atoms. Chew and Sagi (2004) and Grant (1995) are more general in allowing for violations of first stochastic dominance. Compared with Theorem 5.2 in Epstein and Zhang (2001), Theorem 5.5 is more general in allowing for particular finite state spaces and atomless non-convex-ranged measures, in not requiring closedness of  $\mathcal{A}$  under countable (or finite) disjoint unions, and in not requiring monotone continuity (cf. Epstein and Zhang, footnote 17).

### 5.5. Choquet expected utility and rank-dependence

A *capacity*  $W$  maps events to  $[0,1]$  with  $W(\emptyset)=0$ ,  $W(S)=1$ , and  $W(A) \geq W(B)$  whenever  $A \supset B$ . For the definition of the rank  $R$  of event  $A$ , given an act yielding outcome  $\alpha$  under  $A$ , states of nature off  $A$  yielding outcomes equivalent to  $\alpha$  can be allocated to  $R$  as well as to  $L=(A \cup R)^c$ , and in this sense the rank  $R$  is not unique. It is well known for rank-dependent theories that it is immaterial for the evaluation of acts which of the possible events  $R$  is chosen as rank in such a case.  $A^R$  is a *ranked event* if  $R$  is the rank of  $A$ ,  $A \cap R$  is empty, and  $A$ ,  $R$ , and  $A \cup R$  are events. For a ranked event  $A^R$ , the *decision weight*  $\pi(A^R)$  is  $W(A \cup R) - W(R)$ . *Choquet expected utility (CEU)* holds if there exists a capacity  $W$  on  $S$ , and a utility function  $U: \mathcal{E} \rightarrow \mathbb{R}$ , such that Choquet expected utility represents preferences. Here the *Choquet expected utility* of an act  $(A_1 : c_1, \dots, A_n : c_n)$ , with  $c_1 \succ \dots \succ c_n$ , is the *Choquet integral* of  $U \circ f$  under  $W$ . It is  $\sum_{j=1}^n \pi(A_j^{R_j}) U(c_j)$ , with ranks  $R_j = A_1 \cup \dots \cup A_{j-1}$  ( $= \emptyset$  if  $j=1$ ).

Observation 5.6, stated informally in Sarin and Wakker (1998, Section 5), follows from substitution.

**OBSERVATION 5.6.** Under Choquet expected utility:

If  $A^R \succ B^{R'}$ , then  $\pi(A^R) \succ \pi(B^{R'})$ , i.e.  $W(A \cup R) - W(R) > W(B \cup R') - W(R')$ ;

If  $A^R \sim B^{R'}$ , then  $\pi(A^R) = \pi(B^{R'})$ , i.e.  $W(A \cup R) - W(R) = W(B \cup R') - W(R')$ .

To allow for rank dependence, we have to weaken the Archimedean axiom, and impose it on ranked events instead of events. A ranked event  $A^R$  is *nonnull* if  $(R:\gamma, A:\gamma, L:\beta) \succ (R:\gamma, A:\beta, L:\beta)$  for some  $\gamma > \beta$ , and *null* otherwise. This non-nullness obviously implies that  $A$  is nonnull as defined before, but specifies the event  $R$ . In general it may happen that  $A^R$  is nonnull but, for a different rank  $R'$ ,  $A^{R'}$  is null, reflecting the dependence of revealed likelihood on ranks. Under the *rank-dependent Archimedean axiom*, there exist no infinite sequences of disjoint events  $E_0, E_1, \dots$  such that  $E_{j+1}^{E_j \cup \dots \cup E_0} \sim E_j^{E_{j-1} \cup \dots \cup E_0}$  for all  $j \geq 1$ , with all these ranked events nonnull. The existence of such a sequence would imply, with  $\varepsilon > 0$  the decision weight of each ranked event, that  $W(E_j \cup \dots \cup E_0) = W(E_0) + j\varepsilon$  for each  $j$ , implying that  $W(S) = \infty$ , which cannot be. The following theorem extends Theorem 5.3 to the rank-dependent case.

**THEOREM 5.7.** *Let the Structural Assumption 5.1 hold, let there exist at least three nonequivalent outcomes, and assume solvability. Then the following two statements are equivalent:*

- (i) *Choquet expected utility holds.*
- (ii) *The following conditions are satisfied:*
  - (a) *weak ordering;*
  - (b) *monotonicity;*
  - (c) *the rank-dependent Archimedean axiom;*
  - (d) *rank-dependent likelihood consistency.*

*In Statement (i), the capacity is unique and utility is unique up to unit and origin.*

Observation 5.6 clarifies why the intuitive axioms of Theorem 5.7 can be so simple. The observation shows that rank-dependent likelihood consistency concerns consistency of *differences* of capacities, which, as Theorem 5.7 shows, is enough to imply a cardinal model. Because we can restrict attention to capacity differences of nested events, with one event a subset of the other, these differences always reflect decision weights, i.e. willingness to bet. This shows how likelihood consistency, an ordinal-like consistency for willingness to bet, implies a cardinal representation. Likelihood consistency automatically involves set-inclusion, a structural relation that is available on the event space anyhow. Similar observations apply to SEU, with probability-differences of nested events equal to the probabilities of single events: namely, the difference-events. A similar use of the set-theoretic structure on the state space to simplify an intuitive axiom is in Chew and Sagi (2004).

In “outcome-oriented” approaches that assume rich outcome spaces, cardinality is usually obtained through an extra operation on the outcomes, such as addition in Chateaneuf (1991) and de Finetti (1937), and mixing in Anscombe and Aumann (1963), Gilboa and Schmeidler (1989), and Schmeidler (1989). Such operations underly tradeoff-consistency axioms that implicitly endogenize equalities of utility differences (Köbberling and Wakker 2003; Nehring, 2001; Tversky and Kahneman, 1992; Wakker, 1989), and bisymmetry axioms that implicitly endogenize event-weighted mixtures of outcomes (Chew and Karni, 1994; Ghirardato and Marinacci, 2001; Gul, 1992; Planzogl, 1968). Ghirardato et al. (2003) explicitly formalized an endogenous midpoint operation, and through limits of infinitely many such operations defined a general mixture operation. Lemma D.2 and its proof in Appendix D provide an illustration of the way in which, in our “uncertainty-oriented” approach (i.e., with a rich state space), cardinality follows from the interaction of our intuitive axiom with set-union. Thus, we avoid exogenous or endogenized operations. This explains why our axioms are simpler than those in outcome-oriented approaches.

Theorem 5.7 generalizes Kopylov's (2004) characterization of expected utility to rank-dependence (in addition to generalizing the convex-rangedness). It generalizes Gilboa's (1987) preference foundation similar to the way in which Theorem 5.3 generalized Savage's (1954). In particular, Gilboa's result is generalized to mosaics of events. Observation A.4 shows that Gilboa's axioms directly imply ours for finite-valued acts, so that his theorem is an immediate corollary of ours for such acts. The reversed implication does not hold. This can be inferred from Examples 5.4(ii)–(v), which satisfy our axioms but not those of Gilboa (1987). These examples all concerned SEU. We now add similar examples where expected utility is violated.

**EXAMPLE 5.8.** In the following examples, solvability and all other conditions of Theorem 5.7 are satisfied (see Appendix D).

- (i) **CEU WITH ATOMS.**  $\mathcal{C} = \{0, 1, 2\}$ , and  $U$  is the identity. CEU holds,  $S = \{s_1, \dots, s_n\}$ ,  $\mathcal{A}$  contains all subsets of  $S$ ,  $m < n$ ,  $W(A) = 0$  if  $\|A\| \leq n - m$ , and  $W(A) = j/m$  if  $\|A\| = n - m + j$  for  $j > 0$ . The agent is pessimistic in ignoring the  $n - m$  best states of nature.
- (ii) **MOSAICS FOR (CONTINUOUS EXTENSION OF) THE ELLSBERG PARADOX.**  $S = [0, 1] \times [0, 1]$ , where  $(s_1, s_2)$  describes the location ( $s_1$ ) and momentum ( $s_2$ ) of an elementary particle.  $\mathcal{A}$  contains all sets of the form  $A = A_1 \times [0, 1]$  and  $[0, 1] \times A_2$  for  $A_1$  and  $A_2$  Borel-measurable sets. With probability measure  $P$  the Lebesgue-measure, and  $w(p) = p^2$ , assume that  $W(A_1 \times [0, 1]) = P(A_1)$  and  $W([0, 1] \times A_2) = w(P(A_2))$ .  $\mathcal{A}$  is a mosaic (and also a  $\lambda$ -system) but not an algebra. The nonlinear  $w$  may reflect that while the first coordinate, describing location, has a well-known statistical distribution, the second coordinate, describing momentum, does not, leading an ambiguity-averse person to the pessimistic evaluations of momentum-uncertainty as assumed here. All events, including the events regarding momentum, are unambiguous according to the formal definition of Epstein

and Zhang (2001) (our Equation 7.3). Epstein and Zhang (2001, Section 9) argued that ambiguity of momentum events should be inferred from intersections of location and momentum events, and decisions based on these.

Example 5.8(i) shows that our axioms can be satisfied in finite nondegenerate structures without requiring richness of outcomes, contrary to the finite-state CEU-models in Casadesus–Masanell et al. (2000, Lemma A.5), Chew and Karni (1994), Ghirardato and Marinacci (2001), Köbberling and Wakker (2003), Nakamura (1990), or Wakker (1989, Theorem VI.5.1).

Characterizations of prospect theory (Tversky and Kahneman, 1992) under solvability conditions for outcomes seem to be very complex, and are not yet available in the literature (Köbberling and Wakker, 2003). Similarly, a dual version thereof with outcomes and events interchanged, and extending our characterizations to prospect theory, remains an open problem at this stage. The extension to the multiple priors model (Chateauneuf, 1991; Gilboa and Schmeidler, 1989; Wald, 1950) similarly is an open problem.

## 6. APPLICATIONS OF THE LIKELIHOOD METHOD TO RANK-DEPENDENT MODELS

This section shows how the likelihood method can be used to examine the nature of decision weights in rank-dependent models.

### 6.1. *Rank-dependence for uncertainty*

Whereas the measurement of uniform partitions  $\{E_1, \dots, E_n\}$ , and their use for quantitative measurements of probability, is well understood under subjective expected utility, the analog for Choquet expected utility has not yet been expressed in terms of directly observable preference conditions. By means of the likelihood method, this can be done straightforwardly. We take partitions  $(E_1, \dots, E_n)$  for which  $E_j^{R_j} \sim E_{j-1}^{R_{j-1}}$  for all  $j$ , with  $R_j$  denoting  $E_1 \cup \dots \cup E_{j-1}$ . Under

Choquet expected utility, we then know that all decision weights  $\pi(E_j^{R_j})$  must be  $1/n$ , and  $W(R_j) = (j - 1)/n$  for all  $j$ . In this manner, direct quantitative measurements of capacities can be obtained.

We can also test and characterize qualitative properties of capacities. A capacity  $W$  is *convex* if  $W(A \cup B) - W(A) \geq W(B) - W(A \cap B)$  for all events  $A, B$  for which  $A \cup B$  and  $A \cap B$  are events. Convexity holds if and only if  $\pi(A^R)$  is non-decreasing in  $R$ , in the sense that  $\pi(A^{R'}) \geq \pi(A^R)$ , whenever  $R' \supset R$  (substitute  $R' \setminus R = A \setminus B$ ,  $R = A \cap B$ ,  $A = B \setminus A$ ). The condition was characterized for linear utility by Chateauneuf (1991) and Schmeidler (1989), and for continuous utility or convex-ranged capacities by Wakker (2005). The following theorem gives the result for the general model characterized in Theorem 5.7, using a preference condition that excludes violations of convexity in a transparent manner.

**THEOREM 6.1.** *Assume that the conditions and model of Theorem 5.7 hold. Then  $W$  is convex if and only if  $A^{R'} \prec A^R$  for no  $A$  and  $R' \supset R$ .*

The preference condition in the above theorem directly precludes events being weighted less as they are ranked worst. We hope that the transparency and simplicity of the characterizing condition will be taken as an advantage of the likelihood method. Similar results for concavity and the, empirically prevailing, cavex capacities (concave for unlikely events and convex for likely; see Wakker, 2001; Wu and Gonzalez, 1999) can be obtained analogously. For two agents  $\succsim_1$  and  $\succsim_2$ , a condition such as

$$\text{if } R' \supset R \text{ and } A \text{ and } B \text{ are nested, then not} \\ [A^{R'} \succsim_1 B^R \text{ and } A^{R'} \prec_2 B^R]$$

suggests greater pessimism for  $\succsim_2$  than for  $\succsim_1$ , because then the overweighting of badly-ranked events  $A$  relative to better-ranked events  $B$  cannot be less for  $\succsim_2$  than for  $\succsim_1$ . Here, as elsewhere, the symbol  $\succ$  with sub- or superscripts denotes the union of  $\succ$  and  $\sim$  with corresponding sub- and superscripts.

Tversky and Wakker (1995) considered the preference conditions in Figure 7. They imposed these conditions only if  $L$  and  $H$  are not very unlikely, formalized through boundary restrictions, so as to avoid comparisons between events  $A$  when ranked (close to) worst and when ranked (close to) best. These preference conditions are more often satisfied empirically than convexity, both for risk (reviewed by Luce, 2000, Section 3.4) and for uncertainty (reviewed by Wakker, 2004). The configurations of Figure 7 are readily recognized as special cases of the likelihood method, with Figure 7a revealing  $A^\varnothing \succcurlyeq A^B$  and Figure 7b revealing  $A^{A^c} \succcurlyeq A^H$ . Substituting Choquet expected utility yields  $\pi(A^\varnothing) \geq \pi(A^B)$  in Figure 7a and  $\pi(A^{A^c}) \geq \pi(A^H)$  in Figure 7b. Thus, the conditions imply that event  $A$  is weighted more for extreme ranks where it is ranked best or worst, than for moderate ranks where it is ranked in the middle (“extremity-orientedness”). The authors showed that under some richness assumptions, the preference conditions in the figures are not only necessary, but also sufficient, for the inequalities described.

## 6.2. Measuring decision weights for risk

The likelihood method is particularly suited for decision under risk. Under risk, the solvability condition is naturally satisfied owing to the richness of the probability scale. A similar richness is available in the time dimension for intertemporal preferences, and this is another promising domain of application. We will, however, restrict attention to risk and uncertainty in this paper. Strangely enough, only three papers used richness in the probability dimension to char-

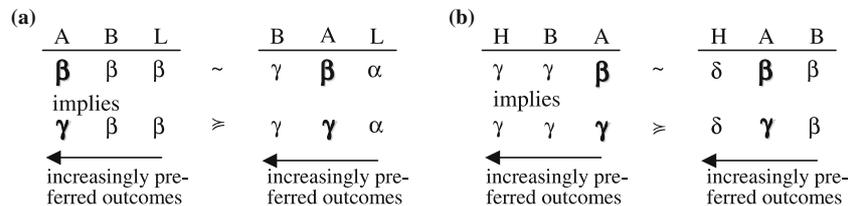


Figure 7. Extremity-orientedness.

acterize rank-dependent utility for risk: namely, Abdellaoui (2002), Nakamura (1995), and Zank (2004). We show how the likelihood method can be used to measure decision weights and probability transformations for Quiggin’s (1981) rank-dependent utility for given probabilities. The corresponding consistency axiom, defined later, is equivalent to the probability tradeoff consistency that Abdellaoui (2002) used to characterize Quiggin’s (1981) rank-dependent utility. The likelihood method thus gives a new interpretation to Abdellaoui’s axiom in the same way as it did to Gilboa’s (1987) axiom  $P2^*$ .

$P = (p_1 : c_1, \dots, p_n : c_n)$  denotes the probability distribution yielding  $c_j$  with probability  $p_j$ , where it is understood that  $c_1 \succ \dots \succ c_n$ . Probabilities  $p_j$  play the same role as events  $A_j$  in acts, with capacities  $W(A_j)$  replaced by  $w(p_j)$  for a probability transformation  $w$ .

Using the more familiar notation of lottery mixtures, Figure 8 can, with the “conditional” lotteries  $P$  and  $Q$  defined appropriately, be rewritten as

$$p_i \beta + (1 - p_i)P \sim q_j \beta + (1 - q_j)Q \quad \text{and}$$

$$p_i \gamma + (1 - p_i)P \succ q_j \gamma + (1 - q_j)Q.$$

Figure 8 shows how we can reveal that the decision weight of probability  $p_i$ , in a rank of probability  $p_1 + \dots + p_{i-1}$ , exceeds that of probability  $q_j$  in a rank of probability  $q_1 + \dots + q_{j-1}$ . If there is equivalence instead of strict preference in the figure, then the two decision weights should be the same.

The consistency axiom corresponding to the above revelation can replace the probability tradeoff consistency of Abdellaoui (2002). By our Theorem 5.7, stochastic dominance, and

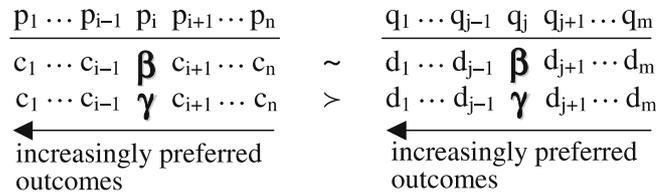


Figure 8. Revealed orderings of decision weights.

Wakker (1990), a characterization of rank-dependent utility then results. Abdellaoui considered measurements with antecedent weak preferences instead of indifferences, as discussed further in the next section (Equation (7.1)). Zank (2004) considered the condition with only indifferences as in this paper, and with a further restriction introduced by Chateauneuf (1999). Abdellaoui (2002) and Zank (2004) formulated their conditions in terms of  $n$ -tuples of probabilities of ranks (probabilities of receiving some outcome or any other better outcome), instead of directly in terms of acts or lotteries as done in this paper.

#### 7. A NEW INTERPRETATION OF OTHER CONDITIONS IN THE LITERATURE

This section reinterprets some conditions for preferences under uncertainty, studied in the literature for some models not considered in the preceding sections, in terms of the likelihood method. A large collection of such models and their applications is available in Gilboa (2004). To simplify comparisons with the literature, we first state a generalization of the likelihood method in Figure 3 that uses weak preference instead of indifference as the point of departure. We write  $A \succsim_w B$  if there exist  $\gamma \succ \beta$  and acts  $f, g$  such that the left two weak preferences in Equation (7.1) hold. We write  $A \prec_w B$  if there exist  $\gamma' \succ \beta'$  and acts  $f', g'$  such that the right two preferences in Equation (7.1) hold. Generalized likelihood conditions can be defined that exclude inconsistencies of the kind  $A \succsim_w B$  and  $A \prec_w B$ . Equation (7.1) states such a condition directly in terms of preferences:

$$\text{For } \gamma \succ \beta \text{ and } \gamma' \succ \beta':$$

$$\text{NOT } \begin{cases} \beta_A f \preceq \beta_B g & \& \beta'_A f' \succ \beta'_B g' & \& \\ \gamma_A f \succ \gamma_B g & \& \gamma'_A f' \prec \gamma'_B g'. \end{cases} \quad (7.1)$$

It is easy to show that the generalized measurements, and their adaptations to various models, still reveal order-

ings of likelihoods. The conditions are stronger than the ones used in our theorems and, because their necessity is always obvious, they can be used as well to characterize various preference models. We chose our likelihood consistencies with indifferences because they are simpler and yield more general theorems. Because the preference conditions considered in the literature were usually of the general form of Equation (7.1), the equation facilitates comparisons. The conditions in the literature were usually further reinforced by imposing them also for  $\gamma \preceq \beta$  and  $\gamma' \preceq \beta'$ , a difference that we will ignore henceforth.

Consider a violation of the sure-thing principle with  $\beta_A f \preceq \beta_A g$  but  $\gamma_A f \succ \gamma_A g$ . The improvement of the outcome on  $A$  implies  $A \succ_w A$ . The sure-thing principle, which excludes this pattern, thus amounts to imposing irreflexivity of the strict revealed-likelihood relation  $\succ_w$ . Examples of violations thereof are in Section 3.2. Given the trivial  $A \sim A$ , the sure-thing principle is equivalent to likelihood consistency restricted to the case of  $A = B$ , and imposed on  $\succ_w$  instead of  $\succ$ . Whereas P4 of Savage (1954) excludes contradictions in basic likelihood revelations between different events, the sure-thing principle excludes contradictions in nonbasic likelihood revelations for single events. For the rank-dependent case, similar observations apply to the comonotonic sure-thing principle and ranked events  $A^R$ .

Under state-dependent expected utility (Karni, 1993; Nau, 1995), outcomes cannot be compared across different events. Hence, it is natural that, to accommodate the likelihood method so as to allow for state dependence, the method be restricted to cases  $A = B$ , which amounts to imposing irreflexivity on the strict relative-likelihood relation – i.e., the requirement that  $A \succ A$  for no event  $A$ . The derivation of state-dependent models from such axioms is a topic of future research.

Machina and Schmeidler (1992) considered probabilistic sophistication, as follows. Write  $A \succ_{ms} B$  if, for disjoint events  $A, B$ , the configuration of Figure 3 holds with

- (i)  $f$  and  $g$  identical (to an act  $h$  in their notation) off  $A \cup B$ ;
- (ii)  $f$  and  $g$  identical to  $\beta$  on  $A \cup B$ .

Such a likelihood revelation, for weak preference  $\succsim_{ms}$ , instead of strict preference  $\succ_{ms}$ , is illustrated by the left indifference  $\sim$  together with the preference  $\succsim$  in Equation (7.2). Restriction (i) ensures that outcome tradeoffs off  $A$  and  $B$  are vacuous, so that, in particular, they are not affected by the change of outcomes under  $A$  or  $B$ . Restriction (ii) ensures that also within  $A$  and  $B$  only tradeoffs of  $\gamma$  against  $\beta$ , and their likelihoods, can play a role, and no other interactions between outcomes or their tradeoffs play a role. Thus, likelihood revelations of probabilities, independently of a decision model for given probabilities, can be obtained in a way alternative to our method in section 3.3. In terms of the likelihood method in Equation (7.1) the strong comparative probability axiom P4\* of Machina and Schmeidler (1992) can be interpreted as a consistency requirement excluding  $A \succsim_{ms} B$  and  $A \prec_{ms} B$ . Machina and Schmeidler stated the condition directly in terms of preferences given hereafter, where we add the upper trivial indifferences to clarify the interpretation in terms of the likelihood method.

For  $\gamma \succ \beta$  and  $\gamma' \succ \beta'$ :

$$\text{NOT } \begin{cases} \beta_A \beta_B h \sim \beta_B \beta_A h & \& \beta'_A \beta'_B h' \sim \beta'_B \beta'_A h' & \& \\ \gamma_A \beta_B h \succsim \gamma_B \beta_A h & \& \gamma'_A \beta'_B h' \prec \gamma'_B \beta'_A h'. \end{cases} \quad (7.2)$$

Equation (7.2) gives another natural condition intermediate between basic likelihood consistency, which is the special case of  $h = \beta$ , and likelihood consistency, from which it results if  $f = \beta_B h$  and  $g = \beta_A h$  in Equation (7.1). Machina and Schmeidler showed that, under common assumptions and Savage's P6, this condition is not only necessary but also sufficient for probabilistic sophistication.

We next consider the definitions of (un)ambiguity of Zhang (2002) and Epstein and Zhang (2001). Event  $T$  is *linearly unambiguous* if not  $(\beta_T f \preceq \beta_T g$  and  $\gamma_T f \succ \gamma_T g)$ , i.e. not  $T \succ_w T$ , with a similar condition imposed on  $T^c$ , i.e. not  $T^c \succ_w T^c$ .

In our terminology, an event is linearly unambiguous if likelihood orderings restricted to this event and its complement contain no inconsistencies. Epstein and Zhang (2001) argued that linear unambiguity is restrictive: Imagine that interactions between tradeoffs of outcomes under  $T^c$  with outcomes under  $T$  are permitted for reasons other than ambiguity, e.g. because of deviations from expected utility for unambiguous events such as in Quiggin's (1981) rank-dependent utility. Then these interactions can explain the reversal of preference, rather than an internal inconsistency (i.e., ambiguity) in revealed likelihood of  $T$ .

Epstein and Zhang (2001) introduced another, weaker, definition of unambiguous events, so as to avoid the interactions just mentioned. An inconsistency in revealed likelihood, as excluded by linear unambiguity, is now excluded only if the preferences between the acts are driven merely by likelihood considerations and not by tradeoffs between outcomes. This led the authors to exclude only inconsistencies of the following kind.

$$\text{For } \gamma > \beta: \text{ NOT } \left\{ \begin{array}{l} \beta_T \lambda_A \mu_B h \preceq \beta_T \mu_A \lambda_B h \text{ \& } \\ \gamma_T \lambda_A \mu_B h \succ \gamma_T \mu_A \lambda_B h \end{array} \right. , \quad (7.3)$$

with a similar condition for  $T^c$ . Here no tradeoffs of outcomes other than  $\lambda$  versus  $\mu$  are involved in each preference, and the preferences are driven merely by the likelihood of  $A$  versus  $B$ , and not by tradeoffs of outcomes in any other way. What happens under  $T$  (if constant so that no subset of  $T$  is involved) should not affect likelihood orderings off  $T$  if  $T$  is unambiguous, in the interpretation of Epstein and Zhang. If we accept this interpretation, then the switch in preference in Equation (7.3) can be explained neither by changes in tradeoffs of outcomes nor by changes in the likelihood ordering of  $A$  and  $B$ . It should, therefore, under the interpretation of the likelihood method, imply that  $T$  is more likely than itself, an inconsistency that is interpreted as ambiguity and is to be excluded.

## NOTES

1. de Finetti (1974, Section 3.1.4) argued for the logic and efficiency of this notation.
2. Under CEU, this case corresponds with  $U(x_\ell) - U(x_{\ell+1}) \geq U(x_m) - U(x_{m+1})$ ; note here that  $W(D^2) = \pi((D^2 \setminus D^1)^{D^1}) = \pi((E^\ell \setminus E^{\ell+1})^{E^{\ell+1}})$ , both being equal to  $\pi((A \setminus B)^B)(U(x_m) - U(x_{m+1})) / (U(x_\ell) - U(x_{\ell+1}))$ .
3. Under CEU, this case corresponds with  $U(x_\ell) - U(x_{\ell+1}) \leq U(x_m) - U(x_{m+1})$ .
4. For  $A \succ E$ , we have  $\beta_A f \sim \beta_E g$  and  $\gamma_A f \succ \gamma_E g$ . We add  $\gamma_E g \succ \beta_E g \sim \beta_A f$ , after which solvability yields event  $A'$  with  $\gamma_{A'} \beta_{A \setminus A'} f \sim \gamma_E g$  which, together with the first indifference, implies  $A' \sim E$ .

## APPENDIX A. PREPARATORY COMMENTS

This Appendix A presents preparatory results. Then first Theorem 5.7 is proved in Appendixes B and C. Next the other results are proved in order of appearance, in Appendix D. In particular, Theorem 5.3, our generalization of Savage's (1954) preference foundation of expected utility, is obtained there as a corollary of Theorem 5.7. We are not aware of an easier way to derive this expected-utility result, because the duality between outcomes and events used in our proof automatically leads to rank dependence. This explains why we derive Theorem 5.7 prior to the other results. The proof of the following lemma is easy and is omitted.

**LEMMA A.1.** *For each act  $(A_1 : x_1, \dots, A_n : x_n)$ , the algebra generated by  $A_1, \dots, A_n$  is contained in  $\mathcal{A}$ . Consequently, for each such act all  $(A_1 : y_1, \dots, A_n : y_n)$  are acts too.*

In the main text, rank-dependence was expressed in terms of ranks. In the literature, rank-dependence is mostly expressed in terms of complete rank-orderings of the state space and comonotonicity. A *rank-ordering*  $r$  is an ordering (called “better”) on  $S$ , which means that it is complete, transitive, and antisymmetric (no different states are ranked the same). Act

$f$  is *compatible* with  $r$  if  $f(s) \succcurlyeq f(t)$  whenever  $r$  ranks  $s$  better than  $t$ . In particular, every constant act is compatible with every  $r$ .  $A \subset S$  is  *$r$ -connected* if  $A$  is an event and every state off  $A$  is either better than all states in  $A$  or worse than all states in  $A$ .  $\mathcal{A}$ -measurability of  $f$  implies that, for a compatible  $r$ , some of the  $r$ -intervals of states must be events and, thus, are  $r$ -connected.

The set of all acts compatible with  $r$  is called the *comoncone* for  $r$ . A set of acts is *comonotonic* if all acts are compatible with a same rank-ordering  $r$ , which holds if and only if the set is a subset of one comoncone. Under weak ordering, it is well known that a set of acts is comonotonic if and only if every pair  $f, g$  is, and the latter holds if and only if  $f(s) \succ f(t)$  and  $g(s) \prec g(t)$  for no states  $s, t$ . For  $r$ , event  $A$  is *ranked better than* event  $B$  if every state in  $A$  is ranked better than every state in  $B$ . Two acts  $(A_1 : x_1, \dots, A_n : x_n)$  and  $(A_1 : y_1, \dots, A_n : y_n)$  are comonotonic if and only if for each event  $A_i$  we can take the same rank for both acts.

Write  $A^r \succ B^{r'}$  if there exist  $\gamma \succ \beta$  and  $f, g$  such that the configuration of Figure 3 holds, where further the left two acts are compatible with  $r$ , and the right two with  $r'$ . Then  $A^R \succ B^{R'}$  as in Figure 6 where: (a)  $R = \{s \in S \setminus A : f(s) \succcurlyeq \gamma\}$  and  $R' = \{t \in S \setminus B : g(t) \succcurlyeq \gamma\}$ ; (b)  $L = (R \cup A)^c = \{s \in S \setminus A : f(s) \preccurlyeq \beta\}$  and  $L' = (R' \cup B)^c = \{t \in S \setminus B : g(t) \preccurlyeq \beta\}$ ; (c) all these sets are events, meaning that they are contained in  $\mathcal{A}$  (follows easily because acts take only finitely many values); (d)  $A$  is  $r$ -connected and  $B$  is  $r'$ -connected. Conversely, if  $A^R \succ B^{R'}$  then  $r$  and  $r'$  as above can be found. Similar observations apply to  $A^r \sim B^{r'}$  and  $A^R \sim B^{R'}$ .

*Comonotonic likelihood consistency (clc)* holds if the implication of Equation (3.2) holds whenever

$$\begin{aligned} &\{\beta_A f, \gamma_A f, \beta'_A f', \gamma'_A f'\} \text{ is comonotonic and} \\ &\{\beta_B g, \gamma_B g, \beta'_B g', \gamma'_B g'\} \text{ is comonotonic.} \end{aligned} \tag{A.1}$$

Clc is implied by rank-dependent likelihood consistency because comonotonicity implies that ( $\mathcal{A}$ -measurable) ranks  $R$  and  $R'$  can be chosen as required. In isolation it is weaker

$$\begin{array}{l}
\text{(a)} \\
A T E_1 \dots E_{j-1} E_j E_{j+1} E_{j+2} \dots L \quad \sim \quad A T E_1 \dots E_{j-1} E_j E_{j+1} E_{j+2} \dots L \\
\varepsilon \gamma \gamma \dots \gamma \beta \beta \beta \dots \beta \quad \sim \quad \delta \gamma \gamma \dots \gamma \gamma \beta \beta \dots \beta \\
\\
\text{(b)} \\
A T E_1 \dots E_{j-1} E_j E_{j+1} E_{j+2} \dots L \quad \sim \quad A T E_1 \dots E_{j-1} E_j E_{j+1} E_{j+2} \dots L \\
\varepsilon \gamma \gamma \dots \gamma \gamma \beta \beta \dots \beta \quad \sim \quad \delta \gamma \gamma \dots \gamma \gamma \gamma \beta \dots \beta
\end{array}$$

Figure A.1. (a) The *weak Archimedean axiom*: If  $\varepsilon > \delta \succ \gamma > \beta$ ,  $\varepsilon_A \delta > \delta$ , and the equivalence holds for each  $j$ , then the sequence  $E_1, E_2, \dots$  must be finite. (b) The same preference for  $j+1$  instead of  $j$  shows that  $E_{j+1} \sim E_j$ .

because it does not impose restrictions on configurations as in Equation (3.2) if, for example,  $f$  and  $f'$  rank-order states within the rank  $R$  of  $A$  differently. In the presence of the other conditions, clc will imply rank-dependent likelihood consistency after all, as follows from Theorem A.3, so that the two conditions are equivalent.

We will use a weakened version of the Archimedean axiom in our proof, so as to enable the derivation of other results in the literature as direct corollaries of our result. The *weak Archimedean axiom* is defined in Figure A.1(a) where, for each  $j$ , receiving  $\gamma$  instead of  $\beta$  under event  $E_j$  exactly offsets receiving  $\varepsilon$  instead of  $\delta$  under event  $A$  (for all other events, the outcomes are the same). The axiom is a joint weakening of the Archimedean axioms in the main text, of Gilboa's (1987) Archimedean axiom, and of the axioms in Chew and Sagi (2004) and Krantz et al. (1971) that exclude infinite "standard sequences."

An interpretation in terms of equally likely events can be inferred by comparing Figure A.1(a) to Figure A.1(b) through the likelihood method: the improvement of  $\beta$  to  $\gamma$  under event  $E_j$  for the left act, and the same improvement under event  $E_{j+1}$  for the right act, maintain indifference, so that  $E_j \sim E_{j+1}$  follows. The assumed  $\varepsilon_A \delta > \delta$  ensures nonnullness of  $A$  and, consequently, of the  $E_j$ 's. Hence, the Archimedean axiom implies the weak Archimedean axiom. If we set  $E_0 = A \cup T$ , then the rank of each event  $E_j$  in all acts above is  $E_0 \cup \dots \cup$

$E_{j-1}$ , from which it follows that the rank-dependent Archimedean axiom also implies the weak Archimedean axiom. Note here that in the rank-dependent Archimedean axiom the event  $E_0$  is not equally likely as the other events, but serves as an arbitrary starting point of the ranks. The following lemma follows.

LEMMA A.2. *The Archimedean axiom and the rank-dependent Archimedean axiom imply the weak Archimedean axiom.*

THEOREM A.3. *In Theorem 5.7, rank-dependent likelihood consistency can be weakened to comonotonic likelihood consistency, and the rank-dependent Archimedean axiom to the weak Archimedean Axiom.*

We will prove Theorem 5.7 in the more general version of Theorem A.3. It will clarify the mathematical relations between our results and preceding results in the literature. We chose the mathematically less general approach with the not-weak Archimedean axiom and ranks in the main text because it is more transparent. Specifications of complete rank-orderings of state spaces are empirically intractable if state spaces are big, for instance if they are continua, and raise  $\mathcal{A}$ -measurability complications. These empirical and mathematical complications of comonotonicity are avoided when using ranks.

OBSERVATION A.4. *The axioms in Gilboa (1987) imply our axioms.*

*Proof.* Gilboa considered the collection of all subsets of  $S$ , which is a mosaic as required in our theorem. His axioms are as follows, where asterixes indicate generalizations of Savage's corresponding axioms. P1: weak ordering; P2\*: Equation (7.1) with indifferences instead of the upper two preferences under the extra comonotonicity restrictions of Equation (A.1), which is still stronger than clc (Equation (7.1) and (A.1) with indifferences for all three weak preferences); P3\*: monotonicity; there is no P4\*; P5\*: three or more nonequivalent outcomes; P6\*: solvability (stronger than ours because also imposed if there are outcomes of  $f$  and  $g$  between  $\beta$  and  $\gamma$ ); P6\*\*\*: Weak Archimedeanity (our axiom is

a special case of his, with substitutions  $f_j$  for the right act in Figure A.1(a),  $x = \varepsilon$ ,  $y = \delta$ ,  $A = A$ ); P7\*: only needed for non-simple acts, which are not considered here.  $\square$

*Strong monotonicity* holds if monotonicity holds and further  $\gamma_A f \succ \beta_A f$  whenever  $\gamma \succ \beta$  and  $A$  is nonnull. *Rank-dependent strong monotonicity* holds if monotonicity holds and further  $\gamma_{A^R} f \succ \beta_{A^R} f$  whenever  $\gamma \succ \beta$  and  $A^R$  is nonnull. For a comoncone with respect to  $r$ , and an  $r$ -connected event  $A$ ,  $A$  is *nonnull on the comoncone* if  $(R : \gamma, A : \gamma, L : \beta) \succ (R : \gamma, A : \beta, L : \beta)$  for some  $\gamma \succ \beta$  with both acts contained in the comoncone, and  $A$  is *null on the comoncone* otherwise. *Comonotonic strong monotonicity* holds if monotonicity holds and, furthermore,  $\gamma_A f \succ \beta_A f$  whenever  $\gamma \succ \beta$  and  $A$  is nonnull on a comoncone that contains both acts (implying that  $A$  is  $r$ -connected for the compatible  $r$ ). Rank-dependent strong monotonicity implies comonotonic strong monotonicity because for comonotonic acts  $\gamma_A f$  and  $\beta_A f$  the same rank can be chosen for event  $A$ .

The following lemma shows that likelihood consistency implies strong forms of monotonicity. The idea is that nullness is captured by likelihood revelations, so that an event cannot manifest itself as null in one situation and nonnull in another.

**LEMMA A.5.** *Assume weak ordering and monotonicity. Then likelihood consistency implies strong monotonicity, rank-dependent likelihood consistency implies rank-dependent strong monotonicity, and clc implies comonotonic strong monotonicity.*

*Proof.* Let  $A$  be nonnull. That is, besides the trivial  $(R : \gamma, \emptyset : \gamma, A : \beta, L : \beta) \sim (R : \gamma, \emptyset : \beta, A : \beta, L : \beta)$ , we also have  $(R : \gamma, \emptyset : \gamma, A : \gamma, L : \beta) \succ (R : \gamma, \emptyset : \gamma, A : \beta, L : \beta)$  for some  $\gamma \succ \beta$ , implying  $A \succ \emptyset$  and  $A^R \succ \emptyset^R$ . If, for some  $\gamma' \succ \beta'$ , besides the trivial  $\gamma' \emptyset \beta'_A f \sim \beta'_A \beta'_\emptyset f$ , we also had  $\gamma' \emptyset \gamma'_A f \sim \beta'_A \gamma'_\emptyset f$ , then  $A \sim \emptyset$  and a violation of likelihood consistency would result, and of rank-dependent and clc if we have  $A^R$  instead of  $A$  in the above indifferences (compare to  $\emptyset^R$ ; the required comonotonicities hold). Preferences  $\gamma'_A f \prec \beta'_A f$  or  $\gamma'_{A^R} f \prec \beta'_{A^R} f$  are excluded by (“weak”) monotonicity. Hence, the strict prefer-

ences  $\gamma'_A f \succ \beta'_A f$  or  $\gamma'_{A^R} f \succ \beta'_{A^R} f$  must hold, as required for strong (rank-dependent/comonotonic) monotonicity.  $\square$

APPENDIX B. PROOF OF THEOREMS 5.7 AND A.3; NECESSITY,  
AND SUFFICIENCY FOR FINITELY MANY OUTCOMES

The implicaton (i)  $\Rightarrow$  (ii) in Theorems 5.7 and A.3 follows from substitution. We, therefore, assume the preference conditions in (ii) of Theorem 5.7, in the weakened version of Theorem A.3, throughout, and derive CEU plus the uniqueness results. In this Appendix B, we assume that  $\mathcal{C}$  has finitely many equivalence classes. Fix  $x_1 \succ \cdots \succ x_{n+1}$  for some  $n \geq 2$ , and assume that every other outcome is equivalent to some  $x_i$ . By monotonicity, we may replace every such outcome by the equivalent  $x_i$  without affecting preference, i.e. we may assume that there are no outcomes other than the  $x_j$ s.

*B.1. Extending preferences to all increasing  $n$ -tuples of events*

A collection of events is *nested* if, for every pair of elements, one is a subset of the other. A collection of events is an  $\mathcal{A}$ -*collection* if the algebra generated by it is contained in  $\mathcal{A}$ . Similarly, a sequence of events is an  $\mathcal{A}$ -*sequence* if the algebra generated by it is contained in  $\mathcal{A}$ , and an  $n$ -*tuple* of events is an  $\mathcal{A}$ - $n$ -*tuple* likewise. We will often use without further mention that every partition  $\{A_1, \dots, A_n\}$  of  $S$  consisting of events is an  $\mathcal{A}$ -*collection*. An  $n$ -tuple of events  $(A_1, \dots, A_n)$  is *increasingly nested* if  $A_1 \subset \cdots \subset A_n$ . For each act  $(F_1^* : x_1, \dots, F_n^* : x_n, F_{n+1}^* : x_{n+1})$ , we define the events  $F_j = F_1^* \cup \cdots \cup F_j^*$  for all  $j$ , and also denote the act by the increasingly nested  $n$ -tuple  $(F_1, \dots, F_n)$ , also denoted  $F$ ; see Figure B.1. The implicit convention is that  $F_0 = \emptyset$  and  $F_{n+1} = S$ .  $F_i$  is the event of receiving  $x_i$  or more, i.e. the rank of outcome  $x_{i+1}$ . By Lemma A.1, all such  $n$ -tuples derived from acts are  $\mathcal{A}$ - $n$ -tuples. Not every increasingly nested  $n$ -tuple  $(A_1, \dots, A_n)$  of events necessarily designates an act, because the differences  $A_{j+1} \setminus A_j$  need not be events for the mosaic  $\mathcal{A}$ . An increasingly nested  $n$ -tuple desig-

The relation between acts and nested n-tuples of ranks.

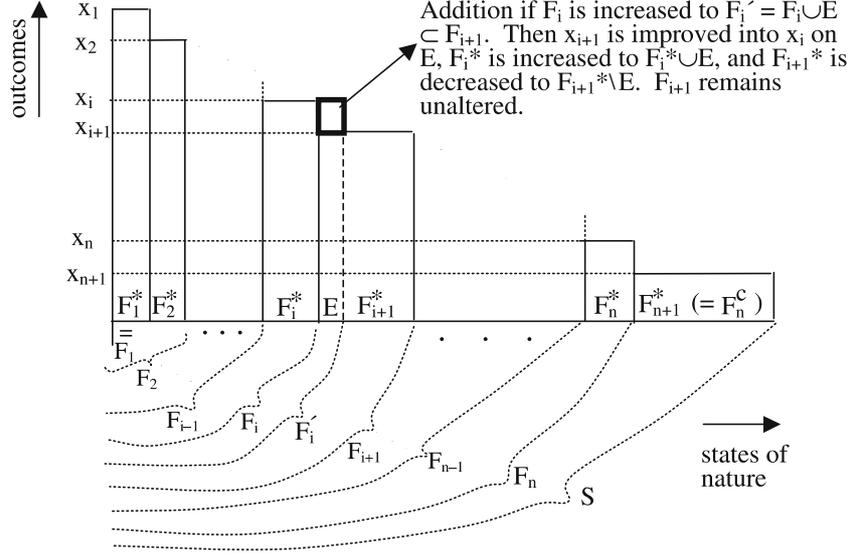


Figure B.1.  $(F_1, \dots, F_n) = (F_1^*: x_1, \dots, F_n^*: x_n, F_{n+1}^*: x_{n+1})$ .  
 $(F_1, \dots, F_i \cup E, F_{i+1}, \dots, F_n) = (F_1^*: x_1, \dots, F_i^* \cup E: x_i, F_{i+1}^* \setminus E: x_{i+1}, \dots, F_n^*: x_n, F_{n+1}^*: x_{n+1})$ .

nates an act if and only if it is an  $\mathcal{A}$ - $n$ -tuple. The following lemma follows immediately from monotonicity.

**LEMMA B.1** [Monotonicity with respect to  $\supset$ ]. *For all acts  $(F_1, \dots, F_n), (G_1, \dots, G_n)$ , we have  $(F_1, \dots, F_n) \succcurlyeq (G_1, \dots, G_n)$  if  $F_j \supset G_j$  for all  $j$ .  $\square$*

**LEMMA B.2** [Savage's (1954) P4, i.e. strong monotonicity and weak ordering for  $\succcurlyeq_b$ ].  *$A \succcurlyeq_b B$  if and only if  $(A: \gamma, A^c: \beta) \succcurlyeq (B: \gamma, B^c: \beta)$  for all  $\gamma \succcurlyeq \beta$ . Further,  $\succcurlyeq_b$  is a weak order.*

*Proof.* Everything for the case  $\gamma \sim \beta$  follows from monotonicity. We, therefore, only consider cases  $\gamma \succ \beta$ . If  $A \sim_b B$  then  $A \sim B$ , and ctc and the trivial  $(A: \beta, A^c: \beta) \sim (B: \beta, B^c: \beta)$  imply  $(A: \gamma, A^c: \beta) \sim (B: \gamma, B^c: \beta)$  for all  $\gamma \succ \beta$ . Using this implication, we consider three cases. In the first case of  $(A: x_1, A^c: x_2) \sim (B: x_1, B^c: x_2)$ , we have  $(A: \gamma, A^c: \beta) \sim (B: \gamma, B^c: \beta)$  for all  $\gamma \succ \beta$ , and everything follows.

The second case concerns  $(A: x_1, A^c: x_2) \succ (B: x_1, B^c: x_2)$ . This,  $(B: x_1, B^c: x_2) \succcurlyeq (A: x_2, A^c: x_2)$ , and solvability imply exis-

tence of  $D \subset A$  with  $(D : x_1, D^c : x_2) \sim (B : x_1, B^c : x_2)$ . As we saw before,  $D \sim_b B$  implies that for all  $\gamma \succ \beta$  we have  $(B : \gamma, B^c : \beta) \sim (D : \gamma, D^c : \beta)$ . Monotonicity and transitivity give  $(B : \gamma, B^c : \beta) \preceq (A : \gamma, A^c : \beta)$ . Equivalence cannot hold (it would lead to the first case), so that  $(A : \gamma, A^c : \beta) \succ (B : \gamma, B^c : \beta)$  follows. The third case,  $(A : x_1, A^c : x_2) \prec (B : x_1, B^c : x_2)$ , similarly implies  $(A : \gamma, A^c : \beta) \prec (B : \gamma, B^c : \beta)$ .

The preferences with  $x_1 \succ x_2$  are always the same as with all  $\gamma \succ \beta$  instead of  $x_1, x_2$  and everything in the lemma follows.  $\square$

**LEMMA B.3** [Solvability for  $\succ_b$ ]. *If  $A \succ_b B \succ_b C$ ,  $A \supset C$ , and  $A \setminus C$  is an event, then  $B' \sim_b B$  for some event  $B'$  with  $A \supset B' \supset C$  and  $A \setminus B'$  and  $B' \setminus C$  events.*

*Proof.* By Lemma B.2,  $(C : x_1, A \setminus C : x_1, A^c : x_2) = (A : x_1, A^c : x_2) \succ (B : x_1, B^c : x_2) \succ (C : x_1, C^c : x_2) = (C : x_1, A \setminus C : x_2, A^c : x_2)$ . By solvability,  $(C : x_1, E_1 : x_1, E_2 : x_2, A^c : x_2) \sim (B : x_1, B^c : x_2)$  for some events  $E_1, E_2$  partitioning  $A \setminus C$ . Take  $B' = C \cup E_1 \in \mathcal{A}$ .  $\square$

Equivalence-dominance (Equation (3.3)) holds if and only if, for all acts  $(F_1, \dots, F_n)$  and  $(G_1, \dots, G_n)$ , we have  $(F_1, \dots, F_n) \sim (G_1, \dots, G_n)$  whenever  $F_j \sim_b G_j$  for all  $j$ .

**LEMMA B.4** *Equivalence-dominance holds.*

*Proof.* For acts  $(F_1, \dots, F_n), (G_1, \dots, G_n)$ , let  $F_j \sim_b G_j$  for all  $j$ . By Lemma B.2,  $(\emptyset, \dots, \emptyset, F_n) \sim (\emptyset, \dots, \emptyset, G_n)$ . Suppose that, for some  $j \leq n$ ,  $(\emptyset, \dots, \emptyset, F_j, \dots, F_n) \sim (\emptyset, \dots, \emptyset, G_j, \dots, G_n)$  (induction hypothesis).  $F_{j-1} \sim_b G_{j-1}$  implies  $F_{j-1} \sim G_{j-1}$ . This, and clc (concerning improvements of  $\beta = x_j$  to  $\gamma = x_{j-1}$  on  $F_{j-1}$  and  $G_{j-1}$ ) implies  $(\emptyset, \dots, \emptyset, F_{j-1}, \dots, F_n) \sim (\emptyset, \dots, \emptyset, G_{j-1}, \dots, G_n)$ . By induction,  $(F_1, \dots, F_n) \sim (G_1, \dots, G_n)$ .  $\square$

**LEMMA B.5** [Nested matching]. *Let  $A_1, A_2, \dots$ , be a finite or infinite sequence of events. Then there exists a nested  $\mathcal{A}$ -sequence of events  $E_1, E_2, \dots$  with  $E_j \sim_b A_j$  for all  $j$ , and  $E_i = E_j$  whenever  $A_i \sim_b A_j$  (so that  $E_i \supset E_j$  whenever  $A_i \succ_b A_j$ ).*

*Proof.* Define  $E_1 = A_1$ . For induction, assume an  $\mathcal{A}$ -sequence  $E_1, \dots, E_{j-1}$  constructed, and  $E_j$  to be constructed next. If

there is some  $i < j$  with  $A_i \sim_b A_j$ , then define  $E_j = E_i$  for any such  $i$ ; suppose henceforth that no such  $i$  exists. Define  $L = E_k$  with  $A_k$  the largest  $A_i$  “below”  $A_j$  for  $i < j$ . That is, if  $A_i <_b A_j$  for some  $i < j$  then define  $L = E_k$  with  $k$  such that  $A_k <_b A_j$  but  $A_k <_b A_i <_b A_j$  for no  $i < j$ . If no such  $i$  exists, then  $L = \emptyset$ . Define  $H = E_\ell$  with  $A_\ell$  the smallest  $A_i$  “above”  $A_j$  for  $i < j$ . That is, if  $A_i >_b A_j$  for some  $i < j$  then define  $H = E_\ell$  with  $\ell$  such that  $A_\ell >_b A_j$  but  $A_\ell >_b A_i >_b A_j$  for no  $i < j$ . If no such  $i$  exists, then  $H = S$ .  $H \succ_b A_j \succ_b L$ . By Lemma B.3, there exists  $E_j$  such that  $E_j \sim_b A_j$  and  $H \supset E_j \supset L$ , and  $H \setminus E_j$  and  $E_j \setminus L$  are events.  $E_1, \dots, E_j$  is an  $\mathcal{A}$ -sequence where the finest partition of  $S$ , generated by its corresponding algebra, has events  $H \setminus E_j$  and  $E_j \setminus L$  as elements, whereas that generated by  $E_1, \dots, E_{j-1}$  is coarser with an event  $H \setminus L = H \setminus E_j \cup E_j \setminus L$  as element. We define the  $E_j$  inductively. Because every finite subcollection thereof is an  $\mathcal{A}$ -collection, so is the collection of all  $E_j$ s.  $\square$

To use existing representation theorems from the literature, we extend  $\succ$  from the nested  $\mathcal{A}$ - $n$ -tuples to a binary relation, also denoted  $\succ$ , on  $\mathcal{R}$ , the set of *rank-ordered  $n$ -tuples of events*  $(E_1, \dots, E_n)$ , i.e. events with  $E_1 \preceq_b \dots \preceq_b E_n$  that need not be nested and, if nested, need not correspond with an act. For each such  $(E_1, \dots, E_n)$ , there exists an increasingly nested  $\mathcal{A}$ - $n$ -tuple  $(F_1, \dots, F_n)$  (so, it is an act) with  $F_j \sim_b E_j$  for all  $j$ , by Lemma B.5. By transitivity of  $\sim_b$  (Lemma B.2) and equivalence-dominance, all such nested  $n$ -tuples are equivalent. We can, therefore, define  $(E_1, \dots, E_n) \sim (F_1, \dots, F_n)$  for any such nested  $n$ -tuple, and extend  $\succ$  to all of  $\mathcal{R}$  as usual through the requirements of weak ordering. The definition of *equivalence-dominance* is extended to all of  $\mathcal{R}$  in the obvious manner. We have:

**LEMMA B.6**  $\succ$  on  $\mathcal{R}$  is a weak order satisfying equivalence-dominance.

The following elementary combinatorial result is given without proof.

**LEMMA B.7** [Equivalence of comonotonicity, compatibility with a joint rank-ordering, and jointly nested events]. *A set of acts  $\{(F_1^i, \dots, F_n^i)\}_{i \in I}$ , for an arbitrary index set  $I$ , is comono-*

tonic if and only if there exists a rank-ordering  $r$  of  $S$  such that all acts are compatible with  $r$ , which holds if and only if the collection of all events  $\{F_j^i\}_{j,i}$  is nested.

The following Corollary of Lemmas B.5, B.6, and B.7 is key to our generalization of existing representation theorems to mosaics of events. The corollary shows that, primarily because of solvability, the set of acts measurable with respect to the mosaic is still sufficiently rich to adopt proof techniques of rank-dependent representations on comonotonic sets, such as the set  $\mathcal{R}$ . As the corollary shows, every condition and theorem about countably many rank-ordered  $n$ -tuples is matched by a corresponding condition and theorem within a comonotonic set of acts.

**COROLLARY B.8.** *For each finite or infinite sequence of rank-ordered  $n$ -tuples  $\{(A_1^i, \dots, A_n^i)\}_{i=1, \dots}$ , there exists a set of  $n$ -tuples  $\{(A_\ell^{i'}, \dots, A_n^{i'})\}_{i=1, \dots}$ , such that the collection of all the primed events is a nested  $\mathcal{A}$ -collection,  $A_j^{i'} \sim_b A_j^i$  for all  $i, j$ , and  $A_j^i \succ_b A_\ell^k$  implies  $A_j^{i'} \supset A_\ell^k$ . All primed  $n$ -tuples correspond to acts, and the set of these acts is comonotonic, that is, it is compatible with one rank-ordering  $r$ . Equivalence-dominance implies that all preferences between primed and non-primed  $n$ -tuples are the same.*

In the remainder of this Section B.1, we demonstrate that  $\succ$  on  $\mathcal{R}$  inherits some properties from  $\succ$  over acts.

**LEMMA B.9.** [Strong monotonicity of  $\succ$  over acts with respect to  $\supset$ ]. *For all acts  $(F_1, \dots, F_n)$ ,  $(G_1, \dots, G_n)$ , we have  $(F_1, \dots, F_n) \succ (G_1, \dots, G_n)$  if  $F_j \supset G_j$  for all  $j$  and  $F_j \succ_b G_j$  for some  $j$ .*

*Proof.* We may assume that  $(F_1, \dots, F_n, G_1, \dots, G_n)$  is an  $\mathcal{A}$ -nested  $2n$ -tuple of events compatible with  $r$  for some rank-ordering  $r$  as in Corollary B.8. Hence, all  $n$ -tuples considered hereafter designate comonotonic acts. By monotonicity (Lemma B.1),  $(G_1, \dots, G_n) \preceq (G_1, \dots, G_{n-1}, F_n) \preceq (G_1, \dots, G_{n-2}, F_{n-1}, F_n) \preceq \dots \preceq (G_1, F_2, \dots, F_n) \preceq (F_1, \dots, F_n)$ . It suffices to establish strict preference when replacing  $G_j$  with  $F_j$ .  $(G_j : x_j, G_j^c : x_{j+1})$ ,  $(F_j : x_j, F_j^c : x_{j+1})$ ,  $(G_1, \dots, G_j, F_{j+1}, \dots, F_n)$ , and  $(G_1, \dots, G_{j-1}, F_j, \dots, F_n)$  are compatible with  $r$ .  $G_j \subset F_j$

and  $G_j \prec_b F_j$  imply that  $(F_j \setminus G_j)^{G_j}$  is nonnull and  $F_j \setminus G_j$  is nonnull on the comoncone for  $r$ . Comonotonic strong monotonicity (Lemma A.5) implies  $(G_1, \dots, G_j, F_{j+1}, \dots, F_n) \prec (G_1, \dots, G_{j-1}, F_j, \dots, F_n)$ .  $\square$

*Dominance* holds on  $\mathcal{R}$  if  $[A_j \succcurlyeq_b B_j \text{ for all } j]$  implies  $(A_1, \dots, A_n) \succcurlyeq (B_1, \dots, B_n)$ , where the latter preference is strict if one of the antecedent  $\succcurlyeq_b$  relationships is strict. Because of the latter requirement, this condition is stronger than the cumulative dominance axiom of Sarin and Wakker (1992). For  $(A_1, \dots, A_n) \in \mathcal{R}$ , and event  $B_i$ ,  $B_i(A_1, \dots, A_n)$  denotes the  $n$ -tuple  $(A_1, \dots, A_{i-1}, B_i, A_{i+1}, \dots, A_n)$ . It is implicit in this notation that the new  $n$ -tuple is in  $\mathcal{R}$  again, i.e.  $A_{i-1} \preccurlyeq_b B \preccurlyeq_b A_{i+1}$ .

#### COROLLARY B.10.

- (1) *Dominance holds.*
- (2) *If  $A_i F \succ E \succ C_i F$ , then  $B_i F \sim E$  and  $A \succcurlyeq_b B \succcurlyeq_b C$  for some event  $B$  ( $\mathcal{R}$ -solvability).*

*Proof.* (1) follows from Lemma B.9 and Corollary B.8.

- (2) We may assume that the  $n$ -tuples except  $B_i F$  (which is yet to be constructed) constitute a nested  $\mathcal{A}$ -collection (Corollary B.8). Relative to  $C_i F$ ,  $A_i F$  improves outcome  $x_{i+1}$  into outcome  $x_i$  on  $A \setminus C$ . We can let  $B$  be  $C \cup G$ , with  $G$  as in the solvability axiom and the substitutions  $A \setminus C$  for  $A$ ,  $x_{i+1}$  for  $\beta$ , and  $x_i$  for  $\gamma$ . (2) follows.

*Technical Observation.* The next point concerns proofs where, as in the above proof, new acts  $B_i F$  or events are constructed, usually through solvability, and will not be discussed explicitly again on future occasions. Although  $B_i F$  is comonotonic with the other acts, it need not be compatible with the rank-ordering  $r$  specified above. There does, however, exist a rank-ordering  $r'$  compatible with all original acts and also with  $B_i F$  (alternatively, one may apply Corollary B.8 anew at that stage), and for the rest of proofs it may then be desirable to replace  $r$  with  $r'$ . In such a case, there were several rank-orderings compatible with the original acts, and an arbitrary  $r$  originally chosen may turn out to be less suited for later purposes.  $\square$

LEMMA B.11 ( $\mathcal{R}$ -likelihood consistency). *If  $C \succ_b A$  and  $D \succ_b B$ , then*

$$\begin{aligned} \mathbf{A}_i F \sim \mathbf{B}_i G \quad \text{and} \quad \mathbf{A}_j Y \sim \mathbf{B}_j Z \quad \text{and} \\ \mathbf{C}_i F \sim \mathbf{D}_i G \\ \text{imply} \quad \mathbf{C}_j Y \sim \mathbf{D}_j Z. \end{aligned} \tag{B.1}$$

*Proof.* By Corollary B.8, we may assume that the  $n$ -tuples constitute a nested  $\mathcal{A}$ -collection, with  $C \supset A$  and  $D \supset B$ . The left two indifferences imply  $(C \setminus A)^r \sim (D \setminus B)^r$  for the appropriate  $r$  (improving  $\beta = x_{i+1}$  to  $\gamma = x_i$  on the respective events; see Figure B.1). The right two indifferences imply  $(C \setminus A)^r \sim (D \setminus B)^r$  for the same  $r$  (improving  $\beta = x_{j+1}$  to  $\gamma = x_j$  on the respective events). By clc, the last indifference must hold indeed.  $\square$

### B.2. Results on $\mathcal{R}$

In this section B.2, we will only consider  $\succ$  on  $\mathcal{R}$  with its properties derived in Section B.1, and not acts. For events  $A, B, C, D$ , we write  $[\mathbf{A}; \mathbf{B}] \sim^* [\mathbf{C}; \mathbf{D}]$  if  $\mathbf{A}_i F \sim \mathbf{B}_i G$  and  $\mathbf{C}_i F \sim \mathbf{D}_i G$  for some  $i$  and some rank-ordered  $n$ -tuples  $F, G$ . The coordinates of  $F$  and  $G$  different than the  $i$ th serve as a kind of gauge, or measuring rod, to show that the likelihood-difference between  $A$  and  $B$  is the same as that between  $C$  and  $D$ . *Tradeoff consistency* holds for  $\succ$  on  $\mathcal{R}$  if the implication of Equation (B.1) holds for all increasing  $n$ -tuples, without the restriction of  $[C \succ_b A \text{ and } D \succ_b B]$ . Cases that do not satisfy the latter restriction correspond to likelihood revelations for worsened outcomes (bets against events). The extension to such cases is complex.

LEMMA B.12 *Tradeoff consistency holds.*

*Proof.* For contradiction, assume that

$$\begin{aligned} \mathbf{A}_i F \sim \mathbf{B}_i G \quad \text{and} \quad \mathbf{A}_j Y \sim \mathbf{B}_j Z \quad \text{and} \\ \mathbf{C}_i F \sim \mathbf{D}_i G \quad \text{and} \quad \mathbf{C}_j Y \succ \mathbf{D}_j Z. \end{aligned} \tag{B.2}$$

*Case 1* [ $C \succ_b A$ ]. By dominance,  $C_i F \succ A_i F$ , which by transitivity implies  $D_i G \succ B_i G$ . Dominance implies  $D \succ_b B$ . The impossibility of Equation (B.2) now follows from Lemma B.11. Tradeoff consistency holds for this case.

*Case 2* [ $C \prec_b A$ ]. Similarly as above, we get  $D \prec_b B$ . We first modify the  $n$ -tuples so as to have  $Y_i \preceq_b Z_i$  for all  $i = j$ .

Assume that there is a  $k < j$  with  $Y_k \succ_b Z_k$ , and take the smallest such  $k$ . Then  $Y_k \succ_b Z_k \succ_b Z_{k-1} \succ_b Y_{k-1}$ , which will imply that the modified  $n$ -tuples  $A_j Y$  and  $C_j Y$  below will be rank-ordered indeed. If replacing  $Y_k$  by  $Z_k$  in the  $n$ -tuple  $C_j Y$  changes the strict preference  $C_j Y \succ D_j Z$  in (B.2) into a reversed preference, strict or weak, then by  $\mathcal{R}$ -solvability we can replace  $Y_k$  by  $Y'_k$  such that  $Y_k \succ_b Y'_k \succ_b Z_k$ , and such that the strict preference  $C_j Y \succ D_j Z$  in (B.2) changes into an equivalence. We also replace  $Y_k$  by  $Y'_k$  in the upper  $n$ -tuple  $A_j Y$  in (B.2), which by dominance generates a strict preference  $\prec$  between the resulting upper right  $n$ -tuples, so that, by Case 1 in this proof (with upper and lower preferences in (B.2) interchanged), a contradiction has resulted. We may, therefore, assume that replacing  $Y_k$  by  $Z_k$  in  $C_j Y$  leaves the strict preference  $C_j Y \succ D_j Z$  in (B.2) unaffected. The modified  $n$ -tuples are denoted as the original ones, and we write  $Y_k$  instead of  $Y'_k$ .

We continue the procedure just described until either an indifference and then a contradiction have resulted, or  $C_j Y \succ D_j Z$  for the modified  $C_j Y$ ,  $D_j Z$ , there is no  $k < j$  left with  $Y_k \succ_b Z_k$ , and  $A_j Y \preceq B_j Z$  (strict preference whenever  $Y_k \succ_b Z_k$  occurred for at least one  $k < j$  in the original situation) for these modified acts. We assume henceforth that the latter modification of Equation (B.2) holds.

Assume that there is a  $k > j$  with  $Y_k \succ_b Z_k$ ; take the largest such  $k$ . Then  $Z_{k+1} \succ Y_{k+1} \succ Y_k \succ Z_k$ , which will imply that the modified  $n$ -tuples below will be rank-ordered indeed. If replacing  $Z_k$  by  $Y_k$  in  $D_j Z$  changes the strict preference in (B.2) into a weak (possibly strict) reversed preference  $\preceq$ , then by  $\mathcal{R}$ -solvability we can replace  $Z_k$  by a  $Z_{k'}$  with  $Z_k \prec_b Z_{k'} \preceq_b Y_k$ , and such that the strict preference  $C_j Y \succ D_j Z$  in (B.2)

changes into an equivalence. We also replace  $Z_k$  by  $Z'_k$  in the upper  $n$ -tuple in (B.2), which by dominance generates or reinforces a strict reversed preference, so that, by Case 1 in this proof (with upper and lower preferences interchanged), a contradiction results. We may, therefore, assume that replacing  $Z_k$  by  $Y_k$  in  $D_jZ$  leaves the strict preference  $\mathbf{C}_jY \succ \mathbf{D}_jZ$  in (B.2) unaffected. We carry out this replacement for  $D_jZ$  and also for  $B_jZ$  in (B.2), strictly improving the latter. We continue this procedure until either a contradiction has resulted, or still  $\mathbf{C}_jY \succ \mathbf{D}_jZ$  in (B.2) for the modified  $\mathbf{C}_jY$  and  $\mathbf{D}_jZ$ , there is no  $k > j$  (or  $k < j$ ) left with  $Y_k \succ_b Z_k$ , and  $\mathbf{A}_jY \preceq \mathbf{B}_jZ$  in (B.2) for these modified  $n$ -tuples. We assume henceforth that the latter case holds. We are in the initial situation of (B.2), but now have  $Y_k \preceq_b Z_k$  for all  $k=j$  and possibly  $<$  instead of  $\sim$  between  $\mathbf{A}_jY$  and  $\mathbf{B}_jZ$ .

Because of dominance,  $C \succ_b D$ .  $C_jZ \succ C_jY \succ D_jZ$  (the former preference also by dominance) imply, by  $\mathcal{R}$ -solvability, existence of  $D'$  such that  $C \succ_b D' \succ_b D$  and  $D'_jZ \sim C_jY$ .  $D'_jZ \succ D_jZ$  implies  $D' \succ_b D$ .  $C_jY \prec A_jY$  implies that  $D'_jZ$  and  $C_jY$  are strictly less preferred than the upper right two  $n$ -tuples in (B.2), which by dominance implies  $D' \prec_b B$ . We replace  $D$  by  $D'$  in  $D_jZ$  and  $D_iG$ , maintaining the notation without primes, and ending up with:

$$\begin{aligned} \mathbf{A}_iF \sim \mathbf{B}_iG \quad \text{and} \quad \mathbf{A}_jY \preceq \mathbf{B}_jZ \quad \text{and} \\ \mathbf{C}_iF \prec \mathbf{D}_iG \quad \text{and} \quad \mathbf{C}_jY \sim \mathbf{D}_jZ, \end{aligned}$$

where still  $C \prec A$  and  $D \prec B$ .

We carry out a procedure for  $\mathbf{C}_iF$  and  $\mathbf{D}_iG$  similar to the one for  $\mathbf{C}_jY$  and  $\mathbf{D}_jZ$ . We now replace  $G_k$  by  $F_k$  if  $G_k \succ_b F_k$  for  $k < i$ , starting with the smallest such  $k$ , etc., as long as this does not affect the  $<$  preference. Each time,  $G_k \succ_b F_k \succ_b F_{k-1} \succ_b G_{k-1}$  ensures that newly constructed  $n$ -tuples are rank-ordered. If all such  $k$  have been exhausted, and still the  $<$  preference holds, we turn to the  $k > i$ . We replace  $F_k$  by  $G_k$  if  $F_k \prec_b G_k$  for  $k > i$ , starting with the largest such  $k$ , etc., as long as this does not affect the  $<$  preference. Each time,  $F_k \prec_b G_k \preceq_b G_{k+1} \preceq_b F_{k+1}$  ensures that newly constructed  $n$ -tuples are

rank-ordered. It is impossible that the two processes, first for  $k < i$  and then for  $k > i$ , can continue until all these  $k$ 's have been exhausted, because in the resulting  $C_i F < D_i G$  all left events would  $\succ_b$  dominate the right events, which violates dominance. The process must stop at some stage, at which through  $\mathcal{R}$ -solvability we obtain  $C_i F \sim D_i G$  for the newly constructed  $n$ -tuples. The replacements in  $F$  and  $G$  are also carried out in  $A_i F$  and  $B_i G$ , resulting in  $A_i F \succ B_i G$  (there was at least one replacement). We now have

$$\begin{aligned} A_i F \succ B_i G \quad \text{and} \quad A_j Y \preceq B_j Z \quad \text{and} \\ C_i F \sim D_i G \quad \text{and} \quad C_j Y \sim D_j Z, \end{aligned}$$

We consider replacing  $A$  by  $C$  in  $A_i F \succ B_i G$ . If  $C_i F \succ B_i G$ , then  $C_i F \succ B_i G \succ D_i G \sim C_i F$  (the second strict preference because  $B \succ_b D$ ) yields a contradiction. So,  $C_i F \preceq B_i G$  must hold. By  $\mathcal{R}$ -solvability,  $A'_i F \sim B_i G$  for an  $A'$  with  $C \preceq_b A' <_b A$ . We replace  $A$  by  $A'$  everywhere, drop the prime, and

$$\begin{aligned} A_i F \sim B_i G \quad \text{and} \quad A_j Y < B_j Z \quad \text{and} \\ C_i F \sim D_i G \quad \text{and} \quad C_j Y \sim D_j Z \end{aligned}$$

results. Because  $A \succ_b C$  and  $B \succ_b D$ , a contradiction with Case 1, with upper and lower preferences interchanged, has resulted.  $\square$

**COROLLARY B.13** *All tradeoff consistency axioms in Köbberling and Wakker (2003) could have been restricted to pairs  $\alpha, \beta$  with  $\alpha \succ \beta$  in  $*$ -relations.*

*Proof.* The proof of Lemma B.12 did not use the set-theoretic structure of events and, besides  $\text{clc}$ , only use weak ordering, monotonicity, and  $\mathcal{R}$ -solvability. All operations in the proof are not only compatible with comonotonicity, but also with the sign-dependence conditions considered in Köbberling and Wakker (2003).  $\square$

The  $\mathcal{R}$ -comonotonic sure-thing principle holds if  $C_i F \sim C_i G$  implies  $C'_i F \sim C'_i G$ .

**COROLLARY B.14** *The  $\mathcal{R}$ -comonotonic sure-thing principle holds.*

*Proof.* By tradeoff consistency (Lemma B.12), if

$$\begin{aligned} C_i F \sim C_i F \quad \text{and} \quad C_i F \sim C_i G \quad \text{and} \\ C'_i F \sim C'_i F \quad \text{then} \quad C'_i F \sim C'_i G. \end{aligned} \quad \square$$

We next turn to the Archimedean axiom. Note that, in the counterfactual event of an infinite sequence  $E_j$  in Figure A.1(a),  $\cup_{j < \infty} E_j$  and  $L$  need not be events if  $\mathcal{A}$  is not a  $\sigma$ -algebra, but this causes no problems. It suffices that  $\cup_{j \geq m} E_j \cup L$  is an event for each  $m$ . In what follows, we will usually let  $E^j$  be cumulative events rather than separate events as in the Archimedean axioms defined before. Thus,  $E^j$  hereafter will correspond with  $A \cup T \cup_{i=1}^{j-1} E_i$  as in Figure A.1, and  $E^{j+1*} = E^{j+1} \setminus E^j$  will correspond with  $E_j$  as in Figure A.1.

A *standard sequence* is a, finite or infinite, sequence of events  $E^1, E^2, \dots$  such that for some  $n$ -tuple  $F = (F_1, \dots, F_n)$ , some  $\ell = m$ , and some events  $A, B, E^{j+1} \ell A_m F \sim E^j \ell B_m F$  for all  $j$ , and either  $E^2 \succ_b E^1$  or  $E^2 \prec_b E^1$ . In the former case the standard sequence is *increasing*, in the latter case it is *decreasing*. The  $m^{\text{th}}$  coordinate is used as a kind of measuring rod. We have  $[E^{j+1}; E^j] \sim^* [E^j; E^{j-1}]$  for each  $j$ . The following lemma follows immediately from dominance.

**LEMMA B.15** *If the standard sequence with notation as above is increasing, then  $B \succ_b A$  and  $E^{j+1} \succ_b E^j$  for all  $j$ . If the standard sequence is decreasing, then  $B \prec_b A$  and  $E^{j+1} \prec_b E^j$  for all  $j$ .*  $\square$

We will derive the traditional Archimedean axiom, excluding infinite standard sequences, from the weaker version of the axiom in the following lemma. In many of the following results,  $m$ th and  $\ell$ th coordinates will be indicated through bold printing and semicolons.

LEMMA B.16 *There are no  $m < \ell$  and infinite increasing standard sequence  $(E^j)$  of the form*

$$\begin{aligned} &(\emptyset, \dots, \emptyset; \mathbf{B}; B, \dots, B; \mathbf{E}^j; S, \dots, S) \sim \\ &(\emptyset, \dots, \emptyset; \emptyset; B, \dots, B; \mathbf{E}^{j+1}; S, \dots, S). \end{aligned}$$

*Proof.* Note that the upper  $\ell$ th coordinate is  $B$ , and the lower is  $\emptyset$ . By Corollary B.8 it suffices to exclude such sequences for the case where all events involved are nested. This case follows immediately from the definition of the weak Archimedean axiom, by substituting

$$\begin{aligned} \varepsilon = x_m, \quad \delta = x_{m+1}, \quad \gamma = x_\ell, \quad \beta = x_{\ell+1}, \quad A = B, \quad T = E^1 \setminus B, \\ E_j = E^{j+1} \setminus E^j \text{ for all } j \geq 1, \quad L = S \setminus (\cup E^k). \end{aligned} \quad \square$$

OBSERVATION B.17 All standard sequences are finite.

*Proof.* The following proof uses two lemmas, ends of which are indicated by *QED*. Throughout this proof, we assume the notation as in the definition of a standard sequence  $E^j$ .

LEMMA B.18 *Assume that an infinite increasing standard sequence  $(E^j)$  exists. Then there exists an infinite decreasing standard sequence  $(D^j)$  with  $D^j {}_\ell A_m F \sim D^{j+1} {}_\ell B_m F$  for all  $j$ .*

*Proof.* We will let  $D^1$  be the  $\succ_b$  smallest upperbound for the  $E^j$ s that is present in the  $n$ -tuples of events. That is, if  $\ell = m - 1$  then  $D^1 = A <_b B$  ( $<_b$  by Lemma B.15); if  $\ell = n$  then  $D^1 = F_{\ell+1} = S$ ; in all other cases,  $D^1 = F_{\ell+1}$ . Assume, for induction, that for some  $k \geq 1$ , and all  $j \leq k$  we have defined  $D^j$  such that  $D^j {}_\ell B_m F \sim D^{j-1} {}_\ell A_m F$  if  $j \geq 2$  and  $D^j \succ E^i$  for all  $i$ . For each  $i$ ,  $E^i {}_\ell B_m F \sim E^{i+1} {}_\ell A_m F \preceq D^k {}_\ell A_m F \preceq D^k {}_\ell B_m F$ . By  $\mathcal{R}$ -solvability, there exists a  $D^{k+1, i}$  with  $E^i \preceq_b D^{k+1, i} \preceq_b D^k$  and  $D^{k+1, i} {}_\ell B_m F \sim D^k {}_\ell A_m F$ . Take any  $i'$  and define  $D^{k+1} = D^{k+1, i'}$ . By dominance, the indifference  $D^{k+1, i} {}_\ell B_m F \sim D^k {}_\ell A_m F$  implies  $D^{k+1} \sim_b D^{k+1, i}$  for all  $i$ . Therefore,  $D^{k+1} \succ E^i$  for all  $i$ . The lemma follows by induction. *QED*

LEMMA B.19 *Assume that an infinite decreasing standard sequence  $(E^j)$  with notation as above exists. Then there exists*

an infinite increasing standard sequence  $(D^j)$  with  $D^j \ell A_m F \sim D^{j+1} \ell B_m F$  for all  $j$ .

*Proof.* We will let  $D^1$  be the  $\succ_b$  largest lowerbound for the  $E^j$ s that is present in the  $n$ -tuples of events. That is, if  $\ell = m + 1$  then  $D^1 = A \succ_b B$ ; if  $\ell = 1$  then  $D^1 = F_{\ell-1} = \emptyset$ ; in all other cases,  $D^1 = F_{\ell-1}$ . Assume, for induction, that for some  $k \geq 1$ , and all  $j \leq k$  we have defined  $D^j$  such that  $D^{j-1} \ell A_m F \sim D^j \ell B_m F$  if  $j \geq 2$  and  $D^j \preceq E^i$  for all  $i$ . For each  $i$ ,  $E^i \ell B_m F \sim E^{i+1} \ell A_m F \succ D^k \ell A_m F \succ D^k \ell B_m F$ . By  $\mathcal{R}$ -solvability, there exists a  $D^{k+1,i}$  with  $E^i \succ_b D^{k+1,i} \succ_b D^k$  and  $D^{k+1,i} \ell B_m F \sim D^k \ell A_m F$ . Take any  $i'$  and define  $D^{k+1} = D^{k+1,i'}$ . By dominance, the indifference  $D^{k+1,i} \ell B_m F \sim D^k \ell A_m F$  implies  $D^{k+1} \sim_b D^{k+1,i}$  for all  $i$ . Therefore,  $D^{k+1} \preceq E^i$  for all  $i$ . The lemma follows by induction. *QED*

Assume, for contradiction, an infinite standard sequence  $(E^j)$  with  $E^{j+1} \ell A_m F \sim E^j \ell B_m F$  for all  $j$ . By the  $\mathcal{R}$ -comonotonic sure-thing principle (Corollary B.14), we may replace  $F_j$  by  $\emptyset$  for all  $j$  smaller than  $\ell$  and  $m$ , and  $F_j$  by  $S$  for all  $j$  exceeding  $\ell$  and  $m$ . These coordinates will be suppressed throughout this proof. In  $n$ -tuple notation, the  $\ell$ th and  $m$ th coordinate will be indicated by bold-printing and semicolons throughout.

*Case 1* [ $m < \ell$ ]. That is, the “measuring-rod” events  $A$  and  $B$  are  $\succ_b$  smaller than the standard sequence, as in Lemma B.16. We may assume, by Lemma B.19, that the standard sequence is increasing. By the  $\mathcal{R}$ -comonotonic sure-thing principle, we may assume that  $F_{m+1} = \dots = F_{\ell-1} = B$  ( $A \preceq_b B \preceq_b E^j$  for all  $j$  so that all  $n$ -tuples are rank-ordered). The following reasoning shows that we may assume that  $A = \emptyset$ .

We have  $(\emptyset; B, \dots, B; \mathbf{E}^1) \preceq (\emptyset; B, \dots, B; \mathbf{E}^2) \preceq (\mathbf{A}; B, \dots, B; \mathbf{E}^2) \sim (\mathbf{B}; B, \dots, B; \mathbf{E}^1)$ . By  $\mathcal{R}$ -solvability, there exists an event  $B'$  with  $\emptyset \preceq B' \preceq B$  and  $(\mathbf{B}'; B, \dots, B; \mathbf{E}^1) \sim (\emptyset; B, \dots, B; \mathbf{E}^2)$ . Assume, for induction, that  $(\emptyset; B, \dots, B; \mathbf{E}^{j+1}) \sim (\mathbf{B}'; B, \dots, B; \mathbf{E}^j)$  for some  $j$ . Because  $[E^{j+2}; E^{j+1}] \sim^* [E^{j+1}; E^j]$ , it follows from tradeoff consistency that  $(\emptyset; B, \dots, B; \mathbf{E}^{j+2}) \sim (\mathbf{B}'; B, \dots, B; \mathbf{E}^{j+1})$ . By induction, this indifference holds for all  $j$ . We may indeed set  $A = \emptyset$  and  $B = B'$ . A contradiction with Lemma B.16 has resulted, and Case 1 cannot occur.

*Case 2* [ $m > \ell$ ]. That is, the “measuring-rod” events  $A$  and  $B$  are  $\succ_b$  ranked bigger than the standard sequence. If the standard sequence is increasing, we can obtain a decreasing standard sequence by applying Lemma B.18. If the standard sequence is decreasing, we can obtain an increasing standard sequence by applying Lemma B.19, and from that a decreasing standard sequence again by applying Lemma B.18. In each case, we have obtained an infinite decreasing standard sequence  $(E^j)$  (so that  $A \succ_b B$ ) of the kind constructed in the proof of Lemma B.18. By the  $\mathcal{R}$ -comonotonic sure-thing principle, we may assume that  $F_{\ell+1} = \dots = F_{m-1} = B$  (the  $\succ_b$ -smaller of  $A$  and  $B$ ) for each decreasing standard sequence. We may further assume that  $E^1 = B$ . This can be inferred from the proof of Lemma B.18, where the first element of the standard sequence is the smallest event in the  $n$ -tuples  $\succ_b$ -dominating the elements of the standard sequence, which is  $B$  in our case.

We end up with an infinite decreasing standard sequence  $(E^j)$  with  $(\mathbf{E}^{j+1}; B, \dots, B; \mathbf{A}) \sim (\mathbf{E}^j; B, \dots, B; \mathbf{B})$  for all  $j \in \mathbb{N}$ ,  $E^1 = B$ , and then an infinite increasing standard sequence  $(D^j)$  as constructed in the proof of Lemma B.19, satisfying  $(\mathbf{D}^{j+1}; B, \dots, B; \mathbf{B}) \sim (\mathbf{D}^j; B, \dots, B; \mathbf{A})$  for all  $j$  and  $D^1 = \emptyset$ . Further,  $D^j \preceq_b E^i$  for all  $j, i$ .

We distinguish two cases.

*Case 2.1*  $[(\mathbf{D}^2; D^2, \dots, D^2; \mathbf{E}^2) \succ (\emptyset; D^2, \dots, D^2; \mathbf{E}^1)]$ .<sup>2</sup> Adding  $(\emptyset; D^2, \dots, D^2; \mathbf{E}^1) \succ (\emptyset; D^2, \dots, D^2; \mathbf{E}^2)$  and applying  $\mathcal{R}$ -solvability gives an event  $T$  with  $D^2 \succ_b T \succ_b \emptyset$  and

$$(\mathbf{T}; D^2, \dots, D^2; \mathbf{E}^2) \sim (\emptyset; D^2, \dots, D^2; \mathbf{E}^1). \quad (\text{B.3})$$

For all  $j$ , we have  $[E^{j-1}; E^j] \sim^* [E^j; E^{j+1}]$  because  $(E^j)$  is a standard sequence. An antecedent indifference

$$(\mathbf{T}; T, \dots, T; \mathbf{E}^j) \sim (\emptyset; T, \dots, T; \mathbf{E}^{j-1}),$$

implies, by tradeoff consistency,

$$(\mathbf{T}; T, \dots, T; \mathbf{E}^{j+1}) \sim (\emptyset; T, \dots, T; \mathbf{E}^j).$$

For  $j=2$ , the antecedent indifference follows from (B.3) and the  $\mathcal{R}$ -comonotonic sure-thing principle. By induction, the

antecedent and implied indifferences hold for all  $j$ . We have constructed an infinite standard sequence of the kind excluded by Case 1.

*Case 2.2*  $[(\mathbf{D}^2; D^2, \dots, D^2; \mathbf{E}^2) \preceq (\emptyset; D^2, \dots, D^2; \mathbf{E}^1)]$ .<sup>3</sup> Adding  $(\emptyset; D^2, \dots, D^2; \mathbf{E}^1) \preceq (\mathbf{D}^2; D^2, \dots, D^2; \mathbf{E}^1)$  and  $\mathcal{R}$ -solvability imply the existence of an event  $E^{2'}$  with  $E^2 \preceq_b E^{2'} \preceq_b E^1$  such that  $(\mathbf{D}^2; D^2, \dots, D^2; \mathbf{E}^{2'}) \sim (\emptyset; D^2, \dots, D^2; \mathbf{E}^1)$ .  $(\mathbf{E}^2; B, \dots, B; \mathbf{A}) \sim (\mathbf{E}^1; B, \dots, B; \mathbf{B})$  and  $E^{2'} \succ_b E^2$  imply  $(\mathbf{E}^{2'}; B, \dots, B; \mathbf{A}) \succ (\mathbf{E}^1; B, \dots, B; \mathbf{B})$ . Adding  $(\mathbf{E}^1; B, \dots, B; \mathbf{B}) \succ (\mathbf{E}^{2'}; B, \dots, B; \mathbf{B})$  and  $\mathcal{R}$ -solvability imply existence of  $A'$  such that  $A \succ_b A' \succ B$  and  $(\mathbf{E}^{2'}; B, \dots, B; \mathbf{A}') \sim (\mathbf{E}^1; B, \dots, B; \mathbf{B})$ . Define  $E^{1'} = E^1$ . Suppose for induction that, for some  $k$  and for all  $j \leq k$ , we have constructed  $E^{j'}$  such that  $E^{j'} \succ E^j$  and  $(\mathbf{E}^{j'}; B, \dots, B; \mathbf{A}') \sim ((\mathbf{E}^{j-1})'; B, \dots, B; \mathbf{B})$ .

We have  $(\mathbf{E}^{k+1}; B, \dots, B; \mathbf{A}') \preceq (\mathbf{E}^{k+1}; B, \dots, B; \mathbf{A}) \sim (\mathbf{E}^k; B, \dots, B; \mathbf{B}) \preceq (\mathbf{E}^{k'}; B, \dots, B; \mathbf{B}) \preceq (\mathbf{E}^{k'}; B, \dots, B; \mathbf{A}')$ . By  $\mathcal{R}$ -solvability, there exists an event  $(E^{k+1})'$  with  $E^{k+1} \preceq_b (E^{k+1})' \preceq_b E^{k'}$  and  $((\mathbf{E}^{k+1})'; B, \dots, B; \mathbf{A}') \sim (\mathbf{E}^{k'}; B, \dots, B; \mathbf{B})$ .

By induction,  $E^{j'}$  can be defined for all  $j \in \mathbb{N}$  and the standard sequence is infinite. As a next induction hypothesis, assume that  $(\mathbf{D}^2; D^2, \dots, D^2; \mathbf{E}^j) \sim (\emptyset; D^2, \dots, D^2; (\mathbf{E}^{j-1})')$ . This,  $[E^{j'}; (E^{j+1})'] \sim^* [(E^{j-1})'; E^{j'}]$  (because the  $(E^{j'})$  are a standard sequence) and tradeoff consistency, imply that  $(\mathbf{D}^2; D^2, \dots, D^2; (\mathbf{E}^{j+1})') \sim (\emptyset; D^2, \dots, D^2; \mathbf{E}^j)$ . Because the induction hypothesis holds true for  $j=2$ , it holds true for all  $j \geq 2$ . We have constructed a standard sequence of the kind excluded by Case 1. The proof of Observation B.17 is complete.  $\square$

**COROLLARY B.20** *The Archimedean axioms in Köbberling and Wakker (2003, Theorem 8) could have been restricted to increasing standard sequences as in Lemma B.16.*

*Proof.* The derivation of Observation B.17 did not use the set-theoretic structure on events. The presence of a minimum  $\emptyset$  and a maximum  $S$  is not crucial because the Archimedean axiom need only be imposed on bounded sets.  $\square$

At this stage, we want to use Theorem 8 of Köbberling and Wakker (2003). All characterizing conditions of this theorem are satisfied. The only remaining problem is that the mentioned theorem is stated only for full product sets, whereas our domain is a rank-ordered subset of a full product set. We, therefore, adapt our domain as follows. Define  $\mathcal{R}'$  as the  $n$ -fold product of events, where the  $n$ -tuples need not be  $\succsim_b$ -rank-ordered. For an  $n$ -tuple  $(A_1, \dots, A_n)$ , define by  $(A'_1, \dots, A'_n)$  a corresponding rank-ordered  $n$ -tuple obtained through permutation such that  $A'_1 \preccurlyeq_b \dots \preccurlyeq_b A'_n$ . For example, if  $A_1 \succsim_b \dots \succsim_b A_n$ , then  $A'_1 = A_n, A'_2 = A_{n-1}, \dots, A'_n = A_1$  can be taken. Extend the binary relation  $\succsim$  from  $\mathcal{R}$  to  $\mathcal{R}'$ , denoted by the same symbol, by setting  $(A_1, \dots, A_n) \sim (A'_1, \dots, A'_n)$  for all  $n$ -tuples, and transitive extension. If  $A_i \sim_b A_j$  for some  $i = j$ , then there are several rank-ordered  $n$ -tuples corresponding with  $(A_1, \dots, A_n)$ . By equivalence-dominance they are all indifferent and the extended binary relation is, therefore, well-defined and unique.

Weak ordering of  $\succsim$  on  $\mathcal{R}'$  holds by definition, and dominance is naturally extended to  $\mathcal{R}'$ . Dominance implies the weak and strong monotonicity conditions of Köbberling and Wakker (2003). For the comonotonic preference conditions of the latter reference, we define, for each rank-ordering  $\rho$  on  $\{1, \dots, n\}$ ,  $\mathcal{R}'^\rho \subset \mathcal{R}'$  as the *comoncone* of  $\rho$ , containing all  $n$ -tuples  $(A_1, \dots, A_n)$  with  $A_{\rho(1)} \preccurlyeq_b \dots \preccurlyeq_b A_{\rho(n)}$ . It follows straightforwardly that the standard sequences within each comoncone must be finite. Further, for any violation of Equation (B.1) with the left  $n$ -tuples  $\{A_i F, B_i G, C_i F, D_i G\}$  in one comoncone and the right  $n$ -tuples  $\{A_j Y, B_j Z, C_j Y, D_j Z\}$  in a possibly different comoncone, and without the restriction that  $C \succsim A$  and  $D \succsim B$ , we can reorder the events for the left quadruple of  $n$ -tuples, and then, possibly differently, those for the right quadruple of  $n$ -tuples, such that all resulting  $n$ -tuples are contained in  $\mathcal{R}$ . Hence, such violations are excluded by our tradeoff consistency. By Corollary 28 of Köbberling and Wakker, their comonotonic tradeoff consistency follows. Köbberling and Wakker require solvability to hold globally, not just within comoncones:  $\mathcal{R}'$ -solvability holds if  $A_i F \succ G \succ$

$C_i F$  implies existence of an event  $B$  with  $A \succ_b B \succ_b C$  and  $B_i F \sim G$ .

LEMMA B.21  $\mathcal{R}'$ -solvability holds.

*Proof.* Assume  $A_i F \succ G \succ C_i F$ . Preferences are not affected by  $n$ -tuple-dependent permutations of coordinates. We may, therefore, assume that  $F$  and  $G$  are elements of  $\mathcal{R}$ , i.e.  $F_1 \preceq_b \dots \preceq_b F_n$  and  $G_1 \preceq_b \dots \preceq_b G_n$ . Because  $S_i F \succ A_i F \succ G \succ C_i F \succ \emptyset_i F$  we can, with  $F_{n+1} = S$  and  $F_0 = \emptyset$  as usual, define  $1 \leq j \leq n+1$  such that  $(F_j)_i F \succ G \succ (F_{j-1})_i F$ . We can ( $n$ -tuples-independently) permute coordinates such that  $(F_j)_i F$  and  $(F_{j-1})_i F$  end up in  $\mathcal{R}$ , with the, originally  $i$ th, coordinates  $F_j$  and  $F_{j-1}$  now at a new coordinate  $k$ , ranked between the original  $j-1$  and  $j$ . By  $\mathcal{R}$ -solvability, an event  $B$  can be found with  $F_{j-1} \preceq_b B \preceq_b F_j$  and  $B_k F \sim G$ . Inverting the last permutation of coordinates,  $B_i F \sim G$ . By dominance,  $A \succ_b B \succ_b C$ .  $\square$

We have verified that  $\succ$  on  $\mathcal{R}'$  satisfies  $\mathcal{R}'$ -solvability, weak ordering, dominance (hence, weak monotonicity of Köbberling and Wakker 2003), the comonotonic Archimedean axiom of Köbberling and Wakker (2003), and comonotonic trade-off consistency. By their Theorem 8, we obtain a Choquet expected utility representation for  $\succ$  on  $\mathcal{R}'$ . We will only use this representation on  $\mathcal{R}$ . On this rank-ordered set, the Choquet expected utility representation is in fact a SEU representation. That is, there exist nonnegative  $p_1^d, \dots, p_n^d$  that sum to one, and a function  $U^d$ , such that  $(A_1, \dots, A_n) \mapsto p_1^d U^d(A_1) + \dots + p_n^d U^d(A_n)$  represents  $\succ$  on  $\mathcal{R}$ . We can normalize  $U^d$  such that  $U^d(\emptyset) = 0$  and  $U^d(S) = 1$ . The superscript  $d$  refers to dual. We define  $W(A) = U^d(A)$  for all events  $A$ ,  $U(x_{n+1}) = 0$ , and  $U(x_j) = p_j^d + \dots + p_n^d$ . Re-arranging terms, we see that, for nested  $(A_1, \dots, A_n)$  corresponding with an act,  $(A_1, \dots, A_n) \mapsto p_1^d U^d(A_1) + \dots + p_n^d U^d(A_n)$  agrees with the Choquet expected utility of the act  $(A_1^*: x_1, A_1^*: x_2, \dots, A_1^*: x_n, A_n^c: x_{n+1})$  under  $W$  and  $U$ , as can easily be seen from Figure B.1 with  $W$ - and  $U$ -transformed axes.

For uniqueness, given  $W(\emptyset) = 0$  and  $W(S) = 1$ ,  $U^d$  and, hence,  $W$ , are uniquely determined. A function  $U^*$  can serve as utility function if and only if  $\frac{U^*(\cdot) - U^*(x_{n+1})}{U^*(x_1) - U^*(x_{n+1})}$  agrees with  $U$  as constructed.  $U$  is unique up to unit and origin. The proof of Theorems 5.7 and A.3 for finite outcome sets is now complete.

**COROLLARY B.22** *Theorem 8 of Köbberling and Wakker (2003) is also valid if the domain is a comoncone instead of a full product set.*

#### APPENDIX C. SUFFICIENCY PROOF OF THEOREMS 5.7 AND A.3 FOR INFINITELY MANY OUTCOMES; AND FURTHER REMARKS

We continue to assume (ii) of Theorem 5.7 in the weakened version of Theorem A.3, and derive CEU plus the uniqueness results. We now assume that  $\mathcal{C}$  has infinitely many equivalence classes. Fix some outcomes  $\sigma \succ \mu \succ \lambda$ . We set  $U(\alpha) = 1$  for all  $\alpha \sim \mu$ , and  $U(\alpha) = 0$  for all  $\alpha \sim \lambda$ . We apply the analysis of Appendix B to acts that only yield the outcomes  $\sigma, \mu, \lambda$ . On this subset, all axioms of Theorem 5.7, including solvability, remain valid, and we obtain a CEU representation and a unique capacity  $W$  on events there. For each outcome  $\alpha$  not equivalent to  $\mu$  and  $\lambda$ , we apply Appendix B to acts yielding only  $\alpha, \sigma, \mu$ , and  $\lambda$ , obtaining a CEU representation there. We always set utility zero at  $\lambda$  and 1 at  $\mu$ , so that the resulting utility for outcome  $\alpha$  is uniquely determined. This is defined as  $U(\alpha)$ .

As we will now show, the capacity and utility  $U$  as just defined give a CEU representation for all acts. Consider any preference between two acts. For the set of all outcomes considered in these two acts, with  $\mu, \lambda$ , and  $\sigma$  added, we consider all acts yielding only those outcomes, for which we can apply Appendix B and get a CEU representation. Because this CEU representation also covers acts only yielding  $\mu, \lambda$ , and  $\sigma$ , its capacity must be identical to the one already defined. Because

this CEU representation also covers acts only yielding  $\sigma, \mu, \lambda$  and  $x_i$  for each outcome  $x_i$  of the two acts considered, its utility function must coincide with the one already defined. We have obtained a CEU representation for all acts.

For uniqueness, note that for each CEU representation we can rescale utility to be 0 at  $\lambda$  and 1 at  $\mu$ , after which it must coincide with the one constructed above. The capacity must coincide with the one constructed above from mere inspection of  $\sigma, \mu$ , and  $\lambda$ . The proof of Theorems 5.7 and A.3 is complete.  $\square$

The expected utility representation that we obtained on  $\mathcal{R}$  at the end of Appendix B is a special case of the additive representation on rank-ordered sets examined by Wakker (1991). Two cases can be distinguished. The *densely-spaced* case holds if for all events  $A \succ_b C$  there exists an event B such that  $A \succ_b B \succ_b C$ . By Eq. 16 in Wakker (1991), the range of  $W$  then is dense in  $[0,1]$ , and finite state spaces and atoms are excluded. We discuss the other case in more detail.

If the densely-spaced case does not hold, then the *equally-spaced* case holds. It means that there exist events  $A \succ_b C$  such that  $A \succ_b B \succ_b C$  for no event B. Corollary 7 of Wakker (1991) then applies. In the notation of Appendix B, then  $p_j^d = 1/n$  for each  $j$ , and the outcomes  $x_1, \dots, x_{n+1}$  are equally-spaced in utility units. Given that  $\emptyset$  and  $S$  are bounds for the events, there must exist  $m \in \mathbb{N}$  such that the range of  $U^d = \{0, 1/m, \dots, (m-1)/m, 1\}$ . Then  $W$  takes all values  $j/m$  ( $0 \leq j \leq m$ ), and no other value. Example 5.8(i) illustrated that all of our axioms can be satisfied for the equally spaced case with no more than three outcomes. It turns out that, without modification of solvability, no other cases are possible.

**OBSERVATION C.1** Assume the equally-spaced case. Then there can be no more than three equivalence classes of outcomes.

*Proof.* Assume four outcomes  $x_1 \succ x_2 \succ x_3 \succ x_4$  with  $u(x_j) = 4 - j$  for all  $j$ . There is an  $m \in \mathbb{N}$  such that for each event  $A$ ,  $W(A) = j/m$  for an integer  $j \geq 0$ . We can take disjoint

events  $A$  and  $B$  such that  $W(A) = 1/m = \pi(B^A)$  (Lemma C.2 hereafter). Let  $L = (A \cup B)^c$ .  $(A : x_1, B : x_4, L : x_4) \succ (A : x_3, B : x_3, L : x_4) \succ (A : x_3, B : x_4, L : x_4)$ , but there exists no subevent of  $A$  with capacity value between 0 and  $1/m$ , in particular not with capacity value  $1/2m$ . Hence,  $(A_1 : x_1, A_3 : x_3, B : x_4, L : x_4) \sim (A : x_3, B : x_3, L : x_4)$  for no partition  $\{A_1, A_3\}$  of  $A$ , and solvability is violated.  $\square$

Our solvability axiom is stronger than the (dual of) Wakker's (1991) solvability axiom. The latter axiom was imposed on coordinates that, for the equally spaced case in our dual version, correspond with smallest increments of outcomes, so that only outcomes  $\gamma \succ \beta$  should be considered for which there is no outcome  $\mu$  with  $\gamma \succ \mu \succ \beta$ . This explains the extra restrictions described in the above observation. Finite state spaces with more than three outcomes can be incorporated if the solvability axiom is relaxed and imposed only on outcomes  $\gamma, \beta$  as above. For the proof in Appendix C, extending from finitely many equivalence classes of outcomes to countably many, then only finite subsets of outcomes should be considered from consecutive equivalence classes. A drawback of this relaxed solvability is that it does not apply to the densely spaced case, leading to two different axioms for the two cases.

A solvability condition that fully covers both the equally spaced case and the densely spaced case, is *weakened solvability*: The implication in Definition 5.2 then is required only for  $\gamma \succ \beta$  such that, for all  $\mu$  with  $\gamma \succ \mu \succ \beta$ , there exists a  $\mu'$  with  $\gamma \succ \mu \succ \mu' \succ \beta$ . This axiom, indeed, allows for all outcome sets  $\mathcal{C}$  that are subintervals (even if unbounded) of the integers, and also allows for the densely spaced case.

In the proof in Appendix B, we used only weakened solvability, not solvability, and in the following observation we will do so likewise. Solvability in the main text was used for its simplicity. The following lemmas illustrate the restrictions imposed by solvability, and prepare for later proofs. Lemma C.2 concerns (rank-dependent) atoms. Lemma C.3 concerns the nonatomic case, generalizing Gilboa's (1987) convex-rangedness.

LEMMA C.2 *Assume the equally-spaced case, with  $\{j/m : 0 \leq j \leq m\}$  the range of  $W$ . For each nonnull ranked event  $E^R$  a partition  $E_1, \dots, E_k$  of  $E$  exists such that  $\pi(E_j^{\cup_{i=1}^{j-1} E_i}) = 1/m$  for all  $j$ .*

*Proof.* As explained before,  $W(R) = i/m$  and  $W(E \cup R) = j/m$  for some  $i \leq j$ . If  $j = i + 1$ , then we simply take  $k = 1$  and  $E_1 = E$ . For induction, assume that the result has been proved for all  $j \leq i + \ell$  for some  $\ell$ , and consider the case  $j = i + \ell + 1$ . Let  $\gamma \succ \beta$ . Write  $L = (E \cup R)^c$ . Take an arbitrary event  $A$  with  $W(A) = (i + \ell)/m$ , and apply solvability to  $(R : \gamma, E : \beta, L : \beta) \preceq (A : \gamma, A^c : \beta) \preceq (R : \gamma, E : \gamma, L : \beta)$  to obtain an event  $E' \subset E$  with  $W(R \cup (E \setminus E')) = W(A) = (i + \ell)/m$ , and  $E \setminus E'$  also an event. We define  $k = \ell + 1$ ,  $E_k = E'$ , and apply the induction hypothesis to  $E \setminus E'$  to obtain events  $E_1, \dots, E_{k-1}$ .  $\square$

LEMMA C.3 *Assume the densely-spaced case. For every non-null ranked event  $E^R$ , and every  $\varepsilon > 0$ , a partition  $E_1, \dots, E_k$  of  $E$  exists such that  $0 < \pi(E_j^{\cup_{i=1}^{j-1} E_i}) < \varepsilon$  for all  $j$ . Consequently, for each ranked event  $E^R$  the set  $\pi(A^R) : A \subset E$  is dense in  $[0, \pi(E^R)]$ .*

*Proof.* If  $\pi(E^R) < \varepsilon$ , then we simply take  $k = 1$  and  $E_1 = E$ . For induction, assume that the result has been proved for all  $\pi(E^R) < \ell\varepsilon$  for some  $\ell \geq 1$ , and consider the case  $\ell\varepsilon \leq \pi(E^R) < (\ell + 1)\varepsilon$ . Let  $\gamma \succ \beta$ . Write  $L = (E \cup R)^c$ . Because, as explained before, the range of  $W$  is dense in  $[0, 1]$ , we can take an event  $A$  with  $W(R \cup E) - \varepsilon < W(A) < W(R) + \ell\varepsilon$ . We apply solvability to  $(R : \gamma, E : \beta, L : \beta) \preceq (A : \gamma, A^c : \beta) \preceq (R : \gamma, E : \gamma, L : \beta)$  to obtain an event  $E' \subset E$  with  $W(R \cup (E \setminus E')) = W(A)$ , and  $E \setminus E'$  also an event.  $\pi(E_j^{\cup_{i=1}^{j-1} E_i}) = W(R \cup E) - W(A) < \varepsilon$  follows. We define  $k = \ell + 1$ ,  $E_k = E'$ , and apply the induction hypothesis to  $E \setminus E_k$  to obtain  $E_1, \dots, E_{k-1}$ .  $\square$

#### APPENDIX D. FURTHER PROOFS

A useful *notation* for rank-dependence is as follows. We write  $A^b$  for  $A^\emptyset$ , where  $A$  is ranked best, and  $A^w$  for  $A^{A^c}$ , where  $A$  is ranked worst.

*Proof of Theorem 5.3* The implication (i)  $\Rightarrow$  (ii) follows from substitution. We, therefore, assume the preference conditions in (ii) and derive SEU plus the uniqueness results. The preference conditions in (ii) of Theorem 5.7 follow immediately, so that we obtain the CEU representation of Theorem 5.7. We will use the following implication of likelihood consistency, a partial sure-thing principle, restricted to indifferences. The lemma extends Corollary B.14.

**LEMMA D.1**  $\beta_A f \sim \beta_A g$  implies  $\gamma_A f \sim \gamma_A g$  for all  $\gamma > \beta$  and acts  $\beta_A f, \beta_A g, \gamma_A f$ , and  $\gamma_A g$ .

*Proof.* The trivial  $\beta_A f \sim \beta_A f$  and  $\gamma_A f \sim \gamma_A f$  imply  $A \sim A$ .  $\beta_A f \sim \beta_A g$  and  $\gamma_A f > \gamma_A g$  would imply  $A > A$ , which, combined with  $A \sim A$ , would violate likelihood consistency. Similarly,  $\beta_A g \sim \beta_A f$  and  $\gamma_A g > \gamma_A f$  would violate likelihood consistency.  $\beta_A f \sim \beta_A g$  must imply  $\gamma_A f \sim \gamma_A g$ . *QED*

*Case 1* [the densely-spaced case]. We prove that  $W(A \cup B) = W(A) + W(B)$  for all disjoint  $A, B$  with  $A \cup B$  an event.

*Case 1.1* [ $W(A \cup B) < 1$ ]. Take three outcomes  $x_1 > x_2 > x_3$ . We write  $L$  for  $(A \cup B)^c$ . By Lemma C.3, we can partition event  $B$  into events  $\{B_1, \dots, B_m\}$  such that, for all  $i$ ,

$$\pi(B_i^{\cup_{j=1}^{i-1} B_j})(U(x_1) - U(x_2)) \leq \pi(L^w)(U(x_2) - U(x_3));$$

here  $\{B_1, \dots, B_m, A, L\}$  is an  $\mathcal{A}$ -partition because all of its elements are events. The inequality means that, in the ranking  $r$  ("ranked better than"),  $B_1 r \dots r B_m r A r L$ , the decision weight of each event  $B_i$  is small relative to that of  $L$ . For each  $i$  the above inequality implies the first preference in  $(\cup_{j=1}^{i-1} B_j : x_1, B_i : x_1, \cup_{j=i+1}^m B_j : x_2, A : x_2, L : x_3) \preceq (\cup_{j=1}^{i-1} B_j : x_1, B_i : x_2, \cup_{j=i+1}^m B_j : x_2, A : x_2, L : x_2) \preceq (\cup_{j=1}^{i-1} B_j : x_1, B_i : x_1, \cup_{j=i+1}^m B_j : x_2, A : x_2, L : x_2)$ .

By solvability, there exists a partition  $\{L_{i2}, L_{i3}\}$  of  $L$  such that (also reranking  $A$ )

$$(\cup_{j=1}^{i-1} B_j : x_1, B_i : x_1, A : x_2, \cup_{j=i+1}^m B_j : x_2, L_{i2} : x_2, L_{i3} : x_3) \sim (\cup_{j=1}^{i-1} B_j : x_1, A : x_2, B_i : x_2, \cup_{j=i+1}^m B_j : x_2, L : x_2).$$

By Lemma D.1,

$(\cup_{j=1}^{i-1} B_j : x_1, B_i : x_1, A : x_1, \cup_{j=i+1}^m B_j : x_2, L_{i2} : x_2, L_{i3} : x_3) \sim (\cup_{j=1}^{i-1} B_j : x_1, A : x_1, B_i : x_2, \cup_{j=i+1}^m B_j : x_2, L : x_2)$ .

The last two indifferences show that  $A^\rho \sim A^{\rho'}$  where  $\rho$  corresponds to the two acts left of the  $\sim$  sign, and  $\rho'$  to the right two acts. Hence,  $\rho$  ranks  $A$  worse than  $B_i$  and  $\rho'$  ranks  $A$  better. By CEU (similar to Observation 5.6),  $A$  receives the same decision weight in both rankings, that is,  $\pi(A^{\cup_{j=1}^{i-1} B_j}) = \pi(A^{\cup_{j=i}^m B_j})$ . By induction,  $\pi(A^\emptyset) = \pi(A^b) = \pi(A^{\cup_{j=1}^m B_j}) = \pi(A^B)$ , i.e.  $W(A) = W(A \cup B) - W(B)$ , and the additivity requirement has been established.

*Case 1.2* [ $W(A \cup B) = 1$ ]. We show that  $W(A) + W(B) = 1$  for disjoint  $A, B$ . Take some  $\varepsilon > 0$ . By Lemma C.3, if  $W(A) < 1$  then there is  $B' \subset B$  with  $1 > W(A \cup B') > 1 - \varepsilon$ . By Case 1.1, then  $W(A) + W(B') > 1 - \varepsilon$ .  $W(B)$  exceeds all  $W(B')$  of this kind and, therefore,  $W(A) + W(B) > 1 - \varepsilon$  for all  $\varepsilon > 0$ , that is,  $W(A) + W(B) \geq 1$  if  $W(A) < 1$ . This inequality obviously also holds if  $W(A) = 1$ .

Assume, for contradiction, that  $W(A) + W(B) > 1$ . We may assume that  $B = A^c$ . By Lemma C.3, there exists  $B' \subset A^c$  with  $1 - W(A) < W(B') < W(A^c)$  and  $\{A, B', A^c \setminus B'\}$  is an  $\mathcal{A}$ -collection. With  $x_1 \succ x_2$ , we have  $(A^c : x_1, A : x_2) \succ (B' : x_1, (B')^c : x_2)$  implying that  $A^c \setminus B'$  is nonnull. However,  $W(A) + W(B') > 1$  implies  $W(A \cup B') = 1$  (by Case 1.1,  $W(A \cup B') < 1$  would imply  $W(A) + W(B') < 1$ ) so that  $(A \cup B' : x_1, A^c \setminus B' : x_2) \sim x_1$ , which violates strong monotonicity (Lemma A.5). Case 1.2 and, thus, Case 1 have been completed.

*Case 2* [The equally-spaced case]. The capacity  $W$  takes only the values  $j/m$  for some  $m$ . Consider disjoint events  $A$  and  $B$  with  $A \cup B$  an event too, and  $W(A) = i/m$  and  $W(A \cup B) = j/m$ . Applying Lemma C.2 to  $A, B$ , and  $(A \cup B)^c$ , we obtain a partition  $E_1, \dots, E_m$  of  $S$  such that  $\pi(E_\ell^{\cup_{k=1}^{\ell-1} E_k}) = 1/m$  for all  $\ell$ ,  $\cup_{\ell=1}^i E_\ell = A$ , and  $\cup_{\ell=i+1}^j E_\ell = B$ . By strong monotonicity (Lemma A.5),  $\pi(E_j^F)$  must be nonzero for each  $j$  and each rank  $F$ ; it must always be at least  $1/m$ . Thus, each rank-ordering of  $E_1, \dots, E_m$  must yield decision weights for each  $E_j$  of at least  $1/m$ . Because the decision weights sum to one, each  $E_j$  must under each such rank-ordering have decision

weight exactly equal to  $1/m$ . This also holds for a rank-ordering where  $E_{i+1}, \dots, E_j$  are ranked first. Then their sum of decision weights is  $W(B)$ , which consequently is  $(j-i)/m$ . It follows that  $W(B) = W(A \cup B) - W(A)$ . SEU has been established for Case 2.

The proof of Theorem 5.3 is now complete.  $\square$

*Proof of Example 5.4* We establish solvability in each case. Assume that  $\beta_A f \prec g \prec \gamma_A f$ , with no outcomes of  $f$  and  $g$  between  $\gamma$  and  $\beta$ .

(i), (ii), and (v). We can take  $G$  such that  $P(G)(\gamma - \beta) = \text{SEU}(g) - \text{SEU}(\beta_A f) < P(A)(\gamma - \beta)$ , in (i) and (v) because of convex-rangeness and in (ii) because this  $P(G)$  is rational. In (i) and (v), for  $\mathcal{C}$  any subset of the reals could have been taken, and in (ii) any subset of the rational numbers.

(iii). Every SEU value is of the form  $j/n$  with  $0 \leq j \leq 2n$ . If  $\gamma = 2$  and  $\beta = 0$ , then all outcomes of the acts are 2 or 0, and in  $\beta_A f$  we change  $\beta$  into  $\gamma$  at so many elements of  $A$  that as many outcomes 2 result as under act  $g$ . If  $\gamma$  and  $\beta$  differ by only 1, then in  $\beta_A f$  we change  $\beta$  into  $\gamma$  at  $j$  states within  $A$  where  $j/n$  is the SEU difference between  $g$  and  $\beta_A f$ .  $A$  contains enough states to make this possible because, with  $\gamma_A f \succ g \succ \beta_A f$ , we have  $\|A\| > j$ .

(iv) Every SEU value and SEU difference are of the form  $a + b\sqrt{2}$  for integers  $a$  and  $b$ . If  $\gamma = 2$  and  $\beta = 0$ , then all outcomes of the acts are 2 or 0. In  $\beta_A f$  we change  $\beta$  to  $\gamma$  on a subset  $A_g$  of  $A$  such that after the change outcome 2 is as probable as for act  $g$ . If  $\gamma$  and  $\beta$  differ by only 1, then we change  $\beta$  into  $\gamma$  on a subset  $A_g$  of  $A$  where  $P(A_g)(\gamma - \beta) = \text{SEU}(g) - \text{SEU}(\beta_A f)$ . Such a set  $A_g$  can be seen to exist because  $P(A)(\gamma - \beta) > \text{SEU}(g) - \text{SEU}(\beta_A f)$  and  $(\text{SEU}(g) - \text{SEU}(\beta_A f))/(\gamma - \beta)$  is of the form  $a + b\sqrt{2}$ ,  $\gamma - \beta$  being 1. Finally, every probability of an event is of the form  $a + b\sqrt{2}$  with  $a$  and  $b$  integers, and can be rational only if it is 0 or 1.  $\square$

*Proof of Theorem 5.5* Monotonicity is implied by stochastic dominance. Necessity of the ps-Archimedean axiom, as

well as the other axioms, is obvious. Hence, (i) implies (ii). We next assume (ii), and derive (i) and the uniqueness result.

*Part 1* [two outcomes; qualitative probability]. Assume that there are only two outcomes  $\gamma \succ \beta$ . We will not use equivalence-dominance (e in Statement (ii)) in this part. This part will follow from Theorem 5.6 of Krantz et al. (1971) applied to our relation  $\succsim$  on events. We first show that all axioms of their theorem are satisfied. The results of Section B.1, in particular Lemma B.5 (nested matching) all remain valid, also for the present case of  $n = 1$ . The proof of the following lemma, deriving Axiom 5.3' of Krantz et al. (p. 215), provides an illustration of the way in which the likelihood method interacts with set union.

**LEMMA D.2** *If  $A$  is disjoint from  $B$  and  $E$  from  $F$ , then  $A \succsim E$  and  $B \succsim F$  imply  $A \cup B \succsim E \cup F$  whenever these unions are events, and the latter preference is strict whenever one of the antecedent preferences is. Hence,  $A \succ B$  if and only if  $A \succ_b B$ .*

*Proof.* First consider the case  $A \sim E$  and  $B \sim F$ . Let  $b(s) = \beta$  for all  $s$ . Consider the indifferences

$$\begin{aligned} \beta_A b &\sim \beta_E b & \beta_A \gamma_B b &\sim \beta_E \gamma_F b, \\ \gamma_A b &\sim \gamma_E b, & \gamma_A \gamma_B b &\sim \gamma_E \gamma_F b. \end{aligned}$$

The first indifference follows from reflexivity, and the second and third from  $B \sim F$  and  $A \sim E$ . Ps-likelihood consistency implies the fourth indifference. It implies  $A \cup B \sim_b E \cup F$  and, hence,  $A \cup B \sim E \cup F$ .

Next assume that exactly one of the antecedent preferences in the lemma is strict, say  $A \succ E$ . By solvability, there exists  $A' \subset A$  with  $A' \sim E$  and  $\{A', A \setminus A', B\}$  an  $\mathcal{A}$ -collection.<sup>4</sup> We get  $\gamma_E \gamma_F b \sim \gamma_{A'} \gamma_B b < \gamma_A \gamma_B b$ , the indifference as above, and the strict preference by nonnullness of  $A \setminus A'$  and strong monotonicity (Lemma A.5).  $A \cup B \succ_b E \cup F$  follows and, hence,  $A \cup B \succ E \cup F$ .

The case  $A \sim E$  and  $B \succ F$  is similar. If  $A \succ E$  and  $B \succ F$ , then we define  $A'$  as above and get  $A \cup B \succ A' \cup B \succ E \cup F$ .

$A \succsim_b B$  obviously implies  $A \succ B$ . Next suppose  $A \succ B$ , so that  $\beta_A f \sim \beta_B g$  and  $\gamma_A f \succ \gamma_B g$ . Let  $E = f^{-1}(\gamma) \setminus A$  and  $F = g^{-1}(\gamma) \setminus B$ . The indifference shows that  $E \sim_b F$  so that  $E \sim F$ . The preference shows that  $A \cup E \succsim_b B \cup F$  so that  $A \cup E \succ B \cup F$ .  $A \prec_b B$  would imply, by the first part of the lemma,  $A \cup E \prec B \cup F$ , violating ps-likelihood consistency.  $A \succsim_b B$  follows. *QED*

Because of the above lemma, the Archimedean axiom of expected utility is equivalent to the ps-Archimedean axiom. Weak ordering of  $\succ$  on the events follows,  $A \succ \emptyset$  for all  $A$  follows from monotonicity, and  $S \succ \emptyset$  follows from nondegeneracy and monotonicity.

The Archimedean axiom (Definition 5.4.4) of Krantz et al. (1971) follows from our ps-Archimedean axiom because an infinite standard sequence  $A_1, \dots$  in their sense can be taken nested by Lemma B.5, that is to say, increasing. It then is, by definition, of the form  $A_j = \cup_{i=1}^j E_i$  for all  $j$ , for a disjoint sequence  $E_j$  with  $E_j \sim E_1$  for all  $j$  and  $E_1$  nonnull. Then  $E_j \sim_b E_1$  for all  $j$  and, by transitivity, all  $E_j$  are equally likely in terms of  $\sim_b$ . This cannot be because of the ps-Archimedean axiom.

Axiom 5.5 of Krantz et al. (1971, p. 207) follows from our Lemma B.5. All conditions of their Theorem 5.3, except  $\mathcal{A}$  being a Dynkin system, have been verified. If  $\mathcal{A}$  is a Dynkin system (QM-algebra in the terminology of Krantz et al.) after all, then our theorem follows from their Theorem 5.3.

Assume next that  $\mathcal{A}$  is not a Dynkin system. The following proof closely follows the proof of Theorem 5.2 in Section 5.3.2 of Krantz et al. (1971). Denote by  $[\mathcal{A}]$  the set  $[A]$  of  $\sim$  equivalence classes of events  $A$ , i.e. sets of events  $B$  with  $B \sim A$ , and extend  $\succ$  to  $[\mathcal{A}]$ . Because of the axioms of qualitative probability and Lemma D.2, we can define  $[A]^c = [A^c]$  independently of the representative  $A$  chosen.  $[A]$  is *nonnull* if  $[A] \succ [\emptyset]$ . We call  $[A]$  and  $[B]$  *disjoint* if  $[A]^c \succ [B]$ , which is equivalent to  $[B]^c \succ [A]$ . By solvability,  $[A]$  and  $[B]$  are disjoint if and only if we can find disjoint representatives  $A, B$ .  $[A]$  can be disjoint from itself. For all disjoint  $[A]$  and  $[B]$  we define a gener-

alized disjoint union operation  $[A] \vee [B]$  (denoted  $\circ$  in Krantz et al.) as  $[A \cup B]$  where  $A$  and  $B$  are disjoint representatives. By Lemma D.2, the particular choice of  $A$  and  $B$  does not matter.  $\vee$  is associative and commutative, and we can define  $n[A]$  in the usual way.

For contradiction, assume that  $[A]$  is nonnull, but  $n[A]$  can be defined for arbitrarily large  $n$ . We define  $E_1$  as any representative from  $[A]$ . By solvability, we can define inductively, for each  $n, E_n$  such that the  $E_1, \dots, E_n$  are disjoint,  $E_1 \cup \dots \cup E_n$  is a representative of  $n[A]$ , and  $E_1, \dots, E_n$  is an  $A$ -sequence. A violation of the ps-Archimedean axiom in our original structure has resulted. It can now easily be established that our structure satisfies all conditions of Theorem 3.3 of Krantz et al. (1971), that a positive-valued ratio scale exists that represents  $\succsim$ , and that this yields a probability measure  $P$  representing  $\succsim$ , exactly as in Section 5.3.2 of Krantz et al. (1971).

Note that in this Part 1, strict stochastic dominance is also satisfied.

*Part 2 [more than two outcomes].* Now there can be more than two outcomes. Fix some  $\gamma \succ \beta$ . Let  $P$  be the probability measure obtained from Part 1 for the acts yielding only  $\gamma$  and  $\beta$ .

To establish stochastic dominance, assume that  $P\{s : f(s) \succ \alpha\} \geq P\{s : g(s) \succ \alpha\}$  for each outcome  $\alpha$ . By finite ranges of  $f$  and  $g$  and Corollary B.8, which also holds if there are not more than two nonequivalent outcomes, there exist  $f'$  and  $g'$  with  $\{s : f(s) \succ \alpha\} \sim_b \{s : f'(s) \succ \alpha\}$  and  $\{s : g(s) \succ \alpha\} \sim_b \{s : g'(s) \succ \alpha\}$ , and  $\{s : g'(s) \succ \alpha\} \subset \{s : f'(s) \succ \alpha\}$  for all  $\alpha$ . By equivalence-dominance,  $f' \sim f$  and  $g' \sim g$ . By monotonicity,  $f' \succ g'$  and, hence,  $f \succ g$ .  $\square$

*Proof of Example 5.8* We establish solvability in each case. Assume that  $\beta_A f \prec g \prec \gamma_A f$ , with no outcomes of  $f$  and  $g$  between  $\gamma$  and  $\beta$ .

- (i) Because  $f$  and  $g$  yield no outcomes between  $\gamma$  and  $\beta$ ,  $\gamma_A f$  and  $\beta_A f$  are comonotonic. By permuting outcomes of  $g$ ,

R	Z	A	L		R	A	Z	L			
f	f	$\beta$	f	~	g	$\beta$	g	g		(D.1)	
f	f	$\gamma$	f	?	g	$\gamma$	g	g		(D.2)	
←				←							
increasingly preferred outcomes				increasingly preferred outcomes							

Figure D.1.  $\gamma > \beta$ ,  $f \succ \gamma$  on  $R \cup Z$ ,  $f \preceq \beta$  on  $L$ ,  $g \succ \gamma$  on  $R$ ,  $g \preceq \beta$  on  $Z \cup L$

we can always replace  $g$  by an equivalent  $g'$  that is comonotonic with the other two acts. Dropping the  $n - m$  best-ranked states of nature (that are “null”), we end up in an SEU substructure and can copy the derivation of Example 5.4. (iii).

- (ii) Solvability follows from convex-rangedness of  $\lambda$  and continuity of  $w$ .  $\square$

*Proof of Theorem 6.1* Necessity of the preference condition is immediate. In preparation for sufficiency, consider some  $A, R, R'$  with  $R' \supset R$  disjoint from  $A$ . Write  $Z = R' \setminus R$  and  $L = (R \cup Z \cup A)^c$ . If we can find  $f, g$ , and  $\gamma > \beta$  such that the configurations of Equations (D.1) and (D.2) in Figure D.1 hold, then  $\pi(A^{R'}) < \pi(A^R)$  if and only if the question mark is a strict preference  $<$ , which then holds if and only if  $A^{R'} < A^R$ . Therefore, whenever we can construct an indifference as in Figure D.1, the assumed preference condition in Theorem 6.1 implies the required  $\pi(A^{R'}) \geq \pi(A^R)$ . From now on, for sufficiency, assume the preference condition.

*Case 1.* [The equally-spaced case]. There exist three outcomes that are equally-spaced in utility units and, by renormalizing, we may assume that these utilities are 0, 1, and 2, and that the outcomes are identical to these utilities. There exists an  $m \geq 2$  such that capacities and decision weights only take values  $j/m$ . The following subcases of Case 1 are exhaustive (but not exclusive).

*Case 1.1.* [ $\pi(Z^R) = 1/m$ ,  $\pi(A^{R \cup Z}) = 0$ ].  $L^w$  or  $R^b$  is non-null (possibly both); say  $L^w$  is ( $R^b$  is similar). By Lemma C.2, there is  $L' \subset L$  with  $\pi((L')^w) = 1/m$ . Then  $(R : 2, Z : 2, A :$

$1, L \setminus L' : 1, L' : 0) \sim (R : 2, Z : 1, A : 1, L \setminus L' : 1, L' : 1) = (R : 2, A : 1, Z : 1, L : 1)$ . This is an indifference as in Figure D.1.  $\pi(A^R) \leq \pi(A^{R \cup Z})$  follows, implying  $\pi(A^R) = \pi(A^{R \cup Z}) = 0$ .

*Case 1.2.* [ $\pi(Z^R) > 1/m, \pi(A^{R \cup Z}) = 0$ ]. By Lemma C.2, we can partition  $Z$  into  $Z_1, \dots, Z_k$  such that  $\pi(Z_j^{R \cup_{i=1}^{j-1} Z_i}) = 1/m$  for all  $j$ . By repeated application of Case 1.1 (“reducing the rank by decision weight  $1/m$  leaves nullness unaffected”),  $\pi(A^{R \cup Z \setminus Z_k}) = 0 = \pi(A^{R \cup Z \setminus (Z_{k-1} \cup Z_k)}) = \dots = \pi(A^{R \cup Z \setminus (Z_1 \cup \dots \cup Z_k)}) = \pi(A^R)$ .

*Case 1.3.* [ $\pi(A^R) = k/m > 0$ ]. By Lemma C.2, we can partition  $A$  into  $A_1, \dots, A_k$  such that  $\pi(A_j^{R \cup_{i=1}^{j-1} A_i}) = 1/m$  for all  $j$ . By repeated application of Case 1.2 (reducing the rank by any decision weight leaves nullness unaffected, so that enlarging the rank by any decision weight leaves nonnullness unaffected),  $\pi(A_j^{R \cup Z \cup_{i=1}^{j-1} A_i}) = 0$  for all  $j$ . Each of the nonzero decision weights must be at least  $1/m$ .  $\pi(A^{R \cup Z})$ , the sum of these decision weights, must be at least  $k/m$ , and exceeds  $\pi(A^R)$ .

*Case 1.4.* [ $\pi(A^R) = 0$ ]. Then  $\pi(A^{R \cup Z}) \geq \pi(A^R)$  is immediate. Case 1, the equally-spaced case, is now complete.  $\pi(A^{R \cup Z}) \geq \pi(A^R)$  then always holds.

*Case 2* [the densely-spaced case]. There exist three non-equivalent outcomes. We may assume that they are identical to their utilities, and that they are 0, 1, and  $1 + \lambda$  for  $\lambda > 0$ .

*Case 2.1.* [ $\pi(A \cup Z)^R < 1$ ]. Then at least one of  $R^b$  and  $L^w$  is nonnull, say the latter (the other case is similar).

*Case 2.1.1.* [ $\pi(Z^R) \leq \pi(L^w)/\lambda$ ]. (If  $L^w$  is null and  $R^b$  is nonnull, then the corresponding case is  $\pi(Z^R) \leq \pi(R^b)\lambda$ .) Then  $(R : 1 + \lambda, Z : 1, A : 1, L : 0) \preceq (R : 1 + \lambda, Z : 1 + \lambda, A : 1, L : 0) \preceq (R : 1 + \lambda, Z : 1, A : 1, L : 1) = (R : 1 + \lambda, A : 1, Z : 1, L : 1)$ . By solvability,  $(R : 1 + \lambda, Z : 1 + \lambda, A : 1, L : 0) \sim (R : 1 + \lambda, A : 1, Z : 1, L_1 : 1, L_0 : 0)$  for a partition  $L_1, L_0$  of  $L$ . We have obtained a configuration as in Figure D.1, with  $\beta = 1$  and  $\gamma = 1 + \lambda$ , and  $\pi(A^R) \geq \pi(A^R)$  follows.

*Case 2.1.2* [ $\pi(Z^R) > \pi(L^w)/\lambda$ ]. By Lemma C.3, we can partition  $Z$  into  $\{Z_1, \dots, Z_k\}$ , ranked from best to worst, such that the decision weight of each event is less than  $\pi(L^w)/\lambda$ .

By repeated application of Case 2.1.1 (“reducing the rank of event  $A$  by a sufficiently small decision weight does not increase  $A$ ’s decision weight”),  $\pi(A^{R \cup_{i=1}^j Z_i}) \geq \pi(A^{R \cup_{i=1}^{j-1} Z_i})$  for all  $k$ .  $\pi(A^{R'}) \geq \pi(A^R)$  follows (reducing the rank of event  $A$  by any decision weight does not increase  $A$ ’s decision weight).

*Case 2.2.* [ $\pi(Z^R) = 0$ ]. Then  $\pi(A^{R \cup Z}) = \pi(A^{R \cup Z}) + \pi(Z^R) = \pi((A \cup Z)^R) \geq \pi(A^R)$ .

*Case 2.3.* [ $\pi(A \cup Z)^R = 1$  and  $\pi(Z^R) > 0$ ]. Consider a partition  $\{Z_1, Z_2\}$  of  $Z$  with  $\pi(Z_1^R) > 0$ . Then  $\pi((A \cup Z_2)^{R \cup Z_1}) < 1$  so that, by Case 2.1,  $\pi(A^{R \cup Z}) \geq \pi(A^{R \cup Z_1})$ .  $\pi(A^{R \cup Z_1}) = \pi(A \cup Z_1)^R - \pi(Z_1^R) \geq \pi(A^R) - \pi(Z_1^R)$ . By Lemma C.3, for each  $\varepsilon > 0$  there exist  $Z_1, Z_2$  as above with  $0 < \pi(Z_1^R) < \varepsilon$ . By letting  $\varepsilon$  tend to zero,  $\pi(A^{R \cup Z}) \geq \pi(A^R)$  follows. This derivation did not use any continuity assumption concerning  $W$ .  $\square$

#### REFERENCES

- Abdellaoui, Mohammed (2000), Parameter-free elicitation of utilities and probability weighting functions, *Management Science* 46, 1497–1512.
- Abdellaoui, Mohammed (2002), A genuine rank-dependent generalization of the von Neumann–Morgenstern expected utility theorem, *Econometrica* 70, 717–736.
- Anscombe, Francis J. and Aumann, Robert J. (1963), A definition of subjective probability, *Annals of Mathematical Statistics* 34, 199–205.
- Blackorby, Charles, Davidson, Russell, and Donaldson, David (1977), A homiletic exposition of the expected utility hypothesis, *Economica* 44, 351–358.
- Casadesus-Masanell, Ramon, Klibanoff, Peter, and Ozdenoren, Emre (2000), Maxmin expected utility over Savage acts with a set of priors, *Journal of Economic Theory* 92, 35–65.
- Chateauneuf, Alain (1991), On the use of capacities in modeling uncertainty aversion and risk aversion, *Journal of Mathematical Economics* 20, 343–369.
- Chateauneuf, Alain (1999), Comonotonicity axioms and rank-dependent expected utility theory for arbitrary consequences, *Journal of Mathematical Economics* 32, 21–45.
- Chew, Soo Hong and Karni, Edi (1994), Choquet expected utility with a finite state space: commutativity and act-independence, *Journal of Economic Theory* 62, 469–479.
- Chew, Soo Hong and Sagi, Jacob (2004), “Event Exchangeability: Small Worlds Probabilistic Sophistication without Continuity or Monotonicity,” Haas School of Business, University of California, Berkeley, CA.

- Cohen, Michèle, Gilboa, Itzhak, Jaffray, Jean-Yves and Schmeidler, David (2000), An experimental study of updating ambiguous beliefs, *Risk, Decision, and Policy* 5, 123–133.
- Davidson, Donald and Suppes, Patrick (1956), A finitistic axiomatization of utility and subjective probability, *Econometrica* 24, 264–275.
- de Finetti, Bruno (1937), La prévision: ses lois logiques, ses sources subjectives, *Annales de l'Institut Henri Poincaré* 7, 1–68. Translated into English by Henry E. Kyburg Jr., Foresight: Its Logical Laws, its Subjective Sources, in Henry E. Kyburg Jr. and Howard E. Smokler (1964, eds.), *Studies in Subjective Probability*, Wiley, New York; pp. 93–158, 2nd edition 1980, pp. 53–118, Krieger, New York.
- de Finetti, Bruno (1974), *Theory of Probability*. Vol. I, Wiley, New York.
- Ellsberg, Daniel (1961), Risk, ambiguity and the Savage axioms, *Quarterly Journal of Economics* 75, 643–669.
- Epstein, Larry G. and Le Breton, Michel (1993), Dynamically consistent beliefs must be Bayesian, *Journal of Economic Theory* 61, 1–22.
- Epstein, Larry G. and Zhang, Jiangkang (2001), Subjective probabilities on subjectively unambiguous events, *Econometrica* 69, 265–306.
- Ergin, Haluk and Gul, Faruk (2004), *A Subjective Theory of Compound Lotteries*, MIT, Cambridge, MA.
- Fishburn, Peter C. (1986), The axioms of subjective probability, *Statistical Science* 1, 335–358.
- Fishburn, Peter C. (1988), *Nonlinear Preference and Utility Theory*. Johns Hopkins University Press, Baltimore, MD.
- Fox, Craig R. and Tversky, Amos (1995), Ambiguity aversion and comparative ignorance, *Quarterly Journal of Economics* 110, 585–603.
- Ghirardato, Paolo and Marinacci, Massimo (2001), Risk, ambiguity, and the separation of utility and beliefs, *Mathematics of Operations Research* 26, 864–890.
- Ghirardato, Paolo, Maccheroni, Fabio, Marinacci, Massimo and Siniscalchi, Marciano (2003), A subjective spin on roulette wheels, *Econometrica* 71, 1897–1908.
- Gilboa, Itzhak (1987), Expected utility with purely subjective non-additive probabilities, *Journal of Mathematical Economics* 16, 65–88.
- Gilboa, Itzhak (ed.) (2004), *Uncertainty in Economic Theory: Essays in Honor of David Schmeidler's 65th Birthday*. Routledge, London.
- Gilboa, Itzhak and Schmeidler, David (1989), Maxmin expected utility with a non-unique prior, *Journal of Mathematical Economics* 18, 141–153.
- Gilboa, Itzhak and Schmeidler, David (2004), Subjective distributions, *Theory and Decision* 56, 345–357.
- Gonzalez, Richard and Wu, George (1999), On the shape of the probability weighting function, *Cognitive Psychology* 38, 129–166.

- Grant, Simon (1995), Subjective probability without eventwise monotonicity: or: how Machina's mom may also be probabilistically sophisticated, *Econometrica* 63, 159–189.
- Grodal, Birgit (1978), Some further results on integral representation of utility functions, Institute of Economics, University of Copenhagen, Copenhagen. Appeared in rewritten form in Chapter 12 of Karl Vind (2003), *Independence, Additivity, Uncertainty. With contributions by B. Grodal*. Springer, Berlin.
- Gul, Faruk (1992), Savage's theorem with a finite number of states, *Journal of Economic Theory* 57, 99–110. (Erratum, 1993, *Journal of Economic Theory* 61, 184).
- Karni, Edi (1993), Subjective expected utility with state-dependent preferences, *Journal of Economic Theory* 60, 428–438.
- Keynes, John Maynard (1921), *A Treatise on Probability* second ed. (1948), McMillan, London.
- Knight, Frank H. (1921), *Risk, Uncertainty, and Profit*. Houghton Mifflin, New York.
- Köbberling, Veronika and Wakker, Peter P. (2003), Preference foundations for nonexpected utility: a generalized and simplified technique, *Mathematics of Operations Research* 28, 395–423.
- Kopylov, Igor (2004), *Subjective Probabilities on "Small" Domains*, Institute for Mathematical Behavioral Sciences, University of California, Irvine CA.
- Krantz, David H., Luce, R. Duncan, Suppes, Patrick and Tversky, Amos (1971), *Foundations of Measurement, Vol. I (Additive and Polynomial Representations)*, Academic Press, New York.
- Luce, R. Duncan (1967), Sufficient conditions for the existence of a finitely additive probability measure, *The Annals of Mathematical Statistics* 38, 780–786.
- Luce, R. Duncan (2000), *Utility of Gains and Losses: Measurement-Theoretical and Experimental Approaches*, Lawrence Erlbaum Publishers, London.
- MacCrimmon, Kenneth R. and Larsson, Stig (1979), Utility Theory: Axioms versus "Paradoxes", In Maurice Allais and Ole Hagen (eds.), *Expected Utility Hypotheses and the Allais Paradox*, Reidel, Dordrecht, The Netherlands, pp. 333-409.
- Machina, Mark J. and Schmeidler, David (1992), A more robust definition of subjective probability, *Econometrica* 60, 745–780.
- Nakamura, Yutaka (1990), Subjective expected utility with non-additive probabilities on finite state spaces, *Journal of Economic Theory* 51, 346–366.
- Nakamura, Yutaka (1995), Rank dependent utility for arbitrary consequence spaces, *Mathematical Social Sciences* 29, 103–129.

- Nau, Robert F. (1995), Coherent decision analysis with inseparable probabilities and utilities, *Journal of Risk and Uncertainty* 10, 71–91.
- Nehring, Klaus D.O. (2001), Ambiguity in the context of probabilistic beliefs, mimeo.
- Pfanzagl, Johann (1968), *Theory of Measurement*. Physica-Verlag, Vienna.
- Quiggin, John (1981), Risk perception and risk aversion among Australian farmers, *Australian Journal of Agricultural Economics* 25, 160–169.
- Sarin, Rakesh K. and Wakker, Peter P. (1992), A simple axiomatization of nonadditive expected utility, *Econometrica* 60, 1255–1272.
- Sarin, Rakesh K. and Wakker, Peter P. (1998), Revealed likelihood and Knightian uncertainty, *Journal of Risk and Uncertainty* 16, 223–250.
- Sarin, Rakesh K. and Wakker, Peter P. (2000), Cumulative dominance and probabilistic sophistication, *Mathematical Social Sciences* 40, 191–196.
- Savage, Leonard J. (1954), *The Foundations of Statistics*. Wiley, New York. (Second ed. 1972, Dover Publications, New York.)
- Schmeidler, David (1989), Subjective Probability and expected utility without additivity, *Econometrica* 57, 571–587.
- Tversky, Amos and Kahneman, Daniel (1992), Advances in prospect theory: cumulative representation of uncertainty, *Journal of Risk and Uncertainty* 5, 297–323.
- Tversky, Amos and Wakker, Peter P. (1995), Risk attitudes and decision weights, *Econometrica* 63, 1255–1280.
- Wakker, Peter P. (1989), *Additive Representations of Preferences, A New Foundation of Decision Analysis*. Kluwer Academic Publishers, Dordrecht, The Netherlands.
- Wakker, Peter P. (1990), Under stochastic dominance Choquet-expected utility and anticipated utility are identical, *Theory and Decision* 29, 119–132.
- Wakker, Peter P. (1991), Additive representation for equally spaced structures, *Journal of Mathematical Psychology* 35, 260–266.
- Wakker, Peter P. (1993), Unbounded utility for Savage’s “foundations of statistics,” and other models, *Mathematics of Operations Research* 18, 446–485.
- Wakker, Peter P. (2001), Testing and characterizing properties of nonadditive measures through violations of the sure-thing principle, *Econometrica* 69, 1039–1059.
- Wakker, Peter P. (2004), On the composition of risk preference and belief, *Psychological Review* 111, 236–241.
- Wakker, Peter P. (2005), Decision-foundations for properties of nonadditive measures for general state spaces or for general outcome spaces, *Games and Economic Behavior* 50, 107–125.
- Wald, Abraham (1950), *Statistical Decision Functions*, Wiley, New York.
- Wu, George and Gonzalez, Richard (1999), Nonlinear decision weights in choice under uncertainty, *Management Science* 45, 74–85.

- Yaari, Menahem E. (1987), The dual theory of choice under risk, *Econometrica* 55, 95–115.
- Zank, Horst (2004), Deriving rank-dependent expected utility through probabilistic consistency, School of Economic Studies, The University of Manchester, Manchester, UK.
- Zhang, Jiangkang (1999), Qualitative probabilities on lambda-systems, *Mathematical Social Sciences* 38, 11–20.
- Zhang, Jiangkang (2002), Subjective ambiguity, expected utility and Choquet expected utility, *Economic Theory*, 20, 159–181.

*Addresses for correspondence:* Mohammed Abdellaoui, GRID, Maison de la Recherche de l'ESTP, 30, Avenue du Président Wilson, 94230 Cachan, France. E-mail: [abdellaoui@grid.ensam.estp.fr](mailto:abdellaoui@grid.ensam.estp.fr)  
Peter P. Wakker, CREED, Dept. of Economics, University of Amsterdam, Amsterdam, The Netherlands. E-mail: [Wakker@few.eur.nl](mailto:Wakker@few.eur.nl)