Derivation of a cardinal utility through a weak tradeoff consistency requirement

Shiri Alon*

Abstract

An axiomatic model is presented, in which a utility function over consequences, unique up to location and unit, is derived. The axioms apply to a binary relation over purely subjective acts, namely no exogenous probabilities are assumed. The main axiom used is a weak tradeoff consistency condition. The model generalizes the biseparable model of Ghirardato and Marinacci (2001).

Keywords: Decision theory, choice under uncertainty, cardinal utility, tradeoff consistency. MSC2000 subject classification :Primary: 91B06 OR/MS subject classification :Primary: Decision analysis

1 Introduction.

Decision models under uncertainty are commonly composed of an abstract set of *states of nature*, an abstract set of *consequences*, and mappings from states to consequences, called *acts*. The decision maker's preferences are often modeled through a binary relation over the set of acts. Preferences over consequences are induced from preferences over constant acts (acts yielding the same consequence in every state of nature).

Given a binary relation over acts, one of the most basic questions is whether there is a utility function that represents the preference induced over consequences. If there is a utility function, then it is further interesting to inquire in what sense, if any, the utility function is unique. Many Decision Theory models derive a utility function that is *cardinal*, namely unique up to location and unit, which has the advantage that utility differences between consequences become comparable. This task is typically easier when

*The department of Economics, Bar Ilan University. email: shiri.aloneron@gmail.com

exogenous probabilities are involved, so that acts map states of nature to *lotteries* over a set of consequences, and more demanding when exogenous probabilities are not assumed.

In the literature, one of the techniques applied to derive cardinal utility functions when no exogenous probabilities are available is the *tradeoffs*, or *standard sequences*, technique. This technique originated in Thomsen [11], further developed in Blaschke and Bol [2] and Debreu [3], and was thoroughly investigated in Krantz, Luce, Suppes and Tversky [8]. The tradeoffs technique yields statements such as 'the tradeoff between consequence x and consequence y is the same as the tradeoff between consequence z and consequence w'. The way to produce such statements is to use particular events and acts as measuring devices. For instance, consider consequences α and β and an event E as your measuring device. If an act yielding a consequence x on E and α otherwise is indifferent to an act yielding a consequence y on the same event E and β otherwise, then α, β and E measure the tradeoff, or the 'preference distance', between x and y. Hence, if an act yielding a consequence z on E and α otherwise is indifferent to an act yielding a consequence w on E and β otherwise, then the tradeoff between z and w equals the tradeoff between x and y. The idea is that these indifferences exhibit that under the same circumstances (i.e., event E and consequences α and β) the decision maker is willing to trade x for y as much as z for w. In a similar manner, standard sequences are sequences (not only pairs), in which consequences are equally spaced in terms of 'preference distance'.

It could be the case that all events and all acts are employed as measuring devices, or that only some events, or some acts, are allowed to play that part. In any case, in order for these tradeoff measurements to be consistent, it should be that any one of these devices yields the same results. Wakker, [12] and [13], and Kobberling and Wakker [6] (abbreviated KW from now on), use tradeoff consistency requirements over different sets of events and acts to obtain different representations: all events and all acts for Subjective Expected Utility, comonotonic acts (see Definition 1 below) for nonadditive expected utility $\acute{a} - la$ Schmeidler [10], and so on. Their work assumes an environment that does not contain exogenous probabilities, and delivers, among other parameters of the representation, a cardinal utility.

This paper examines a tradeoff consistency requirement in which the acts allowed to be used to measure tradeoffs are either bets on, or bets against a state, namely comonotonic acts which obtain one consequence on one state and another consequence on all other states. It is shown that even if tradeoff consistency is restricted to comonotonic bets on or against states, and moreover tradeoffs are measured only over single states (and not over their complements), then a cardinal utility function over consequences may still be

elicited.[1] This utility function respects tradeoff measurements in the sense that equal tradeoffs imply equal utility differences. The model does not assume existence of exogenous probabilities, but instead requires that the set of consequences be a connected topological space. A restriction of the model is that the state space is assumed finite.

In addition to the elicitation of a cardinal utility, the axioms presented imply local additive representations over sets of bets on, and sets of bets against states. More specifically, denote by $\alpha s \beta$ a bet on state s, with consequences α and β such that α is preferred to β. The relation over the set of all bets on state s satisfies that $\alpha s \beta$ is preferred to $\gamma s \delta$, if and only if, $u(\alpha) + V_{-s}(\beta) \geq u(\gamma) + V_{-s}(\delta)$, where u is a cardinal utility over consequences and V_{-s} is a continuous function that (like u) represents the relation over constant acts. The same is satisfied for bets against states.[2] These additive representations are local in that each holds within one set of bets on a state, or one set of bets against a state, and they cannot be employed to compare bets contingent on different states, nor to compare bets on and against the same state.

The model presented in this paper generalizes the biseparable model of Ghirardato and Marinacci [4]. The biseparable model, which encompasses many known choice models as special cases (for instance, non-additive expected utility of Schmeidler [10], maxmin expected utility of Gilboa and Schmeidler [5]), imposes relatively weak assumptions yet extracts one additive representation over all bets on events, and in particular a cardinal utility over consequences (a similar representation is given by the binary rank-dependent utility model of Luce [9]). As discussed in KW (Section 5.2), a tradeoff consistency requirement over comonotonic bets on events, supplemented with a few basic conditions, yields a biseparable representation. This paper shows that a weaker consistency assumption, that restricts attention to bets on or against *states*, still suffices to derive a cardinal utility. It thus suggests a simpler condition that can be tested in order to establish the use of a cardinal utility over consequences. At the same time, it should be noted that the tradeoff consistency condition is not sufficient to obtain a global additive representation over bets, even if only bets on and against states are considered.

The result in this paper serves to strengthen the validity of the tradeoff method as a tool to empirically measure utility. The tradeoff method elicits a utility by finding consequences which are equally spaced in terms

[1] A few basic axioms are also assumed.

[2] That is, the relation over acts $\alpha s \beta$ with β preferred to α also admits an additive representation $u(\alpha) + V_{-s}(\beta)$, but this representation is different than the one over bets *on* s.

of tradeoffs, under a suitable assumption of tradeoff consistency (for a detailed description and a discussion of the method see Wakker and Deneffe [14], Kobberling and Wakker [7], and the references therein). Assuming that the axioms suggested in this paper are satisfied, the tradeoff method may be applied to derive a cardinal utility over consequences by measuring tradeoffs over single states, using bets on or against states. The representation result introduced in this paper thus extends the domain of preferences to which the tradeoff method applies. Not only is it not required to assume an expected utility representation in order to apply the method, but it is even possible to measure a utility when the relation over binary acts does not conform to the biseparable model (see Example 4).

The paper is organized as follows. The next section presents notation and basic axioms. Section 3 presents the tradeoff consistency axiom used and the result. Section 4 contains an example for a relation which obtains a representation as in the theorem, yet is not a biseparable relation. Section 5 discusses the problem of infinite values and a possible generalization. Finally, all proofs appear in Section 6.

2 Notation and basic axioms.

S	a set of *states of nature*, with typical elements $s, t, \ldots$.
$\Sigma = 2^S$	an algebra of all *events* over S, with typical elements $E, F, \ldots$.
X	a nonempty set of *consequences*, with typical elements $x, y, \ldots$.
$X^S = \{f : S \to X\}$	the set of *acts*: mappings from states to consequences, with typical elements $f, g, \ldots$.
$\succsim$	the decision maker's preference relation: a binary relation over the set of acts X^S. As usual, $\sim$ and $\succ$ denote its symmetric and asymmetric components.
xEy	a notation for the act which assigns the consequence x to the states in E and the consequence y otherwise. $x\{s\}y$ is sometimes abbreviated to xsy (more generally, s is sometimes used instead of $\{s\}$).
$\overline{x}$	a constant act assigning consequence x to every state of nature.

With the usual slight abuse of notation, the symbol $\succsim$ is also used to denote a binary relation on X, defined by: $x \succsim y$ if $\overline{x} \succsim \overline{y}$. A state s is said to be *null* if for all $x, y, z \in X$, $xsz \sim ysz$. For a set of acts $\mathcal{A}$, s is said to be null on $\mathcal{A}$ if $xsz \sim ysz$ whenever $xsz, ysz \in \mathcal{A}$. Otherwise the state is *non-null* (on $\mathcal{A}$). Similarly, an event E is null (on $\mathcal{A}$) if for all $x, y, z \in X$, $xEz \sim yEz$ (whenever $xEz, yEz \in \mathcal{A}$).

The first assumption, A0, restricts the domain of decision problems accommodated by the model. It poses structural requirements on the set of states, as well as on the set of consequences.

A0. Structural assumption:

(a) S is finite.

(b) X is a connected topological space, and X^S is endowed with the product topology.

Three basic assumptions are presented first. Discussion of the axioms is delayed until after their statement.

A1. Weak Order:

(a) For all f and g in X^S, $f \succsim g$ or $g \succsim f$ (completeness).

(b) For all f, g, and h in X^S, if $f \succsim g$ and $g \succsim h$ then $f \succsim h$ (transitivity).

A2. Continuity: The sets $\{f \in X^S \mid f \succ g\}$ and $\{f \in X^S \mid f \prec g\}$ are open for all g in X^S.

A3. Monotonicity: For any two acts f and g, $f \succsim g$ holds whenever $f(s) \succsim g(s)$ for all states s in S.

Assumptions A1-A3 are standard. The first requires that the preference relation be complete and transitive, the second states that it is continuous, and the third that monotonicity holds. Monotonicity implies that the preference over consequences is state-independent. If the consequence $f(s)$ is preferred to the consequence $g(s)$ in each state s, where these are compared through constant acts, then the composition of f out of these consequences cannot make it worse than g.

In order to state the fourth assumption the notion of a comonotonic set of acts is defined, generalizing the definition of pairwise comonotonicity.

Definition 1. *A set of acts is a* ***comonotonic set*** *if there are no two acts f and g in the set and states s and t, such that $f(s) \succ f(t)$ and $g(t) \succ g(s)$. Acts in a comonotonic set are said to be* ***comonotonic acts****.*

Comonotonic acts induce essentially the same ranking of states according to the desirability of their consequences (or, more accurately, the same ranking up to indifferences). Given any numeration of the states, say $\pi : S \to \{1, ..., |S|\}$, the set $\{ f \in X^S \mid f(\pi(1)) \succsim \ldots \succsim f(\pi(|S|) \}$ is comonotonic. It is a largest-by-inclusion comonotonic set of acts.

A4. Consistent Essentiality:

(a) For all $s \in S$,

$$\begin{array}{lll} asx \succ bsx & \Rightarrow & csy \succ dsy \text{ for all } c \succ d, \\ & & \text{whenever the set } \{asx, bsx, csy, dsy\} \text{ is comonotonic.} \\ xsa \succ xsb & \Rightarrow & ysc \succ ysd \text{ for all } c \succ d, \\ & & \text{whenever the set } \{xsa, xsb, ysc, ysd\} \text{ is comonotonic.} \end{array}$$

(b) There exist distinct states s', s'' and consequences x, y such that $xs'y \succ \bar{y}$ and $\bar{x} \succ ys''x$.

Part (a) of Consistent Essentiality implies that if a state s is non-null on a set of bets on, or a set of bets against s, then it should influence the decision for every pair of acts from this set. The same holds true for essentiality of the complement. Part (b) of the axiom guarantees that there are at least two non-null states. Moreover, applying Monotonicity, $ys''x \succsim xs'y$, therefore part (b) also yields that for some set of bets on, and some set of bets against states, both the state and its complement are non-null.

Consistent Essentiality and Monotonicity can be replaced by a simpler axiom of Dominance: $f \succ g$ whenever $f(s) \succsim g(s)$ for all states s, with strict preference for at least one state. Though simpler, Dominance is a stronger condition as it excludes the possibility of null states. Therefore the less restrictive axioms of Consistent Essentiality and Monotonicity were chosen here.

3 Tradeoff consistency and results.

In order to present the tradeoff measurement used in this paper and its corresponding consistency requirement, two definitions are introduced.

Definition 2. *An act of the form $\alpha s \beta$ is called a* ***simple binary act****. For a state s, a comonotonic set of acts of the form $\alpha s \beta$, either $\{\alpha s \beta \mid \alpha \succsim \beta\}$ or $\{\alpha s \beta \mid \beta \succsim \alpha\}$, is termed a* ***simple binary comonotonic set****.*

Whenever the notation $\mathcal{C}_s$ is used in the sequel to denote a simple binary comonotonic set of acts, it is to be understood that it denotes such a set of the form $\alpha s \beta$.

Definition 3. *Given a comonotonic set of acts $\mathcal{A}$, an event E is said to be* ***comonotonically non-null on*** *$\mathcal{A}$, if there are consequences x, y, z such that $xEz \succ yEz$, and $\mathcal{A} \cup \{xEz, yEz\}$ is comonotonic. Otherwise, E is* ***comonotonically null on*** *$\mathcal{A}$.*

For instance, $\{s\}$ is comonotonically non-null on $\{asx\}$ if it is non-null on the simple binary comonotonic set of acts containing asx. Similarly, if $\mathcal{C}_s$ is a simple binary comonotonic set of acts and $\mathcal{A} \subseteq \mathcal{C}_s$ is any subset of it, then $\{s\}$ is comonotonically non-null on $\mathcal{A}$ whenever it is non-null on $\mathcal{C}_s$.

The relation which determines how tradeoffs are measured under the model is defined next.

Definition 4. *Define a relation $\sim^*$ over pairs of consequences: for consequences a, b, c, d,*
$< a; b > \sim^* < c; d >$ if there exist consequences x, y and a state s such that,

$$asx \sim bsy \quad \text{and} \quad csx \sim dsy \tag{1}$$

with all four acts comonotonic and $\{s\}$ comonotonically non-null on the set of four acts.

To get intuition for the relation $\sim^*$, assume that $y \succ x$. In such a case, having $asx \sim bsy$ implies that a is preferred to b in the precise amount that makes the act asx indifferent to the act bsy. In other words, gaining a instead of b on s exactly compensates for the advantage of y over x outside state s. If the same is true for consequences c and d then it is concluded that the decision maker is willing to trade a for b as much as c for d, namely that the tradeoff between a and b is the same as the tradeoff between c and d.

An important point to note is that the relation $\sim^*$ measures the tradeoff between consequences only over single states, and only through bets on one particular state, or bets against one particular state. By contrast, for consequences a, b, c, d, state s and acts f, g, indifferences $asf \sim bsg$ and $csf \sim dsg$ do not in general imply similar equivalence of the tradeoffs $< a; b >$ and $< c; d >$. Moreover, tradeoffs are not even allowed to be measured over complements of single states, so that $xsa \sim ysb$ and $xsc \sim ysd$ still do not imply

that the tradeoffs $< a; b >$ and $< c; d >$ are equivalent. The measurement allowed is therefore very cautious. It suggests that indifference relationships which include any violation of comonotonicity, or employ acts which are not simple binary ones, may involve considerations other than the mere tradeoff between consequences.

The definition of $\sim^*$ would be meaningful only if the measurements employed in its definition are independent of the choice of state and consequences outside this state. This is precisely the role of the following axiom.

A5. Simple Binary Comonotonic Tradeoff Consistency (S-BCTC): For any eight consequences a, b, c, d, x, y, z, w, and states s and t,

$$asx \sim bsy, \quad csx \sim dsy, \quad atz \sim btw \;\Rightarrow\; ctz \sim dtw \tag{2}$$

whenever the sets of acts $\{\ asx, bsy, csx, dsy\ \}$ and $\{\ atz, btw, ctz, dtw\ \}$ are comonotonic, $\{s\}$ is comonotonically non-null on the first set and $\{t\}$ is comonotonically non-null on the second set.

As noted above, only measurement over single states, which employs simple binary comonotonic acts, is involved in the axiom. KW (section 5.2) suggest the following stronger consistency assumption, that applies to all events.

Binary Comonotonic Tradeoff Consistency: For any eight consequences a, b, c, d, x, y, z, w, and *events* E and F,

$$aEx \sim bEy, \quad cEx \sim dEy, \quad aFz \sim bFw \;\Rightarrow\; cFz \sim dFw$$

whenever the sets of acts $\{\ aEx, bEy, cEx, dEy\ \}$ and $\{\ aFz, bFw, cFz, dFw\ \}$ are comonotonic, E is comonotonically non-null on the first set and F is comonotonically non-null on the second set.

KW explain that Binary Comonotonic Tradeoff Consistency, together with a few basic axioms, yields the biseparable representation of Ghirardato and Marinacci [4]. By contrast, the weakened version of tradeoff consistency introduced in this paper does not deliver a global additive representation over simple binary acts, but only local additive representations, as is further elaborated in the presentation of Theorem 1 below.

Theorem 1, presented next, is the main result of this paper. The theorem is broken into four parts. Parts one through three are the sufficiency parts and consist of implications of axioms A1 through A5. The fourth part provides necessity.

The first part of Theorem 1 states that axioms A1 through A5 imply existence of a cardinal utility over consequences, which respects tradeoff indifferences in the sense that indifference of tradeoffs is translated to equivalence of utility differences.

Theorem 1, Part 1 (Sufficiency): Cardinal utility.
Assume a binary relation $\succsim$ on X^S, where X and S satisfy the structural assumptions, A0(a) and A0(b). Then (i) below implies (ii).

(i) $\succsim$ satisfies A1-A5.

(ii) There exists a continuous function $u : X \longrightarrow \mathbb{R}$, unique up to location and unit [3], such that for all $x, y \in X$, $x \succsim y \iff u(x) \geq u(y)$, and $< a; b > \sim^* < c; d >$ implies $u(a) - u(b) = u(c) - u(d)$.

The second part of Theorem 1 states that axioms A1 through A5 are sufficient to obtain local additive representations over simple binary comonotonic sets. Focusing on the more interesting case when both s and its complement are non-null on a simple binary comonotonic set (otherwise the result is trivial), $\succsim$ on this set admits an additive representation $\alpha s \beta \mapsto u(\alpha) + V_{-s}(\beta)$. That is, the additive value function operating on the single state consequence is the utility function from (ii), whereas the second additive value function, operating on the complement's consequence, is general. The only known attributes of V_{-s} are that it is continuous and that it represents the relation over consequences. Nothing more can be said about the link between functions V_{-s} for different states s or about the link between local additive representations on different simple binary comonotonic sets. It is thus impossible to use these additive representations to compare acts that belong to different simple binary comonotonic sets.

As mentioned above, in the special case where the relation also satisfies Binary Comonotonic Tradeoff Consistency, a biseparable representation results. According to the biseparable representation, if $\alpha E \beta$ denotes a bet on event E (with $\alpha \succsim \beta$), then the relation over bets on events admits the additive representation $\alpha E \beta \mapsto u(\alpha)\rho(E) + u(\beta)(1 - \rho(E))$, with u a cardinal utility function and ρ a non-additive probability. The biseparable model therefore implies one global additive representation over all bets on events, in contrast to the model suggested here, which delivers only *local* additive representations (and only on bets on and against states).

Theorem 1, Part 2 (Sufficiency): Local additive representations on simple binary comonotonic sets.

[3] In other words, unique up to a positive linear transformation.

Assume a binary relation $\succsim$ on X^S, where X and S satisfy the structural assumptions, A0(a) and A0(b). Then (i) above also implies:

(iii) Suppose that $\mathcal{C}_s$ is a simple binary comonotonic set of acts of the form $\alpha s \beta$, on which both s and $\{s\}^c$ are non-null. Then given a function u as in (ii), there exists a function
$V_{-s} : X \longrightarrow \mathbb{R} \cup \{-\infty, +\infty\}$, unique up to location[4], such that for all $asx, bsy \in \mathcal{C}_s$,

$$asx \succsim bsy \iff u(a) + V_{-s}(x) \geq u(b) + V_{-s}(y) \ . \tag{3}$$

Moreover, the function V_{-s} is continuous where it is finite and represents $\succsim$ on X.

If only s ($\{s\}^c$) is non-null on $\mathcal{C}_s$ then for all $asx, bsy \in \mathcal{C}_s$, $asx \succsim bsy$ if and only if
$u(a) \geq u(b)$ ($u(x) \geq u(y)$).

The third component of the theorem establishes that the axioms imply a general representation of $\succsim$ over all acts, and that general representation identifies with the utility function over constant acts. Some additional attributes of the representation, necessary for the axioms to be satisfied, are stated. One of these attributes requires a definition.

Definition 5. *A function $J : X^S \longrightarrow \mathbb{R}$ is* self-monotonic *if $J(f) \geq J(g)$ whenever*
$J(\overline{f(s)}) \geq J(\overline{g(s)})$ for all states $s \in S$.

Theorem 1, Part 3 (Sufficiency): A general representation.
Assume a binary relation $\succsim$ on X^S, where X and S satisfy the structural assumptions, A0(a) and A0(b). Then (i) above also implies:

(iv) Given a function u as in (ii), there exists a unique continuous and self-monotonic function $J : X^S \longrightarrow \mathbb{R}$, such that for all $f, g \in X^S$, $f \succsim g \iff J(f) \geq J(g)$, and $J(\bar{x}) = u(x)$ for all $x \in X$.

Furthermore, there are distinct states s' and s'' and consequences x and y such that
$u(x) > J(ys''x)$ and $J(xs'y) > u(y)$.

[4]That is, unique up to an additive constant.

Finally the fourth part of the theorem maintains that statements (ii), (iii) and (iv) are necessary for axioms A1-A5 to be satisfied.

Theorem 1, Part 4 (Necessity).
Assume a binary relation $\succsim$ on X^S, where X and S satisfy the structural assumptions, A0(a) and A0(b). Then (ii),(iii) and (iv) together imply (i).

The next corollary is a supplement to the main theorem, and added since it may prove useful in case the theorem is applied.

Corollary 1. *Assume a binary relation $\succsim$ on X^S, where X and S satisfy the structural assumptions, A0(a) and A0(b). Suppose further that the relation satisfies axioms A1-A5. Then there exists $\delta > 0$ such that if $u(a) - u(b) = u(b) - u(c)$, and $|u(a) - u(b)| < \delta$, then $< a; b > \sim^* < b; c >$.*

4 Example.

The following example depicts a relation that satisfies axioms A1 through A5, hence admits the representation and satisfies the conditions in (ii),(iii) and (iv) of Theorem 1. Nonetheless, the relation in the example is not a biseparable one.

Let $S = \{1, 2, \ldots, n\}$ $(n \geq 3)$ and $X = \mathbb{R}_+$. Suppose that $\succsim$ is represented by the following functional $(0 < \varepsilon < 1)$:

$$J(f) = \varepsilon \left(\frac{1}{2}\sqrt{f(s_1)} + \frac{1}{2}\sqrt{f(s_n)} \right) + (1 - \varepsilon) \left(\frac{1}{n-2} \sum_{i=2}^{n-1} f(s_i) \right),$$

for any act f and an ordering $f(s_1) \succsim f(s_2) \succsim \ldots \succsim f(s_n)$.

First observe that this preference relation satisfies all the axioms listed above. It obviously defines a weak order over $(\mathbb{R}_+)^{\{1,2,\ldots,n\}}$, and, being a continuous, monotonic functional over $\mathbb{R}_+$, the implied preference relation satisfies Monotonicity and Continuity. Moreover, for any two acts f and g, $f \succ g$ whenever $f(s) \succsim g(s)$ in all states s and $f(s) \succ g(s)$ for at least one state s. Thus Consistent Essentiality follows, and all states are non-null on any simple binary comonotonic set. Last, assume that for eight consequences a, b, c, d and x, y, v, w, and two states s and t, $asx \sim bsy$, $csx \sim dsy$ and $atv \sim btw$. The first two indifferences imply

$$\frac{\varepsilon}{2}(\sqrt{a} - \sqrt{b}) = \frac{\varepsilon}{2}(\sqrt{y} - \sqrt{x}) + (1 - \varepsilon)(y - x) = \frac{\varepsilon}{2}(\sqrt{c} - \sqrt{d}) \quad ,$$

while the third one renders

$$\frac{\varepsilon}{2}(\sqrt{a}-\sqrt{b}) = \frac{\varepsilon}{2}(\sqrt{w}-\sqrt{v}) + (1-\varepsilon)(w-v) \ .$$

The indifference $ctv \sim dtw$ is thus implied, and delivers axiom S-BCTC (A5). Still, this preference relation is *not* a biseparable relation. Biseparable preference relations maintain tradeoff consistency across *events*. Therefore, a biseparable preference would satisfy, for instance, that if $asx \sim bsy$, $csx \sim dsy$ and $vta \sim wtb$, then also $vtc \sim wtd$, whenever the sets $\{asx, bsy, csx, dsy\}$ and $\{vta, wtb, vtc, wtd\}$ are comonotonic. This, however, is not the case here, since $\sqrt{a}-\sqrt{b} = \sqrt{c}-\sqrt{d}$ does not imply $\frac{\varepsilon}{2}(\sqrt{a}-\sqrt{b}) + (1-\varepsilon)(a-b) = \frac{\varepsilon}{2}(\sqrt{c}-\sqrt{d}) + (1-\varepsilon)(c-d)$. [5]

5 Comments.

5.1 Infinite values.

Note that the representation (u, V_{-s}) in (3) may obtain a value of $\pm\infty$. Still, an infinite value may only be obtained on *extreme acts*, as defined below.

Definition 6. *A consequence y is termed* minimal *if $x \succsim y$ for all $x \in X$, and* maximal *if $y \succsim x$ for all $x \in X$. A consequence which is either minimal or maximal is called* extreme. *Correspondingly, a* minimal act *is one which obtains a minimal consequence in every state, whereas a* maximal act *obtains a maximal consequence in every state. An act is said to be* extreme *if it is either minimal or maximal.*

If (u, V_{-s}) represents $\succsim$ on $\{\alpha s \beta \mid \alpha \succsim \beta\}$, then V_{-s} may obtain a value of $+\infty$ on a maximal consequence, thus the representation on a maximal act may equal $+\infty$. Similarly, if (u, V_{-s}) represents $\succsim$ on $\{\alpha s \beta \mid \beta \succsim \alpha\}$, then V_{-s} may obtain a value of $-\infty$ on a minimal consequence, consequently the representation may obtain $-\infty$ on a minimal act. Nevertheless, in any case the utility u is bounded.

The following example shows that an infinite value may be obtained under the conditions of the model (The example is based on Example 3.8 from Wakker [13]). Let $X = [0,1]$ and $S = \{1,2,3\}$. Consider a relation $\succsim$ that over simple binary comonotonic sets of the form $\{\alpha s \beta \mid \alpha \succsim \beta\}$ admits the additive representation $\alpha + \beta$, and over simple binary comonotonic sets

[5] The preference relation defined by J satisfies also tradeoff consistency across complements, though that attribute is not required by the axioms: if $xsa \sim ysb$, $xsc \sim ysd$ and $vta \sim wtb$, then also $vtc \sim wtd$.

$\{\alpha s\beta \mid \beta \succsim \alpha\}$ admits the additive representation $\alpha + \ln\beta$. The resulting relation over constant acts is represented by $u(\alpha) = \alpha$, and satisfies that $< a; b > \sim^* < c; d > \Rightarrow\ a - b = c - d$. Indifference curves on the entirety of $[0,1]^3$ may be completed in any monotonic and continuous manner (for the remaining acts (x_1, x_2, x_3) where the x_i's all differ). The obtained relation satisfies the conditions of Theorem 1.

Denote $V_{-s}(\beta) = \ln\beta$. The additive representation $u(\alpha) + V_{-s}(\beta)$ over simple binary comonotonic sets $\{\alpha s\beta \mid \beta \succsim \alpha\}$ returns the value $-\infty$ on the constant act 0. To see that the consequence 0 must be assigned a value of $-\infty$ by V_{-s}, consider the sequence of consequences $\beta_k = e^{-k}$. For this sequence, $V_{-s}(\beta_k) - V_{-s}(\beta_{k+1}) = V_{-s}(\beta_{k+1}) - V_{-s}(\beta_{k+2})$, therefore V_{-s} must be unbounded from below.

5.2 A generalization of the tradeoff consistency axiom.

Let s be some state. Denote by $\mathcal{A}_s^M$ the set of all acts which obtain their best consequence on s (that is, acts f for which $f(s) \succsim f(t)$ for all states $t \in S$), and by $\mathcal{A}_s^m$ the set of all acts which obtain their worst consequence on s. Consider the following tradeoff consistency axiom, which is a strengthening of Simple Binary Comonotonic Tradeoff Consistency (S-BCTC):[6]

Extreme Consequence Tradeoff Consistency: For any four consequences a, b, c, d, four acts f, g, f', g', and states s and t,

$$asf \sim bsg, \quad csf \sim dsg, \quad atf' \sim btg' \ \Rightarrow\ ctf' \sim dtg'$$

whenever all acts $\{asf, bsg, csf, dsg\}$ belong to $\mathcal{A}_s^M$ or all belong to $\mathcal{A}_s^m$, all acts $\{atf', btg', ctf', dtg'\}$ belong to $\mathcal{A}_t^M$ or all belong to $\mathcal{A}_t^m$, and correspondingly $\{s\}$ is non-null on $\mathcal{A}_s^M$ or on $\mathcal{A}_s^m$, and $\{t\}$ is non-null on $\mathcal{A}_t^M$ or on $\mathcal{A}_t^m$.

Extreme Consequence Tradeoff Consistency strengthens S-BCTC as it imposes consistency when acts are not necessarily binary. Although tradeoffs are still measured over single states, now the acts employed in the measurement are any acts which obtain their extreme consequence – either best for all acts or worst for all acts – on these single states. An interesting feature, satisfied under this axiom and axioms A0 through A4, may be identified: by connectedness of X, weak order and continuity, for each act f and state s there exists a *conditional certainty equivalent* $x \in X$ that satisfies $f \sim fsx$. Extreme Consequence Tradeoff Consistency implies that this conditional certainty equivalent is independent of the specific consequence obtained on s, as

[6] I thank an anonymous referee for pointing out to me this generalization and its interest.

long as this consequence is extreme. More accurately, $asf \sim asx$ if and only if $bsf \sim bsx$, whenever both asf and bsf belong to $\mathcal{A}_s^M$ or both belong to $\mathcal{A}_s^m$. This independence is implied by applying the axiom to indifference relationships $asf \sim asf$, $bsf \sim bsf$, and $asf \sim asx$, which result in $bsf \sim bsx$, if $asf, bsf \in \mathcal{A}_s^M$ or $asf, bsf \in \mathcal{A}_s^m$.[7] Further implications of this axiom remain a matter for future research.

6 Proofs.

6.1 Proof of Theorem 1: Sufficiency.

Two standard observations are listed first. These observations will be used in the sequel, sometimes without explicit reference.

Observation 1. *Consistent Essentiality and Monotonicity imply that there are two consequences* $x^*, x_* \in X$ *such that* $x^* \succ x_*$.

Observation 2. *Weak order, Continuity and Monotonicity imply that each act* f *has a certainty equivalent, i.e., a constant act* $\overline{x}$ *such that* $f \sim \overline{x}$.

The proof that (i) of Theorem 1 implies (ii), (iii) and (iv) is conducted in two logical steps. First, an additive representation is shown to hold on simple binary comonotonic sets. Second, a utility is derived, satisfying the conditions in (ii), and yielding the specific representation in (iii) on simple binary comonotonic sets. A general representation as in (iv) follows by employing certainty equivalents.

6.1.1 Additive representation on simple binary comonotonic sets.

This subsection contains a proof that $\succsim$ on simple binary comonotonic sets is represented by an additive functional, every component of which is a representation of $\succsim$ over constant acts.

Let s be some state and $\mathcal{C}_s$ a corresponding simple binary comonotonic set of acts. Suppose that both s and its complement $\{s\}^c$ are non-null on $\mathcal{C}_s$.

Lemma 2. *The binary relation* $\succsim$ *on* $\mathcal{C}_s$ *is a continuous weak order, satisfying Monotonicity.*

Proof. All attributes on $\mathcal{C}_s$ follow from their counterparts on X^S, assumed in A1, A2 and A3. ■

[7] The other acts involved must, by the other axioms, belong to the same sets. Some nullity conditions are omitted for the sake of brevity.

The proof that $\succsim$ on $\mathcal{C}_s$ admits an additive representation makes use of Theorem 3.2 and Remark 3.7 from Wakker [13]. In order to apply these results two additional attributes are required.

Definition 7. *The binary relation $\succsim$ satisfies the* ***hexagon condition*** *on $\mathcal{C}_s$ if, for all acts $\nu s\alpha,\ \mu s\beta,\ \nu s\beta,\ \mu s\gamma,\ \xi s\alpha,\ \nu s\beta,\ \xi s\beta,\ \nu s\gamma$ in $\mathcal{C}_s$,*

$$\nu s\alpha \sim \mu s\beta, \quad \nu s\beta \sim \mu s\gamma, \quad \xi s\alpha \sim \nu s\beta \quad \Rightarrow \quad \xi s\beta \sim \nu s\gamma \ .$$

Definition 8. *The binary relation $\succsim$ satisfies* ***strong monotonicity*** *on $\mathcal{C}_s$ if, for all consequences α, β and acts $\alpha sx, \beta sx, ys\alpha$ and $ys\beta$ in $\mathcal{C}_s$,*

$$\begin{aligned} \alpha \succsim \beta \quad &\Leftrightarrow \quad \alpha sx \succsim \beta sx \ \text{, and} \\ \alpha \succsim \beta \quad &\Leftrightarrow \quad ys\alpha \succsim ys\beta \ . \end{aligned}$$

Claim 3. *The binary relation $\succsim$ on $\mathcal{C}_s$ satisfies strong monotonicity and the hexagon condition.*

Proof. Strong monotonicity follows from Monotonicity, and the assumption that both s and $\{s\}^c$ are non-null on $\mathcal{C}_s$ together with part (a) of Consistent Essentiality. The hexagon condition is implied by assigning in (2) both states to s, and $a = d = \nu,\ b = \mu,\ c = \xi$,
$x = \alpha,\ y = v = \beta,\ w = \gamma$. ■

Lemma 4. *(A combination of Theorem 3.2 and Remark 3.7 from Wakker [13], applied to $\mathcal{C}_s$) The following two statements are equivalent:*

(a) The binary relation $\succsim$ on $\mathcal{C}_s$ is a continuous weak order, satisfying strong monotonicity and the hexagon condition.

(b) There are functions $V_s, V_{-s} : X \to \mathbb{R} \cup \{-\infty, +\infty\}$, such that the binary relation $\succsim$ on $\mathcal{C}_s$ is represented by the additive functional:

$$(V_s, V_{-s})(\alpha s\beta) = V_s(\alpha) + V_{-s}(\beta) \ , \tag{4}$$

with V_s and V_{-s} that are continuous where they are finite.

In case $\mathcal{C}_s = \{\alpha s\beta \mid \alpha \succsim \beta\}$ then V_s and V_{-s} are finite, except possibly V_s at minimal consequences and V_{-s} at maximal consequences. In case $\mathcal{C}_s = \{\alpha s\beta \mid \beta \succsim \alpha\}$, then V_s and V_{-s} are finite, except possibly V_s at maximal consequences and V_{-s} at minimal consequences.

The representation (V_s, V_{-s}) *is unique up to locations and joint unit on* $\mathcal{C}_s \setminus \{$*extreme acts*$\}$.[8]

Remark 5. *If values (finite or infinite) at extreme consequences are chosen in the following manner, then the conditions of Lemma 4 are still satisfied (see section 4.2 in Wakker [13]): In case* $\mathcal{C}_s = \{\alpha s \beta \mid \alpha \succsim \beta\}$*, if an extreme consequence exists, set* $V_s(\textit{minimal consequence}) = \inf\{V_s(x) \mid x \in X \setminus \{\textit{minimal consequences}\}\}$ *for any minimal consequence, and* $V_{-s}(\textit{maximal consequence}) = \sup\{V_{-s}(x) \mid x \in X \setminus \{\textit{maximal consequences}\}\}$ *for any maximal consequence. In case* $\mathcal{C}_s = \{\alpha s \beta \mid \beta \succsim \alpha\}$*, if an extreme consequence exists, set* $V_s(\textit{maximal consequence}) = \sup\{V_s(x) \mid x \in X \setminus \{\textit{maximal consequences}\}\}$ *for any maximal consequence, and* $V_{-s}(\textit{minimal consequence}) = \inf\{V_{-s}(x) \mid x \in X \setminus \{\textit{minimal consequences}\}\}$ *for any minimal consequence.*

In the remainder of the proof, whenever there is a reference to an additive representation over a simple binary comonotonic set, it is assumed that values on extreme consequences are set as in the remark.

Conclusion 6. *Each of* V_s, V_{-s} *represents* $\succsim$ *on* X.

Proof. Suppose that $\mathcal{C}_s = \{\alpha s \beta \mid \alpha \succsim \beta\}$. By strong monotonicity and the non-nullity assumptions, applying the additive representation,

$$\alpha \succsim \beta \;\Leftrightarrow\; V_s(\alpha) + V_{-s}(\beta) \geq V_s(\beta) + V_{-s}(\beta) \;.$$

In case β is not maximal, $V_{-s}(\beta)$ is finite and $V_s(\alpha) + V_{-s}(\beta) \geq V_s(\beta) + V_{-s}(\beta) \Leftrightarrow V_s(\alpha) \geq V_s(\beta)$. If β is maximal then $\alpha \succsim \beta \Leftrightarrow \alpha \sim \beta$, and by Monotonicity and Consistent Essentiality there exists $x \in X$ such that $\beta \succ x$. Strong Monotonicity yields $\alpha \sim \beta \Leftrightarrow \alpha s x \sim \beta s x$, yielding $V_s(\alpha) + V_{-s}(x) = V_s(\beta) + V_{-s}(x) \Leftrightarrow V_s(\alpha) = V_s(\beta)$, as $V_{-s}(x)$ is finite.

Similarly,

$$\alpha \succsim \beta \;\Leftrightarrow\; V_s(\alpha) + V_{-s}(\alpha) \geq V_s(\alpha) + V_{-s}(\beta) \;.$$

In case α is not minimal, $V_s(\alpha)$ is finite and the above holds if and only if $V_{-s}(\alpha) \geq V_{-s}(\beta)$. If α is minimal then so is β, and there exists $y \in X$ such that $y \succ \alpha$. Strong Monotonicity again yields $\alpha \sim \beta \Leftrightarrow y s \alpha \sim y s \beta$, implying $V_s(y) + V_{-s}(\alpha) = V_s(y) + V_{-s}(\beta)$. As $V_s(y)$ is finite, it follows that $V_{-s}(\alpha) = V_{-s}(\beta)$, and V_{-s} too represents $\succsim$ on X.

The proof for the case $\mathcal{C}_s = \{\alpha s \beta \mid \beta \succsim \alpha\}$ is completely analogous. ■

[8] That is, if (W_s, W_{-s}) is another additive representation as above, then $W_s = \sigma u + \mu^1$, $W_{-s} = \sigma V_{-s} + \mu^2$, $\sigma > 0$, on non-extreme acts.

6.1.2 Completion of the proof that (i) implies (ii), (iii) and (iv).

Lemma 7. *(A partial version of Lemma VI.8.2 from Wakker [12]):*
Let $\mathcal{C}_s$ and $\mathcal{C}_t$ be two simple binary comonotonic sets of acts. Assume that s and its complement $\{s\}^c$ are non-null on $\mathcal{C}_s$, and that t and $\{t\}^c$ are non-null on $\mathcal{C}_t$. Let (V_s, V_{-s}) and (V_t, V_{-t}) be the implied continuous cardinal additive representations of $\succsim$ on $\mathcal{C}_s$ and $\mathcal{C}_t$, respectively (according to Lemma 4). Then on $X \setminus \{$extreme consequences$\}$, $V_t = \sigma V_s + \tau$, $\sigma > 0$.

Proof. By applying Lemma VI.8.2 of Wakker [12] for V_s, V_t (in his notation, for $V_l^\pi = V_t$,
$V_k^{\pi'} = V_s$) on $X \setminus \{$extreme consequences$\}$ (in case extreme consequences exist, Proposition VI.9.5 from Wakker [12] guarantees that the required topological conditions on $X \setminus \{$extreme consequences$\}$ are satisfied, and the Lemma may be applied there). ■

Let s' be a state, and x, y consequences, as characterized in part (b) of Consistent Essentiality. By assumption, $xs'y \succ \overline{y}$, which by Monotonicity implies $x \succ y$. Part (b) of Consistent Essentiality also states that for some other state s'', $\overline{x} \succ ys''x$. Applying Monotonicity once more, $ys''x \succsim xs'y$, hence $\overline{x} \succ xs'y \succ \overline{y}$, implying that both s' and $\{s'\}^c$ are non-null on $\{\alpha s'\beta \mid \alpha \succsim \beta\}$. According to the proof in the previous subsection there exists a continuous cardinal additive representation $(V_{s'}, V_{-s'})$ of $\succsim$ on the simple binary comonotonic set in question, where $V_{s'}, V_{-s'}$ represent $\succsim$ on X. Let $u(x) = V_{s'}(x)$ for all $x \in X \setminus \{$minimal consequences$\}$. If a minimal consequence exists, set $u(\text{minimal consequence}) = \inf\{u(x) \mid x \in X \setminus \{\text{minimal consequences}\}\}$. According to Lemma 4, Conclusion 6 and Remark 5, u represents $\succsim$ on X, it is finite except possibly at minimal consequences, and continuous where it is finite.

Let s'' be a state as characterized in part (b) of Consistent Essentiality. Applying Monotonicity as is done above, both s'' and $\{s''\}^c$ are non-null on $\{\alpha s''\beta \mid \beta \succsim \alpha\}$, thus according to the previous subsection there exists a continuous cardinal additive representation $(V_{s''}, V_{-s''})$ of $\succsim$ on this simple binary comonotonic set, where $V_{s''}, V_{-s''}$ represent $\succsim$ on X, and $V_{s''}$ is finite except possibly at maximal consequences. By Lemma 7 $V_{s''}$ is a positive linear transformation of u on non-extreme consequences. If a minimal consequence exists, then $V_{s''}$, being finite there, is bounded from below. Hence u is bounded from below and by its definition on minimal consequences must be finite on those as well. The definition of u on minimal consequences extends its continuity to X.

It follows that $\succsim$ on X is represented by a continuous utility function, $u : X \to \mathbb{R}$. Uniqueness of u and the fact that it respects tradeoff indifferences

is proved in the following lemma.

Lemma 8. *If* $< a; b > \sim^* < c; d >$*, then* $u(a) - u(b) = u(c) - u(d)$*.* u *is unique up to a positive linear transformation.*

Proof. If $< a; b > \sim^* < c; d >$, then there exist consequences x, y and a state t such that $atx \sim bty$ and $ctx \sim dty$, with $\{atx, bty, ctx, dty\}$ comonotonic, and $\{t\}$ comonotonically non-null on this set of acts. If $\{t\}^c$ is comonotonically *null* on $\{atx, bty, ctx, dty\}$, then having $bty \sim btx$ delivers $atx \sim bty \;\Leftrightarrow\; a \sim b$, and similarly $ctx \sim dty \;\Leftrightarrow\; c \sim d$. The equality $u(a) - u(b) = u(c) - u(d)$ is trivially implied.

Otherwise, assume that $\{t\}^c$ is comonotonically non-null on $\{atx, bty, ctx, dty\}$. By the above subsection there exists a continuous additive representation (V_t, V_{-t}) for $\succsim$ on the simple binary comonotonic set containing $\{atx, bty, ctx, dty\}$. Translating the indifference relationships with this representation obtains $V_t(a)+V_{-t}(x) = V_t(b)+V_{-t}(y)$ and $V_t(c)+V_{-t}(x) = V_t(d)+V_{-t}(y)$. If V_t or V_{-t} obtain a value of $\pm\infty$ on one of these consequences, then the comonotonicity restrictions and Consistent Essentiality imply that $a \sim b$ and $c \sim d$, immediately implying the required result. Else, if all values involved are finite, $V_t(a) - V_t(b) = V_t(c) - V_t(d)$. According to Lemma 7, $V_t = \sigma V_{s'} + \tau,\ \sigma > 0$, on non-extreme consequences, thus, as all values are finite and recalling the definition on extreme consequences, also $u(a) - u(b) = u(c) - u(d)$.

For uniqueness of u, let $\hat{u}$ be some other representation of $\succsim$ on X which satisfies that if $< a; b > \sim^* < c; d >$ then $\hat{u}(a) - \hat{u}(b) = \hat{u}(c) - \hat{u}(d)$. Both u and $\hat{u}$ represent $\succsim$ on X, therefore $\hat{u} = \varphi \circ u$, with φ a continuous and strictly increasing transformation. Non-triviality of $\succsim$, connectedness of X and continuity of u imply that $u(X) = V_{s'}(x)$ is a non-degenerate interval. For the same reasons $V_{-s'}(X)$ is a non-degenerate interval. Let ξ be an internal point of $u(X)$. There exists an interval R small enough around ξ such that, for all $u(a), u(c)$ and $u(b) = [u(a)+u(c)]/2$ in R, there are $x, y \in X$ for which $as'x, bs'y, bs'x, cs'y$ are comonotonic, with $a \succsim x,\ b \succsim y,\ b \succsim x$ and $c \succsim y$, and

$$u(a)-u(b) = V_{s'}(a)-V_{s'}(b) = V_{-s'}(y)-V_{-s'}(x) = V_{s'}(b)-V_{s'}(c) = u(b)-u(c).$$

That is, $as'x \sim bs'y$ and $bs'x \sim cs'y$ with $\{as'x, bs'y, bs'x, cs'y\}$ comonotonic, which is precisely the definition of $< a; b > \sim^* < b; c >$. By the assumption on $\hat{u}$, $\hat{u}(a) - \hat{u}(b) = \hat{u}(b) - \hat{u}(c)$ as well, implying that for all $\alpha, \gamma \in R$, φ satisfies $\varphi((\alpha + \gamma)/2) = [\varphi(\alpha) + \varphi(\gamma)]/2$. By Theorem 1 of section 2.1.4 of Aczel [1], φ must be a positive linear transformation. ■

Let s be some state and $\mathcal{C}_s$ a corresponding simple binary comonotonic set of acts. If both s and its complement $\{s\}^c$ are non-null on $\mathcal{C}_s$, then

according to Lemma 4, Lemma 7 and the definition of u, $\succsim$ on $\mathcal{C}_s$ admits a representation as in (4) (by choosing for V_s the same location and unit as those of u). If only s is non-null on $\mathcal{C}_s$ then for every asx, bsy in $\mathcal{C}_s$, $asx \succsim bsy$ if and only if $a \succsim b$. Since u represents $\succsim$ on X, it follows that $asx \succsim bsy$ if and only if $u(a) \geq u(b)$. If only $\{s\}^c$ is non-null on $\mathcal{C}_s$ then $asx \succsim bsy$ if and only if $x \succsim y$, which is true if and only if $u(x) \geq u(y)$.

Given u, a unique continuous representation J may be defined using certainty equivalents: for every act f, set $J(f) = u(x)$ for $x \in X$ such that $f \sim \bar{x}$ (such a consequence x exists by Observation 2). The resulting functional J represents $\succsim$ on X^S, is unique given a specific u, and continuous due to continuity of $\succsim$. Self-monotonicity of J follows from Monotonicity (A3). Part (b) of Consistent Essentiality (A4) implies that there are states s' and s'' and consequences x and y such that $J(x) = u(x) > J(ys''x)$ and $J(xs'y) > J(y) = u(y)$.

6.2 Proof of Theorem 1: Necessity.

Let (ii), (iii) and (iv) of Theorem 1 hold. By (iv), $\succsim$ satisfies Weak Order (A1), Continuity (A2) and Monotonicity (A3). By existence of states s', s'' and consequences x, y as detailed in (iv), part (b) of A4 holds. Part (a) of A4 is implied by (iii).

For S-BCTC (A5), let consequences a, b, c, d and x, y, v, w, and states s and t, be such that

$$asx \sim bsy, \quad csx \sim dsy, \quad atv \sim btw \tag{5}$$

where the sets $\{\ asx, bsy, csx, dsy\ \}$ and $\{\ atv, btw, ctv, dtw\ \}$ are comonotonic, $\{s\}$ is comonotonically non-null on the first set and $\{t\}$ is comonotonically non-null on the second set. The first two indifference relations imply, by definition, $< a; b > \sim^* < c; d >$, and thus, according to (ii), $u(a) - u(b) = u(c) - u(d)$. If $\{t\}^c$ is comonotonically null on $\{\ atv, btw, ctv, dtw\ \}$, then $atv \sim btw$ implies $u(a) = u(b)$, which renders $u(a) - u(b) = u(c) - u(d) = 0$ and thus $c \sim d$. From nullity of $\{t\}^c$ and Monotonicity it follows that $ctv \sim ctw \sim dtw$. Otherwise, by Lemma 7, $V_t = \sigma u + \tau$, $\sigma > 0$, hence $V_t(a) - V_t(b) = V_t(c) - V_t(d)$. The indifference $atv \sim btw$ implies $V_t(a) - V_t(b) = V_{-t}(w) - V_{-t}(v)$, thus also $V_t(c) - V_t(d) = V_{-t}(w) - V_{-t}(v)$, yielding $ctv \sim dtw$.

6.3 Proof of Corollary 1.

Let a, b, c be consequences satisfying $u(a) - u(b) = u(b) - u(c)$, and assume w.l.o.g. $u(a) > u(b) > u(c)$, so equivalently $a \succ b \succ c$. Suppose further that

$u(a) - u(b) < \delta$. Let s' be a state satisfying $\bar{x} \succ xs'y \succ \bar{y}$ whenever $x \succ y$ (exists by Essentiality (A4) and Monotonicity), and $(u, V_{-s'})$ the corresponding continuous cardinal additive representation of $\succsim$ on $\{\alpha s'\beta \mid \alpha \succsim \beta\}$. Similarly, let s'' be a state satisfying $\bar{x} \succ ys''x \succ \bar{y}$ whenever $x \succ y$, and $(u, V_{-s''})$ the corresponding continuous cardinal additive representation of $\succsim$ on $\{\alpha s''\beta \mid \beta \succsim \alpha\}$.

Employing the above representations, tt suffices to show that there are consequences x, y that satisfy either

$$u(a) - u(b) = V_{-s'}(y) - V_{-s'}(x), \quad a \succ b \succ c \succsim y \succ x \tag{6}$$

or

$$u(a) - u(b) = V_{-s''}(y) - V_{-s''}(x), \quad y \succ x \succsim a \succ b \succ c\ . \tag{7}$$

Note that according to the conditions on the consequences and Lemma 4, $V_{-s'}$ and $V_{-s''}$ in the above equations must obtain finite values.

Let $z_1 > z_2 > z_3 > z_4$ be interior consequences (there are x^*, x_* for which $x^* \succ z_1 \succ z_2 \succ z_3 \succ z_4 \succ x_*$). If $u(a) - u(c) < \min_{i=1,2,3}(u(z_i) - u(z_{i+1}))$, then either $a \prec z_2$, or $a \succsim z_2$ and $c \succ z_3$. In the first case, letting $u(a) - u(b) < V_{-s''}(z_1) - V_{-s''}(z_2)$ guarantees existence of x, y as required by (7). In the second case, $u(a) - u(b) < V_{-s'}(z_3) - V_{-s'}(z_4)$ guarantees that there are consequences x, y such that (6) is satisfied. Hence, choosing

$$\delta = \min\{\frac{1}{2}\min_{i=1,2,3}(u(z_i) - u(z_{i+1})), V_{-s'}(z_3) - V_{-s'}(z_4), V_{-s''}(z_1) - V_{-s''}(z_2)\}$$

yields the desired result. That is, if $u(a) - u(b) < \delta$ then there are x, y satisfying either (6) or (7), implying $< a; b > \sim^* < b; c >$. ■

Acknowledgments.

The author is thankful to David Schmeidler and Peter Wakker for very helpful and detailed remarks and advice, and to an anonymous referee and an associate editor for many useful suggestions and comments. Thanks are also due to the Israeli Science Foundation for partial financial support.

References

[1] Aczel, J. (1966), Lectures on functional equations and their applications. Academic Press, New York.

[2] Blaschke, W. and G. Boll (1938), Geometric der Gewebe. Springer-Verlag, Berlin.

[3] Debreu, Gérard (1960), "Topological Methods in Cardinal Utility Theory." In Kenneth J. Arrow, Samuel Karlin, & Patrick Suppes (1960, Eds), Mathematical Methods in the Social Sciences, 16Uf02d26, Stanford University Press, Stanford, CA.

[4] Ghirardato, P. and M. Marinacci (2001), "Risk, Ambiguity, and the Separation of Utility and Beliefs", *Mathematics of Operations Research* 26, 864-890.

[5] Gilboa, I. and D. Schmeidler (1989), "Maxmin expected utility with nonunique prior", *Journal of Mathematical Economics*, 18, 141-153.

[6] Kobberling, V. and P. Wakker (2003), "Preference foundations for nonexpected utility: A generalized and simplified technique", *Mathematics of Operations Research*, 28, 395-423.

[7] Kobberling, V. and P. Wakker (2004), "A Simple Tool for Qualitatively Testing, Quantitatively Measuring, and Normatively Justifying Savage's Subjective Expected Utility", *Journal of Risk and Uncertainty*, 28, 135-145.

[8] Krantz, D.H., R.D. Luce, P. Suppes, and A. Tversky (1971), Foundations of Measurement. New York: Academic Press.

[9] Luce, R.D. (2000), Utility of Gains and Losses: Measurement-Theoretical and Experimental Approaches. Lawrence Erlbaum Publishers, London.

[10] Schmeidler, D. (1989), "Subjective probability and expected utility without additivity", *Econometrica*, Vol. 57, No. 3, 571-587.

[11] Thomsen, G. (1927), "Un teorema topologico sulle schiere di curve e una caratterizzazione geometrica delle superficie isotermo-asintotiche", *Bollettino della Unione Matematica Italiana*, 6, 80-85.

[12] Wakker, P. (1989), Additive Representations of Preferences, A New Foundation of Decision Analysis. Kluwer Academic Publishers, Dordrecht, The Netherlands.

[13] Wakker, P. (1993), "Additive representations on rank-ordered sets II. The topological approach.", *Journal of Mathematical Economics*, 22, 1-26.

[14] Deneffe, D. and P. Wakker (1996), "Eliciting von Neumann-Morgenstern Utilities when Probabilities Are Distorted or Unknown", *Management Science*, 42, 11311150.

Subjective multi-prior probability: A representation of a partial likelihood relation*

Shiri Alon† and Ehud Lehrer‡

April 1, 2013

Abstract

This paper deals with an incomplete relation over events. Such a relation naturally arises when likelihood estimations are required within environments that involve ambiguity, and in situations which engage multiple assessments and disagreement among individuals' beliefs. The main result characterizes binary relations over events, interpreted as likelihood relations, that can be represented by a unanimity rule applied to a set of prior probabilities. According to this representation an event is at least as likely as another if and only if there is a consensus among all the priors that this is indeed the case. A key axiom employed is a cancellation condition, which is a simple extension of similar conditions that appear in the literature.

Keywords: Multi-prior probability, incomplete relation, ambiguity, cancellation.
JEL classification: D81

*The authors acknowledge helpful discussions with Gabi Gayer, Tzachi Gilboa, Fabio Maccheroni, David Schmeidler and Peter Wakker.

†Department of Economics, Faculty of Social Sciences, Bar-Ilan University, Ramat-Gan 5290002, Israel. e-mail: shiri.aloneron@gmail.com

‡The School of Mathematical Sciences, Tel Aviv University, Tel Aviv 69978, Israel and INSEAD, Boulevard de Constance, 77305 Fontainebleau, France. e-mail: lehrer@post.tau.ac.il

Lehrer acknowledges support of the ISF through Grant #538/11.

1 Introduction

1.1 Motivation and Background

Estimating the odds and comparing the likelihood of various events are essential parts of processes carried out by many organizations. For instance, the US intelligence community produces National Intelligence Estimates, in which the likelihood of various events is assessed. Questions such as, 'Is it more likely that democracy will prevail in Libya, or that a military regime will be established?', and the like, seem natural to ask. These questions and many others call for likelihood comparisons of different events. Other examples in which probabilistic estimates are used include forecasts published by central banks, that address issues such as the odds of inflation or recession, assessments of market trends supplied by committees of experts, estimated likelihood of natural events such as global warming, that are based on individual opinions of scientists and on many experiments, and so forth. In all these instances, statements of the type 'event A is more likely than event B' seem fundamental to the respective context.

Typically, assessments of the kind given above rely on 'objective' data, such as reports of military movements, temperature measurements and the like. Such assessments are usually intended to be based as closely as possible on the data, in order to be considered 'objective'. Frequently, though, the events examined involve some degree of ambiguity. Knowledge or available information might be insufficient to determine which of two events under consideration is more likely. For example, due to insufficient or conflicting data about customers' preferences, a committee of experts might be unable to determine which of the smart phones and the tablet PCs markets will grow more rapidly in the next few years. A likelihood relation in such situations might therefore leave the comparison between some pairs of events unspecified. A question arises, as to what kind of probabilistic representation, if any, may describe such likelihood relations, and reflect indeterminacy with regard to some pairs of events.

The Bayesian approach suggests that an individual's likelihood relation over events can be represented by a prior probability measure. This probability is subjective in nature, as it emerges from subjective likelihood comparisons. Ramsey (1931), de Finetti (1931,1937), Savage (1954), and others who followed in their footsteps, introduced conditions on a binary relation over events that characterize when it may be represented by a prior probability. However, situations in which some events are incomparable cannot be captured by a single prior probability, as a single probability determines the likelihood order

between any two events.

Arguments that question the validity of the completeness assumption of preference relations, are by now well-known. Von Neumann and Morgenstern (1944) already doubted that an individual can always decide from among all alternatives. Aumann (1962) stated that 'of all the axioms of utility theory, the completeness axiom is perhaps the most questionable', and doubted completeness on descriptive as well as normative grounds.

In later models, involving alternatives with unknown probabilities, completeness was challenged based on ambiguity considerations. The leading rationale was that when the decision situation is unclear (due, for instance, to lack of information) an individual might be unable, or unwilling, to make decisions among some alternatives. In most of these models, ambiguity was reflected in a difficulty to assess probabilities and choose one prior probability to describe the decision maker's belief. As a result, belief was represented by a set of prior probabilities. Those models that are closely related to the current work are further discussed in subsection 1.4 below.

Notwithstanding the above arguments, incompleteness in decisions is often criticized on the grounds that a decision should ultimately be made, and that 'no decision' is in itself a decision. However, when likelihood judgements are concerned it is not clear how an individual can be forced to express an honest opinion regarding the likelihood comparison of events. For reasons of ambiguity as mentioned above, an individual may feel unable to tell which of two events is more likely, and prefer to remain silent by declaring them to be incomparable. Contrary to decision making, refraining from judgement in this context cannot be claimed to be a likelihood judgement by itself. Incompleteness of a likelihood relation over events is therefore a reasonable and even a desirable assumption.

1.2 Subjective multi-prior probability

This paper proposes a characterization of a subjective probability, which differs from the aforementioned classical works. Here, completeness is not assumed. The paper formulates conditions on a binary relation over events, that are necessary and sufficient for the relation to be represented by a *set* of prior probabilities. This set of priors should be considered subjective, as it emerges from subjective likelihood comparisons. As an aside, we comment that a binary relation over events, that is taken here as a primitive of the model, can easily be deduced from a binary relation over bets on events (as it is in fact done by Savage (1954), for instance). Taking a binary relation over bets as a primitive, event A would be

considered as more likely than event B whenever a bet on A is preferred to a bet on B. [1]

The nature of the representing set of priors is that one event is considered at least as likely as another, precisely when all the priors in the representing set agree that this is the case. When the representing set contains more than one prior, there may exist pairs of events over which there is no consensus among the priors. Hence, the likelihood order induced by two or more different priors might be incomplete.

Formally, let $\succsim$ denote a binary relation over events, where $A \succsim B$ for events A and B is interpreted as 'A is at least as likely as B'. The main theorem of the paper introduces necessary and sufficient conditions ('axioms') on $\succsim$ that guarantee the existence of a set of prior probabilities, $\mathcal{P}$, for which,

$$A \succsim B \quad \Leftrightarrow \quad \mu(A) \geq \mu(B) \text{ for all } \mu \in \mathcal{P}. \tag{1}$$

When this happens it is said that the 'at least as likely as' relation admits a *subjective multi-prior probability* representation.

The representing set of prior probabilities is not necessarily unique. This is the reason why reference is made to the maximal, w.r.t. inclusion, representing set of priors. In the case of a finite state set, the set of priors is not restricted in any sense. However, when the state space is infinite, the priors should agree on any null-set, or more formally, to satisfy a kind of uniform absolute continuity condition. Still, even in this case, the priors need not agree on the probability of any event rather than null and universal events.

The main result of the paper is separated into two cases: when the state space is finite and when it is infinite. While for the finite case three basic assumptions and the main axiom, which is a Cancellation condition, suffice to characterize those relations that are representable by a multi-prior probability, the infinite case requires an additional assumption regarding the richness of the state space, akin to the Archimedean assumption of Savage (his P6). The Cancellation condition used is discussed in the next subsection.

The use of a set of prior probabilities to describe an individual's belief elicits likelihood judgements that are more robust to ambiguity than a description by a single probability measure. Suppose that an individual is uncertain as to the prior probability to choose. Choice of a single prior probability is sensitive to uncertainty in the sense that two different probabilities might exhibit reversal of likelihood order between some pair of events. On the other hand, representing the uncertainty through a set of prior probabilities leaves

[1] An appropriate consistency condition, such as Savage's postulate P4, should obviously be imposed. See elaboration in Subsection 4.1.

such conflicting comparisons undetermined, reflecting the individual's lack of knowledge regarding the 'right' prior to choose.

The technique of employing sets of probabilities to produce robust models is used extensively in Bayesian Analysis. Robust Bayesian Analysis replaces a single Bayesian prior with a set of priors so as to make the statistical model less sensitive to the prior chosen. A well-known example of this approach is the ε-contamination class (see, for instance, Berger 1994). The axiomatization presented here may therefore lend normative ground to these robust, multi-prior Bayesian techniques[2].

Two additional results, somewhat complementary to the elicitation of a subjective multi-prior probability, are contained in the paper. One compares the ambiguity perceived by two individuals, holding subjective multi-prior probability beliefs, by means of inclusion of sets of priors. One individual is considered to perceive more ambiguity than another if he or she is less decisive, that is, leaves more comparisons undetermined. In the representation, the maximal (w.r.t. set inclusion) set of probabilities of the first individual contains that of the second individual. The other result proposes a scheme to complete a partial likelihood relation, that obtains a subjective multi-prior probability. This scheme suggests aversion to ambiguity.

1.3 Axiomatization

Some notation is needed to facilitate the following discussion. Let S denote a nonempty state-space with a typical element s, and Σ an algebra of events over S. For an event E, $\mathbf{1}_E$ denotes the indicator function[3] of E. A binary relation $\succsim$ is defined over Σ, with $\succ$ denoting its asymmetric part. A probability measure P *agrees* with $\succsim$ if it represents it, in the sense that $A \succsim B \quad \Leftrightarrow \quad P(A) \geq P(B)$. A probability measure P *almost agrees* with $\succsim$ if the former equivalence is relaxed to $A \succsim B \quad \Rightarrow \quad P(A) \geq P(B)$. In other words, P almost agrees with $\succsim$ if it cannot be the case that $A \succsim B$, and at the same time, $P(B) > P(A)$. It is possible, however, to have $P(A) \geq P(B)$ yet $\neg(A \succsim B)$.

De Finetti introduced four basic postulates that must be satisfied by an 'at least as likely as' relation. These postulates define an entity known since as a *qualitative probability*. The four postulates are:

[2]We thank Tzachi Gilboa for bringing this example to our attention

[3]That is, $\mathbf{1}_E$ is the function that attains the value 1 on E and 0 otherwise.

Complete Order:	The relation $\succsim$ is complete and transitive.
Cancellation:	For any three events, A, B and C, such that $A \cap C = B \cap C = \emptyset,\ \ A \succsim B \quad \Leftrightarrow \quad A \cup C \succsim B \cup C.$
Positivity:	For every event A, $A \succsim \emptyset$.
Non Triviality:	$S \succ \emptyset$.

De Finetti assumed that the 'at least as likely as' relation is complete, and, in addition, satisfies three basic assumptions: Transitivity; Positivity, which states that any event is at least as likely as the empty event; and Non-triviality, which states that 'all' (the universal event) is strictly more likely than 'nothing' (the empty event).

Beyond these three assumptions, de Finetti introduced another, more substantial postulate, Cancellation. Cancellation implies a form of separability over events in the following sense. Any event has its own likelihood weight, which is unrelated to other disjoint events. In other words, any event has the same marginal contribution, no matter what other disjoint events it is annexed to[4].

The four de Finetti's assumptions are necessary for the relation to have an agreeing probability. de Finetti posed the question whether these postulates are also sufficient to guarantee existence of an agreeing probability, or even of an almost agreeing probability.

The question was answered negatively by Kraft et al.(1959). They introduced a counter example with a relation over a finite state space, that satisfies all the above conditions and yet has no almost agreeing probability. Kraft et al.(1959) suggested a strengthening of the Cancellation condition, and showed that with a finite state space, the strengthened condition together with the above basic conditions imply that the relation has an almost agreeing probability. Later on, other conditions that derive agreeing or almost agreeing probabilities were proposed by Scott (1964), Kranz et al. (1971) and Narens (1974). These conditions stated that for two sequences of events, $(A_i)_{i=1}^n$ (the A-sequence) and $(B_i)_{i=1}^n$ (the B-sequence),

$$\begin{aligned} &\text{If} \quad \sum_{i=1}^{n} \mathbf{1}_{A_i}(s) = \sum_{i=1}^{n} \mathbf{1}_{B_i}(s) \ \text{ for all } s \in S, \\ &\text{and} \quad A_i \succsim B_i \quad \text{for } i = 1, \ldots, n-1, \\ &\text{then} \ \ B_n \succsim A_n. \end{aligned} \tag{2}$$

[4]This assumption is violated, for instance, when states are evaluated through a nonadditive probability v. Under such an evaluation, the marginal contribution of event C when added to event A, $v(A \cup C) - v(A)$, is not necessarily the same as its marginal contribution when added to event B, $v(B \cup C) - v(B)$.

The idea behind Finite Cancellation is similar in essence to that lying at the basis of de Finetti's Cancellation condition. Like Cancellation, Finite Cancellation is based on the assumption that each state has always the same marginal contribution of likelihood, no matter to which other states it is added. However, Finite Cancellation takes this rationale a step further.

The equality, that appears in the axiom, between the two sums of indicators, means that each state appears the same number of times in each of the sequences. Following the rationale explained above, the weight of each state does not depend on the order and manner in which it appears in each sequence. Hence, the equality between the two sums suggests that it cannot be that the A-sequence has an overall likelihood weight greater than that of the B-sequence. In order to explain this point we turn to an analogy from accounting.

Consider a simple double-entry booking procedure with credit and debit that express likelihood weights. In each of the $n-1$ first lines of the ledger there are pairs of events with $A_i \succsim B_i$: A_i is recorded on the credit side, while B_i on the debit side. Another pair of events A_n and B_n is now considered. Suppose that A_n is recorded on the credit column and B_n on the debit column. This reflects the fact that A_n is at least as likely as B_n. It is argued that this ranking contradicts the rationale behind Cancellation, of invariant marginal contribution.

When A_n is ranked more likely than B_n, each event in the A-sequence is as likely as its counterpart in the B-sequence. Therefore, the credit balance outweighs the debit balance. Such an assertion contradicts the conclusion that the two sequences should have the same overall likelihood weight, unless each pair consists of equally likely events. It implies, in particular, that if A_n and B_n are comparable, it must be that $B_n \succsim A_n$. Finite Cancellation explicitly states that A_n and B_n are comparable, and that B_n is at least as likely as A_n. For further details on Cancellation axioms and almost agreeing probabilities see Fishburn (1986) for a thorough survey, and Wakker (1981) for a discussion and related results.

This paper follows the path of studies described above and introduces an axiom termed *Generalized Finite Cancellation* (GFC). This axiom postulates that for two sequences of

events, $(A_i)_{i=1}^n$ and $(B_i)_{i=1}^n$, and an integer $k \in \mathbb{N}$,

$$\begin{aligned}
&\text{If} \quad \sum_{i=1}^{n-1} \mathbf{1}_{A_i}(s) + k\mathbf{1}_{A_n}(s) = \sum_{i=1}^{n-1} \mathbf{1}_{B_i}(s) + k\mathbf{1}_{B_n}(s) \quad \text{for all } s \in S, \\
&\text{and} \quad A_i \succsim B_i \quad \text{for } i = 1, \ldots, n-1, \\
&\text{then} \quad B_n \succsim A_n.
\end{aligned}$$

GFC strengthens Finite Cancellation. Similarly to Finite Cancellation, it is concerned with sequences of events with identical accumulation of indicators, which again should have the same overall likelihood weight. However, the conclusion in GFC applies to multiple repetitions of the last pair of events, A_n and B_n. As in Finite Cancellation, if the first $n-1$ events in one sequence are judged to weigh at least as much as the first $n-1$ events in the other sequence, then the last pair of events, repeated more than once in the aggregation, should balance the account. Thus, B_n is regarded at least as likely as A_n.

This generalization of Finite Cancellation is required since the relation in this paper is not assumed to be complete. For a complete relation, GFC is implied by Finite Cancellation: $\neg(B_n \succsim A_n)$ translates to $A_n \succ B_n$, and under the assumptions of the axiom, a contradiction to Finite Cancellation is inflicted. When the relation is incomplete, such a contradiction may not arise, thus GFC specifically requires that $B_n \succsim A_n$ be concluded.

As this work is concerned with incomplete relations, the role of GFC is two fold. The first is to preserve consistency of the 'at least as likely as' relation, as explained above. The second is to allow for extensions of the relation to yet undecided pairs of events, on the basis of others. Suppose that a pair of yet-undecided events can play the part of A_n and B_n for some sequences of events $(A_i)_{i=1}^n$ and $(B_i)_{i=1}^n$. GFC then explicitly prescribes a completion rule that imposes the principle of each state having its own marginal contribution, unrelated to other states.

Any relation $\succsim$ with a subjective multi-prior probability representation as in (1) satisfies GFC. In the opposite direction, Theorems 1 and 2, the main results of the paper, show that GFC, along with basic conditions (and when S is infinite an additional richness assumption), imply that the relation $\succsim$ admits a subjective multi-prior representation.

An analogue axiom to GFC, formulated on mappings from states to outcomes, appeared in Blume et al. (2009) (under the name 'extended statewise cancellation'). In this paper, the axiom was applied in a different framework, and was used to obtain a representation

that contains a *subjective* state space. Hence, the result of Blume et al. cannot be employed to characterize a likelihood relation over a given, primitive state space.

1.4 Related literature

Incomplete Expected Utility models over objective lotteries (on a set of prizes) began with Aumann (1962). Dubra, Maccheroni and Ok (2004) gave a full characterization of an incomplete preference relation over lotteries, which obtains a unanimity representation over a set of utility functions.

The path of dropping completeness for the sake of ambiguity, as suggested in this paper, was taken in the literature by several authors. Ok, Ortoleva and Riella (2008) axiomatized incomplete relations that admit either an expected multi-prior utility or an expected multi-utility representations. Galaabaatar and Karni (2011) presented a model of incomplete preferences, in which incompleteness stems both from ambiguity regarding a prior probability (which is the rationale advocated here) and from incompleteness of tastes. Their axiomatization characterizes an incomplete preference relation in an uncertain environment, that admits a multi-probability expected multi-utility representation. The special cases of expected multi-utility (w.r.t. a single probability measure), and of multi-probability expected utility representation (w.r.t. a single utility), were given specific axiomatizations. Other papers, which are the most related to the current work, are Bewley (2002), Ghirardato et al (2003), and Nehring (2009). These papers are discussed next.

Bewley (2002) axiomatized preference among alternatives that are mappings from events to a rich set of consequences. Bewley worked in an environment with exogenous probabilities, as in Anscombe and Aumann (1963; as rephrased by Fishburn 1970), and showed that dropping completeness and maintaining the rest of Anscombe-Aumann axioms yields a multi-prior expected utility representation. That is, for every pair of alternatives f and g,

$$f \succsim g \quad \Leftrightarrow \quad \mathbb{E}_\mu(u(f)) \geq \mathbb{E}_\mu(u(g)) \ \text{ for every } \mu \in \mathcal{P}, \tag{3}$$

for a unique convex and closed set of priors $\mathcal{P}$ and a vN-M utility function u ($\mathbb{E}_\mu$ denotes the expectation operator w.r.t. the probability measure μ). In other words, under a quite 'standard' set of axioms, which is known to imply an expected utility representation with a unique probability measure (as Anscombe and Aumann show), giving up the completeness

assumption leads to an expected utility representation with a set of probability measures.

Prior to Bewley, Giron and Rios (1980) phrased similar axioms over alternatives which map an abstract state space to the real line. Giron and Rios took the alternatives to be bounded, and assumed that they consist of a convex set (they mentioned randomization as justification for convexity). Giron and Rios showed that their set of axioms implied a multi-prior expected utility representation as in (3).

Both models, that by Giron and Rios and the one by Bewley, if applied with only two consequences, may be used to extract beliefs without any implication to risk attitude. However, they both assume convexity, which amounts to the fact that events may be mixed using exogenous, objective probabilities. Hence, their setup is less general than the purely subjective setup of Savage.

Ghirardato et al (2003) axiomatized multi-prior expected utility preferences in an environment without exogenous probabilities. Instead, they assumed that the set of consequences is connected, and that there exists an event on which all probabilities agree. The structure and axioms they employed do not permit applying the model for two consequences alone. In order to obtain the mixtures used in their axioms, it necessitates specifying preference on all binary alternatives contingent on the agreed-upon event (that is, on all alternatives that yield one consequence on this event and another outside of it). Thus, the model cannot be used to identify 'pure' belief.

Nehring (2009) presented a model based on an incomplete 'at least as likely as' relation over events. Nehring formulated assumptions on the relation, and proved a multi-prior representation result as in (1). In Nehring's work, there is a unique convex and closed set of priors that represents the relation. The model is placed in a general setup, where events belong to an algebra over a state space. No assumptions a-priori are made on this space, and in particular no objective probabilities are assumed. Nonetheless, one of the axioms Nehring applied to characterize the multi-prior representation is *Equidivisibility*, which hinges upon an explicit assumption that any event can be divided into two equally-likely events. Nehring's assumption is quite a strong one. It states that any event can be split into two events in a way that *all* the priors involved agree that they are equally probable. Formally, the set of priors in Nehring's representation theorem is explicitly assumed to have the property that, for any event A, there exists an event $B \subset A$ such that all priors μ agree that $\mu(B) = \mu(A)/2$. It implies, in particular, that all the priors agree on a rich algebra of events: the one generated by dividing the entire space into 2^n equally likely events, for any integer n. This assumption restricts the plurality of the representing set

of priors. Moreover, it does not work in the finite case.

The setup employed in this paper is as general as that of Savage[5]. Neither exogenous probabilities nor any manner of mixture is assumed or required. The assumptions apply to events alone.

1.5 Outline of the paper

Section 2 describes the essentials of the subjective multi-prior probability model. It details the setup and assumptions, and then formulates representation theorems, separately for the cases of a finite and an infinite spaces. Section 3 presents some examples, and Section 4 elaborates on three issues: a relative notion of 'more ambiguous than'; an embedding of ambiguity attitude that yields a complete relation; and minimal sets of priors. All proofs appear in the last section.

2 The subjective multi-prior probability model

2.1 Setup and assumptions

Let S be a nonempty set, Σ an algebra over S, and $\succsim$ a binary relation over Σ. A statement $A \succsim B$ is to be interpreted as 'A is at least as likely as B'. For an event $E \in \Sigma$, $\mathbf{1}_E$ denotes the indicator function of E. In any place where a partition over S is mentioned, it is to be understood that all atoms of the partition belong to Σ.

The following assumptions ('axioms') are employed to derive a subjective multi-prior probability belief representation.

P1. Reflexivity:
For all $A \in \Sigma$, $A \succsim A$.

P2. Positivity:
For all $A \in \Sigma$, $A \succsim \emptyset$.

P3. Non Triviality:
$\neg(\emptyset \succsim S)$.

[5]In fact the setup is even somewhat less restrictive than that of Savage, in that it requires Σ to be an algebra and not a σ-algebra.

These first three assumptions are standard. Positivity and Non-Triviality are two of de Finetti's suggested attributes. Since Completeness is not supposed, Reflexivity is added in order to identify the relation as a weak one. Transitivity is implied by the other axioms, hence it is not written explicitly.

The next assumption is the central one in the derivation of the main result of the paper.

P4. Generalized Finite Cancellation:
Let $(A_i)_{i=1}^n$ and $(B_i)_{i=1}^n$ be two sequences of events from Σ, and $k \in \mathbb{N}$ an integer. Then

$$\begin{aligned}
&\text{If} \quad \sum_{i=1}^{n-1} \mathbf{1}_{A_i}(s) + k\mathbf{1}_{A_n}(s) = \sum_{i=1}^{n-1} \mathbf{1}_{B_i}(s) + k\mathbf{1}_{B_n}(s) \quad \text{for all } s \in S, \\
&\text{and} \quad A_i \succsim B_i \quad \text{for } i = 1, \ldots, n-1 \ , \\
&\text{then} \quad B_n \succsim A_n \quad .
\end{aligned}$$

Generalized Finite Cancellation is based on the basic assumption that every state has the same marginal contribution of likelihood, no matter to which other states it is added. Generalized Finite Cancellation applies this underlying assumption to sequences of events that have exactly the same aggregation for each and every state, that is, in cases where each state appears the same number of times in each of the sequences. As each state should add the same likelihood weight to each of these aggregations, the axiom requires that likelihood judgements be balanced among the two sequences.

Basic Cancellation (de Finetti's condition) is obtained from Generalized Finite Cancellation by letting $A_1 = A$, $B_1 = B$, $A_2 = B \cup C$ and $B_2 = A \cup C$, for events A, B and C such that $(A \cup B) \cap C = \emptyset$. The indicators sum is identical,

$$\mathbf{1}_{A_1} + \mathbf{1}_{A_2} = \mathbf{1}_A + \mathbf{1}_{B \cup C} = \mathbf{1}_A + \mathbf{1}_B + \mathbf{1}_C = \mathbf{1}_{A \cup C} + \mathbf{1}_B = \mathbf{1}_{B_1} + \mathbf{1}_{B_2} \ ,$$

and hence $A \succsim B$ implies $A \cup C \succsim B \cup C$ and vice versa. Transitivity is obtained by letting $A_1 = B_3 = A$, $A_2 = B_1 = B$, and $A_3 = B_2 = C$. The following remark sums several implications of Generalized Finite Cancellation.

Remark 1. Generalized Finite Cancellation implies that $\succsim$ satisfies Transitivity, and together with Positivity yields that $A \succsim B$ whenever $A \supset B$. It also results in additivity: $A \succsim B \ \Leftrightarrow \ A \cup E \succsim B \cup E$, whenever $A \cap E = B \cap E = \emptyset$. In particular[6],

[6]To understand the following, note that $A = (A \cap B^c) \cup (A \cap B) \succsim (B \cap A^c) \cup (A \cap B) = B \quad \Leftrightarrow$

$A \succsim B \quad \Leftrightarrow \quad B^c \succsim A^c$, hence $S \succsim B$ for all events B.

2.2 Subjective multi-prior probability representation

2.2.1 The case of a finite space S

The next theorem states that when S is finite assumptions **P1-P4** are necessary and sufficient to obtain a subjective multi-prior probability representation of $\succsim$. For simplicity, Σ is assumed to be the collection of all subsets of S (that is, $\Sigma = 2^S$).

Theorem 1. *Suppose that S is finite, and let $\succsim$ be a binary relation over events in S. Then statements (i) and (ii) below are equivalent:*

(i) $\succsim$ satisfies axioms P1 through P4.

(ii) There exists a nonempty set $\mathcal{P}$ of additive probability measures over events in S, such that for every $A, B \subseteq S$,

$$A \succsim B \quad \Leftrightarrow \quad \mu(A) \geq \mu(B) \quad \textit{for every } \mu \in \mathcal{P}.$$

The set of prior probabilities need not be unique. However, the union of all representing sets is itself a representing set, and is maximal w.r.t. inclusion by its definition. When analyzing judgements made under ambiguity, the maximal set w.r.t. inclusion seems to be a natural choice to express belief, as it takes into consideration all priors that may be relevant to the case at hand. In fact, the maximal w.r.t. inclusion set includes all the probability measures that almost agree with the relation.

Suppose, for instance, that after identifying a subjective multi-prior probability, an individual is interested also in the extreme probabilities of each event. The maximal set yields the largest range of probabilities for each event, without casting unnecessary limits. On that account, the notion of a maximal w.r.t. inclusion set is claimed to be a satisfactory notion of unique belief.

When S is finite, then even if completeness is assumed, there need not be one unique probability measure that represents the relation. That is to say, even if $\succsim$ is complete, the maximal representing set is not necessarily a singleton (see Example 1 below).

$A \cap B^c \succsim B \cap A^c \quad \Leftrightarrow \quad B^c = (A \cap B^c) \cup (A^c \cap B^c) \succsim (B \cap A^c) \cup (A^c \cap B^c) = A^c.$

2.2.2 The case of an infinite space S

To obtain representation when S is infinite, an additional richness assumption is required. Without this assumption, for an infinite S, it is possible to obtain a set $\mathcal{P}$ of probability measures that only *almost agrees* with $\succsim$, in the sense that for all events A and B, $A \succsim B \;\Rightarrow\; \mu(A) \geq \mu(B)$ for every $\mu \in \mathcal{P}$, but not necessarily the other way around. To obtain the full 'if and only if' representation, a Non-Atomicity axiom is added. The assumption requires a definition of strong preference for its formulation (the definition originates in Nehring, 2009).

Definition 1. *For two events $A, B \in \Sigma$, the notation $A \succ\succ B$ states that there exists a finite partition $\{G_1, \ldots, G_r\}$ of S, such that $A \setminus G_i \succsim B \cup G_j$ for all i, j.*

In the representation (thus when all the axioms are assumed to hold), whenever the set $\mathcal{P}$ is the maximal w.r.t. inclusion set[7], having $A \succ\succ B$ is equivalent to the condition that there exists $\delta > 0$ for which $\mu(A) - \mu(B) > \delta > 0$ for every $\mu \in \mathcal{P}$.

P5. Non-Atomicity:
If $\neg(A \succsim B)$ then there exists a finite partition of A^c, $\{A'_1, \ldots, A'_m\}$, such that for all i, $A'_i \succ\succ \emptyset$ and $\neg(A \cup A'_i \succsim B)$.

Remark 2. As $\neg(A \succsim B) \;\Leftrightarrow\; \neg(B^c \succsim A^c)$, Non-Atomicity can equivalently be phrased as:
If $\neg(A \succsim B)$ then there exists a finite partition of B, $\{B_1, \ldots, B_m\}$, such that for all i, $B_i \succ\succ \emptyset$ and $\neg(A \succsim B \setminus B_i)$.

Non-Atomicity is the incomplete-relation version of Savage's richness assumption P6. Adding Completeness makes **P5** (along with the definition of strict preference) identical to Savage's P6, as negation of preference simply reduces to strict preference in the other direction. The setup used here is somewhat weaker than that in Savage, as Σ is assumed to be an algebra and not necessarily a σ-algebra. Still, adding Completeness yields a unique probability that represents the 'at least as likely as' relation $\succsim$ (see Kopylov (2007) for this result for an even more general structure of Σ)[8]. The difference from Savage's theorem is that the derived probability need not be convex-ranged, only *locally dense*:

[7] or any other compact set

[8] Generalized Finite Cancellation obviously implies additivity (de Finetti's Cancellation), thus when adding completeness to axioms P2 through P5, Savage's result follows.

Definition 2. A probability measure μ over Σ is *locally dense* if, for every event B, the set $\{\mu(A) \mid A \subset B\}$ is dense in $[0, \mu(B)]$.

In a similar fashion, Non Atomicity implies that each probability measure in the subjective multi-prior probability set is locally dense.

Definition 3. A set $\mathcal{P}$ of probability measures is said to be *uniformly absolutely continuous* if:

(a) For any event B, $\mu(B) > 0 \Leftrightarrow \mu'(B) > 0$ for every pair of probabilities $\mu, \mu' \in \mathcal{P}$.

(b) For every $\varepsilon > 0$, there exists a finite partition $\{G_1, \ldots, G_r\}$ of S, such that for all j, $\mu(G_j) < \varepsilon$ for all $\mu \in \mathcal{P}$.

The following is a representation result of the subjective multi-prior probability model for the case of an infinite set S.

Theorem 2. *Let $\succsim$ be a binary relation over Σ. Then statements (i) and (ii) below are equivalent:*

*(i) $\succsim$ satisfies axioms **P1-P5**.*

(ii) There exists a nonempty, uniformly absolutely continuous set $\mathcal{P}$ of additive probability measures over Σ, such that for every $A, B \in \Sigma$,

$$A \succsim B \quad \Leftrightarrow \quad \mu(A) \geq \mu(B) \text{ for every } \mu \in \mathcal{P}. \tag{4}$$

Corollary 1. *Assume that $\succsim$ admits a subjective multi-prior probability representation as in (ii) of Theorem 2. Then all the probability measures in the representing set are locally dense.*

It is a known fact that existence of fine partitions, as depicted in the definition of uniform absolute continuity, imply local denseness (see, for instance, Lehrer and Schmeidler, 2005). Hence, if $\succsim$ admits a subjective multi-prior probability representation as in (ii) of Theorem 2, then all the probability measures in the representing set are locally dense.

Observation 1. The union of all sets of probabilities that represent $\succsim$ as in (ii) of Theorem 2 is itself a representing set of $\succsim$. Thus, the union of all sets that represent $\succsim$ satisfies (ii) of Theorem 2. By its definition, every other representing set is contained in it, namely it is the maximal w.r.t. inclusion representing set of $\succsim$.

Note that for general sets of probability measures, the fact that all sets in the union are uniformly absolutely continuous does not imply that the union itself is uniformly absolutely continuous (as the union may be over an infinite number of sets). However, if the sets represent $\succsim$, then (i) of Theorem 2 guarantees uniform absolute continuity of the union. As already explained in the introduction, the set of probabilities which represents $\succsim$ need not be unique. Nevertheless, the maximal representing set, which is the union of all representing sets, is unique by its definition.

A few words on the reasons for lack of uniqueness, even when Non-Atomicity is assumed, are in order. The difficulty in obtaining uniqueness is that it is impossible, under the assumptions made, to produce an 'objective measuring rod' for probabilities. The probabilities in the priors set need not agree on any event which probability is strictly between zero and one. Moreover, there even need not be events with probability that is ε-close to some value $0 < p < 1$, according to all measures (see Example 2 in section 3). In addition, the technique used in the proof of the main theorem translates events to indicator vectors. That is, an event E is represented by the vector in $\mathbb{R}^S$ which assigns one to any state s included in E and zero otherwise. Under this translation, the preference relation considered in this paper applies only to indicator vectors, and preference calls among vectors with values other than zero and one are meaningless. In that sense, the domain to which the relation applies is too scarce to obtain uniqueness. The problem is not only technical, for there are examples with two distinct (convex and closed) representing sets of priors. An example may be found in Nehring (2009; Example 1).

3 Examples

The following examples demonstrate some features of the model. In the first example it is shown that in the case of a finite state space, even if assumptions P1 through P4 are supplemented with completeness, there need not be one unique probability measure representing the relation. The second example is a simple example of an incomplete relation over an infinite state space, satisfying assumptions P1 through P5. The third and last example, which again employs an infinite state space, demonstrates the problem of producing an 'objective measuring rod'. It shows that one cannot hope to be able to produce events which are unambiguous (other than null or universal events), in the sense that all probabilities in the representing set assign those events the same probability.

Example 1. Let $S = \{H, T\}$, $\Sigma = 2^S$. If it is the case that $H \succsim T$, then any probability measure of the type $(H : p\ ;\ T : 1-p)$, for $0.5 \leq p \leq 1$, represents the relation. In that

case, even though the relation is complete, there is no unique representing probability measure, or unique representing probability measures set.

In the following examples, λ denotes the Lebesgue measure.

Example 2. Let $S = [0, 1)$ and denote by Σ the algebra generated by all intervals $[a, b)$ contained in $[0, 1)$. Define measures π_1 and π_2 through their densities:

$$f_1(s) \;=\; \begin{cases} \frac{1}{2} & s \in [0, 0.5) \\ \frac{3}{2} & s \in [0.5, 1) \end{cases} \quad , \quad f_2(s) \;=\; \begin{cases} \frac{3}{2} & s \in [0, 0.5) \\ \frac{1}{2} & s \in [0.5, 1) \end{cases}$$

In other words, π_1 distributes uniform weight of $\frac{1}{4}$ on $[0, 0.5)$ and uniform weight of $\frac{3}{4}$ on $[0.5, 1)$, and π_2 distributes uniform weight of $\frac{3}{4}$ on $[0, 0.5)$ and uniform weight of $\frac{1}{4}$ on $[0.5, 1)$. Let $\mathcal{P}$ denote the convex set generated by π_1 and π_2. A set $\mathcal{P}$ of this form may express an individual's ambiguity as to the way the probabilistic weight is divided between the lower and upper halves of a considered range, with the constraint that at least $\frac{1}{4}$ of the weight is placed on each half, and with otherwise uniform belief.

For every pair of measures $\pi, \pi' \in \mathcal{P}$, $\pi(A) = 0 \;\Leftrightarrow\; \pi'(A) = 0$, and for $\varepsilon > 0$, $\lambda(E) < \frac{2}{3}\varepsilon$ guarantees $\pi_i(E) < \varepsilon$ for $i = 1, 2$, therefore for all measures in $\mathcal{P}$. Thus, fine partitions as required exist, and $\mathcal{P}$ is uniformly absolutely continuous. The subjective multi-prior probability induced by $\mathcal{P}$ is:
$A \succsim B$, if and only if,

$$\begin{aligned} \frac{1}{2}\,[\,\lambda(A) - \lambda(B)\,] \quad &\geq \quad \lambda(B \cap [0.5, 1)) - \lambda(A \cap [0.5, 1)) \\ &\text{and} \\ \frac{1}{2}\,[\,\lambda(A) - \lambda(B)\,] \quad &\geq \quad \lambda(B \cap [0, 0.5)) - \lambda(A \cap [0, 0.5))\ . \end{aligned}$$

In this example, a necessary condition to obtain $A \succsim B$ is $\lambda(A) \geq \lambda(B)$, but this condition is not sufficient. It should also be that on each half separately, event B does not have a 'large-enough edge'. For instance, $\neg\,(\,[0, 0.5) \succsim [0.8, 1)\,)$, but $[0, 0.45) \cup [0.95, 1) \succsim [0.8, 1)$. Events A and B are considered equally likely if they have the same Lebesgue measure, divided in the same manner between $[0, 0.5)$ and $[0.5, 1)$.

Example 3. Let $S = [0, 1)$ and Σ the algebra generated by all intervals $[a, b)$ contained in $[0, 1)$. For $A \in \Sigma$ such that $\lambda(A) = \frac{1}{2}$, let π_A to be the probability measure defined by

the density:

$$f_A(s) = \begin{cases} \frac{1}{2} & s \in A \\ \frac{3}{2} & s \notin A \end{cases}$$

Let $\mathcal{P}$ be the convex and closed set generated by all probability measures π_A. All measures in the set are mutually absolutely continuous with λ. For $\varepsilon > 0$, letting $\{E_1, \ldots, E_n\}$ be a partition with $\lambda(E_i) < \frac{2}{3}\varepsilon$ obtains $\pi_A(E_i) < \varepsilon$ for all measures π_A, hence for all measures in $\mathcal{P}$. The set $\mathcal{P}$ is therefore uniformly absolutely continuous.

The resulting subjective multi-prior probability representation satisfies assumptions **P1-P5**. Note that for any event $B \in \Sigma$ for which $0 < \lambda(B) < 1$,

$$\begin{aligned} \max_{\pi \in \mathcal{P}} \pi(B) &= \frac{3}{2}\min(\lambda(B), \frac{1}{2}) + \frac{1}{2}\max(0, \lambda(B) - \frac{1}{2}) \\ &> \frac{1}{2}\min(\lambda(B), \frac{1}{2}) + \frac{3}{2}\max(0, \lambda(B) - \frac{1}{2}) = \min_{\pi \in \mathcal{P}} \pi(B). \end{aligned}$$

That is, the measures in $\mathcal{P}$ do not agree on any event which is non-null and non-universal. Moreover, there are even no events with probability ε-close to a fixed value $0 < p < 1$. Note also that for events A and B, it cannot be that all measures in $\mathcal{P}$ agree that A and B have equal probabilities, therefore events cannot be partitioned into equally likely events.

4 Extensions and comments

4.1 Link to choice behavior

The model presented in this paper takes as a primitive a binary relation over events, which is interpreted as an 'at least as likely as' relation. The same 'at least as likely as' judgements could be derived from preferences over bets, in the following manner: if a bet on A is preferred to a bet on B (when both bets involve the same prizes), then it is concluded that the individual judges A to be at least as likely as B. This is the way Savage (1954) defined the likelihood relation in his Subjective Expected Utility model (See also de Finetti (1931,1937) and Ramsey (1931)). If the primitive notion considered is a binary relation over bets (or, more generally, a binary relation over mappings from events to a set of consequences) then some consistency assumption should be imposed, in order to maintain that the derived likelihood relation is independent of the choice of prizes defining the bets (one such well-known consistency assumption is Savage's (1954) axiom P4).

Formally, let X be a nonempty set, denoting the set of *consequences*, let S be the set of states endowed with an algebra Σ, and denote by $\mathcal{F}$ the set of (simple) *acts*, which are all the Σ-measurable and finite-valued functions from S to X. Let $\succsim$ denote a binary relation over $\mathcal{F}$. An event $A \in \Sigma$ is said to be *null* if $xAf \sim f$ for all $x \in X$ and $f \in \mathcal{F}$. Otherwise the event is *non-null*. For consequences $x, y \in X$ we write $x \succsim y$ whenever $\bar{x} \succsim \bar{y}$, where $\bar{x}$ denotes the constant act yielding x in all states, and similarly for $\bar{y}$. For consequences $x, y \in X$ denote $x \succ y$ if $x \succsim y$ but $\neg(y \succsim x)$. For simplicity, assume that the relation over consequences is complete. The assumption implies that incomparability of acts is not a result of incompleteness of tastes, but rather of uncertainty about events.

Completeness of Tastes:
For all consequences $x, y \in X$, either $x \succsim y$ or $y \succsim x$.

An 'at least as likely as' relation $\succsim^l$ over Σ may be derived from $\succsim$, by:

$$\text{For events } A, B \in \Sigma,\ A \succsim^l B \text{ whenever } x^*Ax \succsim x^*Bx \text{ for consequences } x^* \succ x\ . \quad (5)$$

Suppose that the 'at least as likely as' relation $\succsim^l$ satisfies axioms P1-P4 in case S is finite, and P1-P5 in case S is infinite. According to theorems 1 and 2, there exists a nonempty set $\mathcal{P}$ of additive probability measures over Σ such that for every $A, B \in \Sigma$, $A \succsim B$ if and only if $\mu(A) \geq \mu(B)$, for every $\mu \in \mathcal{P}$.

The next step is to define when the preferences of an individual are governed by a set of prior probabilities. In order to do that, some notation and definitions are introduced. denote by $L = [\ x_1, p_1\ ;\ \ldots\ ;\ x_n, p_n\ ]$ the lottery which obtains consequence x_i with probability p_i, and by $\mathcal{L}(X)$ the set of all lotteries over consequences, that is, the set of all probability distributions over X with finite support. Every act $f \in \mathcal{F}$ and additive probability measure μ over Σ induce a lottery $L(f, \mu)$ in $\mathcal{L}(X)$, by
$L(f, \mu) = [\ x_1, \mu(f^{-1}(x_1))\ ;\ \ldots\ ;\ x_n, \mu(f^{-1}(x_n))\ ]$, where $x_1, \ldots, x_n$ are the consequences obtained under act f. For lotteries $L_1, L_2 \in \mathcal{L}(X)$, L_1 *first order stochastically dominates* L_2 with respect to $\succsim$ if $L_1(\{y \in X \mid y \succsim x\}) \geq L_2(\{y \in X \mid y \succsim x\})$ for every $x \in X$. That is, L_1 first order stochastically dominates L_2 if, for every x, the probability of receiving a consequence x or better (when 'better' is determined according to $\succsim$) is higher under L_1 than under L_2. A definition of *multi-prior probabilistic sophistication* can now be stated, akin to probabilistic sophistication of Machina and Schmeidler (1992).

Definition 4. *A binary relation over acts $\succsim$ is said to be* ***multi-prior probabilistically***

sophisticated *if there exists a nonempty set $\mathcal{P}$ of probability measures over Σ such that the derived 'at least as likely as' relation $\succsim^l$ admits a subjective multi-prior representation with the set $\mathcal{P}$, and for two acts f and g, $f \succsim g$ whenever $L(f,\mu)$ first order stochastically dominates $L(g,\mu)$ w.r.t. $\succsim$, for every $\mu \in \mathcal{P}$.*

Probabilistic Sophistication of Machina and Schmeidler (1992) maintains that the individual holds a subjective belief in the form of an additive probability measure, and employs it to translate each act to the lottery it induces over consequences. Preference between acts is then dictated by preference between lotteries, where preference between lotteries is required to abide by the rule of first order stochastic dominance.[9] Multi-prior probabilistic sophistication generalizes this idea by allowing for a multi-prior belief instead of a single prior one. Consequently, each act induces a multitude of lotteries over consequences, each through a different prior in the set. Preference between acts is dictated by preference between these probability-lottery mappings, where preference between mappings is required to be monotonic with respect to pointwise (i.e., per-probability) first order stochastic dominance.

Note that the notion of stochastic dominance used links every lottery to the probability measure from which it originates. That it to say, f is preferred to g whenever its induced lottery under the probability measure μ first order stochastically dominates the lottery induced by g *under the same probability* μ, for every μ. By contrast, it is not required that every lottery induced by f (under the entire priors set) first order stochastically dominates every lottery induced by g. Rather, each prior in the set is viewed as a conceivable belief over the states of nature, and accordingly the lotteries induced by f and g under each conceivable belief *separately* should exhibit first order stochastic dominance.

Next, multi-prior probabilistic sophistication is characterized axiomatically. The main axiom in the characterization is *Cumulative Dominance*, formulated by Sarin and Wakker (1992 and 2000). The axiom is an intuitive preference-version of first order stochastic dominance, and asserts that an act f is preferred to an act g whenever, for every consequence x, the individual believes that it is more likely to obtain a consequence x or better under f than under g. For an act f and a consequence x, denote by $\{f \succsim x\}$ the set $\{s \in S | f(s) \succsim x\}$, which is the set of all states on which f obtains a consequence x or better.

[9] Machina and Schmeidler (1992) pose an additional continuity requirement, which is omit here.

Cumulative Dominance:
For all acts $f, g \in \mathcal{F}$, if $\{f \succsim x\} \succsim^l \{g \succsim x\}$ for every $x \in X$, then $f \succsim g$.

Cumulative Dominance implies that the 'at least as likely as' relation $\succsim^l$ is consistent, in the sense that the likelihood judgements it entails are independent of the choice of prizes involved in the bets. To see this, suppose that for events $A, B \in \Sigma$ and consequences $x^* \succ x$, $x^*Ax \succsim x^*Bx$. Let y^* and y be consequences such that $y^* \succ y$, and consider the bets y^*Ay and y^*By. According to the definition of $\succsim^l$, the relationship between the bets x^*Ax and x^*Bx implies that $A \succsim^l B$, thus for every consequence z, $\{y^*Ay \succsim z\} \succsim^l \{y^*By \succsim z\}$, and by Cumulative Dominance $y^*Ay \succsim y^*By$.

Sarin and Wakker (2000) show that probabilistic sophistication *á*-la Machina and Schmeidler (1992) holds if and only if there is a probability measure agreeing with the 'at least as likely as' relation $\succsim^l$, and Cumulative Dominance is satisfied. Similarly, Cumulative Dominance together with a set of prior probabilities that represents the 'at least as likely as' relation $\succsim^l$ delivers multi-prior probabilistic sophistication, as stated in the next theorem.

Theorem 3. *Let $\succsim$ be a binary relation over $\mathcal{F}$. Then statements (i) and (ii) below are equivalent:*

(i) $\succsim^l$, the derived binary relation over Σ defined according to (5), satisfies axioms P1-P4, and also P5 if S is infinite. $\succsim$ and $\succsim^l$ satisfy Cumulative Dominance.

(ii) $\succsim$ is multi-prior probabilistically sophisticated. If S is infinite, then the corresponding probabilities set is uniformly absolutely continuous.

4.2 A relative notion of 'more ambiguous than'

Next, a relative notion of 'more ambiguous than' is presented. One subjective multi-prior probability is considered to be more ambiguous than another, if it is less decisive, and leaves more comparisons open. As can be expected, the 'more ambiguous than' relation is characterized by inclusion of the $\succsim$-maximal representing sets of priors. Being more ambiguous means having a larger, in the sense of set inclusion, set of priors.

Proposition 1. *Let $\succsim_1$ and $\succsim_2$ be two binary relations over Σ, satisfying P1-P5, with $\mathcal{P}_1$ and $\mathcal{P}_2$ the maximal representing sets (unions of all representing sets) of $\succsim_1$ and $\succsim_2$, respectively. Then the following statements are equivalent:*

(i) $\succsim_1 \subseteq \succsim_2$.

(ii) $\mathcal{P}_2 \subseteq \mathcal{P}_1$.

A result similar to this one is presented in Ghirardato, Maccheroni and Marinacci (2004). The setup in Ghirardato et al. contains alternatives which are mappings from a state space to a set of consequences. In that paper, a definition of 'reveals more ambiguity than' is given, that amounts to inclusion of one relation in another, where both relations admit a multiple-priors expected utility representation. Ghirardato et al. show that such inclusion of the relation implies reversed inclusion of the priors sets.[10]

4.3 Ambiguity attitude

A subjective multi-prior probability may be viewed as representing the unambiguous judgements of an individual. According to this view, some comparisons between events are left open due to ambiguity, and those that are determined are the 'unambiguous part' of the individual's likelihood judgements.

In the representation, a set of prior probabilities represents the scenarios the individual considers possible. An event is judged at least as likely as another if this is the case under every possible scenario. The set of priors thus represents all the ambiguity that the individual perceives in the situation, and the unanimity rule characterizes those comparisons that are not affected by this ambiguity.

After ambiguity itself is identified, attitudes towards it may be added. One obvious possibility is completion using one probability measure from the multi-prior probability set. Another possibility, in a setup with an infinite space S, is suggested next. It exhibits aversion to ambiguity, and describes an individual who judges events by their lowest possible probability. A similar treatment, of completing an incomplete relation, is suggested by Gilboa et al. (2010) and Ghirardato, Maccheroni and Marinacci (2004). In these models, a set of priors is obtained as a description of the ambiguity perceived by an individual. The relations discussed in those papers apply to alternatives which map states to lotteries over an abstract set of consequences, and the representations obtained involve a procedure of expected utility. Here, the relation applies to events alone, but the treatment of ambiguity separately from ambiguity attitude is similar in spirit.

In order to obtain an ambiguity averse completion, the framework of Sarin and Wakker (1992) is adopted. In that framework, an exogenous sub-sigma-algebra is taken as part of

[10] In order to obtain the other direction in Ghirardato et al., an additional assumption of identical utility for both relations is required.

the model primitives, and is meant to represent a (rich) set of unambiguous events. An assumption of *Ambiguity Aversion* is further used. The aversion to ambiguity is expressed by supposing that the individual ranks unambiguous events strictly above other events, whenever this does not contradict consistency with the incomplete, 'original' relation.

Let $\mathcal{B}$ denote a sigma-algebra contained in Σ, and $\succsim$ a binary relation over Σ. Events in $\mathcal{A}$ should be thought of as *unambiguous*. For a complete relation $\succsim'$, the relations $\succ'$ and $\sim'$ denote its asymmetric and symmetric components, respectively.

First, a weak order assumption is required for the relation $\succsim'$.

P1'. Weak Order:
For any two events A and B, either $A \succsim' B$ or $B \succsim' A$. (Completeness)
For any three events A, B and C, if $A \succsim' B$ and $B \succsim' C$ then $A \succsim' C$. (Transitivity)

Next, a richness assumption, Savage's postulate P6 confined to unambiguous events, is stated.

P6-UA. Unambiguous Non-Atomicity.:
If $D \succ' C$ for events $C, D \in \Sigma$, then there exists a finite partition of S, $\{B_1, \ldots, B_m\}$ with $B_i \in \mathcal{B}$ for all i, such that $D \succ' C \cup B_i$ for all i.

Last, two linking assumptions are added. The first assumption, *Consistency*, states that $\succsim'$ is a completion of $\succsim$, so $\succsim'$ is consistent with the preferences exhibited by $\succsim$. The second, *Ambiguity Aversion*, subscribes a completion rule that favors events with known probabilities. This assumption imitates Epstein's **?** definition of uncertainty aversion for alternatives, mapping events to an abstract set of consequences. It uses as a benchmark for uncertainty neutrality the relation $\succsim$, and suggests to complete it by preferring unambiguous events. The completion in fact identifies, for every event, an *'unambiguous equivalent'* event, much in the same manner as is done with certainty equivalents. The 'unambiguous equivalent' of an event gives what is known as its 'matching probability' (see, for instance, Wakker (2010; Equation 4.9.1)). As noted above, analogous development, applied to alternatives mapping events to consequences, may be found in Gilboa et al. (2010; the relevant axiom is termed there *Default to Certainty*, or *Caution* in its weaker version).

P7. Consistency:
For any two events A and B, $A \succsim' B$ whenever $A \succsim B$.

P8. Ambiguity Aversion:
Let $E \in \Sigma$ and $B \in \mathcal{B}$ be events. If $\neg(E \succsim B)$ then $B \succ' E$.

Proposition 2. *Let $\succsim$ and $\succsim'$ be binary relations over Σ, and let $\mathcal{B}$ be a σ-algebra contained in Σ. Then statements (i) and (ii) below are equivalent.*

(i) The following conditions are satisfied:

(a) $\succsim$ satisfies axioms P1 through P5. On $\mathcal{B}$, $\succsim$ is complete.

(b) $\succsim'$ satisfies P6-UA.

(c) $\succsim$ and $\succsim'$ satisfy Consistency (P7) and Ambiguity Aversion (P8).

(ii) There exists a nonempty, uniformly absolutely continuous set of probabilities $\mathcal{P}$, such that for any pair of events A and B,

$$A \succsim B \quad \Leftrightarrow \quad \mu(A) \geq \mu(B) \ \text{ for every } \mu \in \mathcal{P} \tag{6}$$

$$A \succsim' B \quad \Leftrightarrow \quad \min_{\mu \in \mathcal{P}} \mu(A) \geq \min_{\mu \in \mathcal{P}} \mu(B) \tag{7}$$

Furthermore, the probabilities in $\mathcal{P}$ agree on $\mathcal{B}$, and their common part on $\mathcal{B}$ is unique and convex-ranged there.

The proposition suggests how to complete a subjective multi-prior probability, in a manner that expresses ambiguity aversion. An ambiguity seeking rule can analogously be described through the use of maximum instead of minimum in (7). As in Theorem 2, the set of probability measures $\mathcal{P}$ is not necessarily unique, yet a maximal (w.r.t. inclusion) set may be identified, which represents the relation (see Observation 1 above).

4.4 Minimal set of priors

Nehring (2009) presents an example with two sets of priors $\mathcal{P}_1 \subsetneqq \mathcal{P}_2$ that represent the same likelihood order. The priors in $\mathcal{P}_2$ that are not in $\mathcal{P}_1$ do not restrict the order induced by $\mathcal{P}_1$ in the sense that when all $p \in \mathcal{P}_1$ agree that $p(A) \geq p(B)$, every $q \in \mathcal{P}_2 \setminus \mathcal{P}_1$ concurs.

In this example the intersection of the two representing sets is also a representing set: it is $\mathcal{P}_1$. The question arises whether the intersection of every two representing sets is always a representing one.

The following example answers this question in the negative.

Example 4. Suppose that $S = \{0,1\}^{\mathbb{N}}$, the set of all sequences consisting of 0's and 1's. The algebra Σ is the one generated by events of the type $C(i,a) = (x_1, x_2, ...) \in S;\ x_1 = a$, where $a = 0, 1$. The probability measure, denoted $(p_1, p_2, ...)$, is the one induced by an infinite sequence of independent tosses of coins, where the probability of getting 1 in the i-th toss is p_i ($p_i \in (0,1)$) and the probability of getting 0 is $1 - p_i$. For simplicity a coin that assigns probability p to the outcome 1 and probability $1 - p$ to the outcome 0 is referred to as p-coin.

We define two disjoint sets of priors. Let $\mathcal{P}_1$ be the convex hull of the probability measures of the type $(p_1, ..., p_n, \frac{1}{2}, \frac{1}{2}, ...)$, where $p_1, ..., p_n \in \{\frac{1}{3}, \frac{1}{2}\}$. In other words, any distribution in $\mathcal{P}_1$ is such that the coins from time $n+1$ and onwards are $\frac{1}{2}$-coins, while up to time n the coins could be any combination of $\frac{1}{2}$-coins and $\frac{1}{3}$-coins.

The set $\mathcal{P}_2$ is defined in a way similar to that of $\mathcal{P}_1$ with the difference that coins from time $n+1$ and onwards are $\frac{1}{3}$-coins. Note that the distributions in $\mathcal{P}_1$ are generated by only finitely many $\frac{1}{3}$-coins and infinitely many $\frac{1}{2}$-coins. On the other hand, the distribution in $\mathcal{P}_2$ are generated by only finitely many $\frac{1}{2}$-coins and infinitely many $\frac{1}{3}$-coins. Thus, $\mathcal{P}_1$ and $\mathcal{P}_2$ are disjoint.

We argue that $\mathcal{P}_1$ and $\mathcal{P}_2$ induce the same likelihood relation over Σ. Let A, B be two events in Σ. There is time, say n, such that A and B are defined by conditions imposed only on the first n coordinates. Thus, the probability of A and B assigned by a particular distribution depends only on its behavior on the n first coordinates. However, for every distribution $\mu \in \mathcal{P}_1$ there is $\nu \in \mathcal{P}_2$ that behaves on the first n coordinates like μ. In particular, if $\mu(A) \geq \mu(B)$ (or $\mu(A) > \mu(B)$), then $\nu(A) \geq \nu(B)$ (or $\nu(A) > \nu(B)$). This implies that if $\nu(A) \geq \nu(B)$ for every $\nu \in \mathcal{P}_2$, then $\mu(A) \geq \mu(B)$ for every $\mu \in \mathcal{P}_1$. For a similar reason the inverse of the previous statement is also true. We conclude that $\mathcal{P}_1$ and $\mathcal{P}_2$ induce the same likelihood relation over Σ.

This example shows that two representing sets can be disjoint. In particular, the intersection of two representing sets is not representing the same relation.

A careful look at this example reveals that the sets $\mathcal{P}_1$ and $\mathcal{P}_2$ are not closed (w.r.t. the weak* topology). For instance, the probability $(\frac{1}{3}, \frac{1}{3}, ...)$ is in the closure of $\mathcal{P}_1$. The reason

is that $(\frac{1}{3}, \frac{1}{3}, ...)$ is a cumulative point of the following sequence of probability measures in $\mathcal{P}_1$: $(\frac{1}{3}, \frac{1}{2}, \frac{1}{2}, \frac{1}{2}...)$, $(\frac{1}{3}, \frac{1}{3}, \frac{1}{2}, \frac{1}{2}, \frac{1}{2}...)$, $(\frac{1}{3}, \frac{1}{3}, \frac{1}{3}, \frac{1}{2}, \frac{1}{2}, \frac{1}{2}...)$,... .

It turns out that the closure of one set contains the other and therefore, the closures of $\mathcal{P}_1$ and of $\mathcal{P}_2$ coincide. This conclusion leaves us with two unanswered questions: whether the intersection of two closed (w.r.t. the weak* topology) representing sets of priors can be disjoint, and whether the intersection itself is a representing set for the same relation.

5 Proofs

Let $B_0(S, \Sigma)$ denote the vector space generated by linear combinations of indicator functions $\mathbf{1}_A$ for $A \in \Sigma$, endowed with the supremum norm. Define a subset $D_{\succsim}$ of $B_0(S, \Sigma)$,

$$D_{\succsim} = closure\{ \sum_{i=1}^{n} \alpha_i \ [\ \mathbf{1}_{A_i} - \mathbf{1}_{B_i} \ ] \mid A_i \succsim B_i, \ \alpha_i \geq 0, \ n \in \mathbb{N} \ \} \ .$$

That is, $D_{\succsim}$ is the closed convex cone generated by indicator differences $\mathbf{1}_A - \mathbf{1}_B$, for $A \succsim B$. By Reflexivity, Positivity and Nontriviality, $D_{\succsim}$ has vertex at zero, $D_{\succsim}$ is not the entire space $B_0(S, \Sigma)$, and it contains every nonnegative vector $\psi \in B_0(S, \Sigma)$.

The next claims show that under assumptions P1 through P4, preference is preserved under convex combinations (the proof is joint for the finite and infinite cases). In some of the claims, conclusions are more easily understood considering the following alternative formulation of Generalized Finite Cancellation:

Let A and B be two events, and $(A_i)_{i=1}^n$ and $(B_i)_{i=1}^n$ two sequences of events from Σ, that satisfy,

$$\begin{aligned} &A_i \succsim B_i \ \text{ for all } i, \text{ and for some } \ k \in \mathbb{N}, \\ &\sum_{i=1}^{n} [\mathbf{1}_{A_i}(s) - \mathbf{1}_{B_i}(s)] = k[\mathbf{1}_A(s) - \mathbf{1}_B(s)] \ \text{ for all } s \in S. \end{aligned}$$

Then $A \succsim B$.

Claim 1. *Suppose that*[11] $r_i \in \mathbb{Q}_{++}$, *and* $A_i \succsim B_i$ *for* $i = 1, \ldots, n$. *If* $\mathbf{1}_A - \mathbf{1}_B = \sum_{i=1}^{n} r_i(\mathbf{1}_{A_i} - \mathbf{1}_{B_i})$, *then* $A \succsim B$.

[11] $\mathbb{Q}_{++}$ is the set of strictly positive rational numbers.

Proof. Let k denote the common denominator of $r_1, \ldots, r_n$, and write $r_i = \frac{m_i}{k}$ for $m_i \in \mathbb{N}$ and $i = 1, \ldots, n$. It follows that $k(\mathbf{1}_A - \mathbf{1}_B) = \sum_{i=1}^n m_i(\mathbf{1}_{A_i} - \mathbf{1}_{B_i})$ for all $s \in S$. By GFC applied to sequences $(A_i)_{i=1}^N$ and $(B_i)_{i=1}^N$, where each A_i and each B_i repeats m_i times ($N = m_1 + \ldots + m_n$), it follows that $A \succsim B$. ■

Claim 2. *Suppose there are* $\alpha_i > 0, i = 1, \ldots, n,$ *such that* $\mathbf{1}_A - \mathbf{1}_B = \sum_{i=1}^n \alpha_i(\mathbf{1}_{A_i} - \mathbf{1}_{B_i})$. *Then there are* $r_i \in \mathbb{Q}_{++}, i = 1, \ldots, n,$ *such that* $\mathbf{1}_A - \mathbf{1}_B = \sum_{i=1}^n r_i(\mathbf{1}_{A_i} - \mathbf{1}_{B_i})$.

Proof. Consider the partition induced by $A_1, \ldots, A_n, B_1, \ldots, B_n, A, B$, and denote it by $\mathcal{A}$, with atoms denoted by a. The assumed indicators identity for all $s \in S$ translates to the following finite system of linear equations, with the variables $\alpha_1, \ldots, \alpha_n$.

$$\sum_{i=1}^n \delta_i(a)\alpha_i = \delta(a), \quad a \in \mathcal{A},$$
$$\delta_i(a) = \mathbf{1}_{A_i}(s) - \mathbf{1}_{B_i}(s), \ s \in a, \quad \delta(a) = \mathbf{1}_A(s) - \mathbf{1}_B(s), \ s \in a.$$

Since all coefficients in the above equations, $\delta(a)$ and $\delta_i(a)$, are integers, by denseness of the rational numbers in the reals it follows that if a nonnegative solution to this system exists, then there also exists a nonnegative *rational* solution. Thus, if indeed $\mathbf{1}_A - \mathbf{1}_B = \sum_{i=1}^n \alpha_i(\mathbf{1}_{A_i} - \mathbf{1}_{B_i})$ for some $\alpha_i > 0$, then there is a solution $r_i \in \mathbb{Q}_{++}$. ■

Conclusion 1. *Suppose that* $A_i \succsim B_i$ *for* $i = 1, \ldots, n$. *If there are* $\alpha_i > 0$, $i = 1, \ldots, n$, *such that* $\mathbf{1}_A - \mathbf{1}_B = \sum_{i=1}^n \alpha_i(\mathbf{1}_{A_i} - \mathbf{1}_{B_i})$, *then* $A \succsim B$.

In order to show that $D_{\succsim}$ contains exactly those indicator differences which correspond to preference, it should further be proved that preference is preserved under the closure operation. This is done separately for the finite and infinite cases.

5.1 Proof of Theorem 1

First it is proved that under assumptions **P1-P4**, the multiple-priors representation follows (direction (i)⇒(ii)).

When S is finite, there are finitely many pairs of events. Therefore $D_{\succsim}$ is generated by a finite number of vectors. In fact,

$$D_{\succsim} = \sum_{(A,B):A \succsim B} \alpha_{(A,B)}(\mathbf{1}_A - \mathbf{1}_B), \quad \alpha_{(A,B)} \geq 0 \ ,$$

hence $D_{\succsim}$ is closed, and by Conclusion 1, $\mathbf{1}_A - \mathbf{1}_B \in D_{\succsim}$ if and only if $A \succsim B$. Define[12] $\mathcal{V} = \{v \in \mathbb{R}^S \mid v \cdot \varphi \geq 0 \text{ for all } \varphi \in D_{\succsim}\}$. The set $\mathcal{V}$ is a closed convex cone, and contains the zero function.

Claim 3.

$$\varphi \in D_{\succsim} \quad \Leftrightarrow \quad v \cdot \varphi \geq 0 \ \ \textit{for every } v \in \mathcal{V}.$$

Proof. By definition of $\mathcal{V}$, if $\varphi \in D_{\succsim}$ then $v \cdot \varphi \geq 0$ for every $v \in \mathcal{V}$. Now suppose that $\psi \notin D_{\succsim}$. Since $D_{\succsim}$ is a closed convex cone then by a standard separation theorem (see Dunford and Schwartz, Corollay V.2.12) there exists a nonzero vector separating $D_{\succsim}$ and ψ. Since $0 \in D_{\succsim}$ and $a\psi \notin D_{\succsim}$ for all $a > 0$, there exists $w \in \mathbb{R}^S$ such that $w \cdot \varphi \geq 0 > w \cdot \psi$, for every $\varphi \in D_{\succsim}$. It follows that $w \in \mathcal{V}$, and the proof is completed. ■

Conclusion 2.

$$A \succsim B \quad \Leftrightarrow \quad v \cdot (\mathbf{1}_A - \mathbf{1}_B) \geq 0 \ \ \textit{for every } v \in \mathcal{V}.$$

For every event $A \in \Sigma$ and vector $v \in \mathcal{V}$, denote $v(A) = v \cdot \mathbf{1}_A = \sum_{s \in A} v(s)$. According to the Non-Triviality assumption, $\mathcal{V} \neq \{0\}$. By Positivity, $\mathbf{1}_A \in D_{\succsim}$ for all $A \in \Sigma$, therefore $v(A) \geq 0$ for every $v \in \mathcal{V}$. It follows that the set $\mathcal{P} = \{\pi = v/v(S) \mid v \in \mathcal{V} \setminus \{0\}\}$ is a nonempty set of additive probability measures over Σ, such that:

$$A \succsim B \quad \Leftrightarrow \quad \pi(A) \geq \pi(B) \text{ for every } \pi \in \mathcal{P}.$$

By its definition, $\mathcal{P}$ is the maximal set w.r.t. inclusion that represents the relation.

The other direction, from the representation to the axioms, is trivially implied from properties of probability measures (GFC directly follows by taking expectation on both sides).

5.2 Proof of Theorem 2

5.2.1 Proof of the direction (i)⇒(ii)

Claim 4. *If* $A \succ\succ B$ *, then* $\mathbf{1}_A - \mathbf{1}_B$ *is an interior point of* $D_{\succsim}$.

[12]For $x = (x_1, ..., x_{|S|}), y = (y_1, ..., y_{|S|}) \in \mathbb{R}^S$, $x \cdot y$ denotes the inner product of x and y. That is, $x \cdot y = \sum_{i=1}^{|S|} x_i y_i$.

Proof. By definition, $A \succ\succ B$ implies that there exists a partition $\{A_1, \ldots, A_k\}$ of A and a partition $\{B'_1, \ldots, B'_l\}$ of B^c, such that for all i, j, $A \setminus A_i \succsim B \cup B_j$. First observe that it cannot be that $A = \emptyset$, for it would imply, on the one hand, that $A = \emptyset \succsim B$, by definition of strong preference, and on the other hand, by Generalized Finite Cancellation, $\emptyset \succsim B^c \;\Leftrightarrow\; B \succsim S$, contradicting Non-Triviality. Similarly $B = S$ is impossible. Hence, $k, l \geq 1$. Using the definition of strong preference, monotonicity of $\succsim$ w.r.t set inclusion and the structure of $D_{\succsim}$ obtains:

$$\begin{aligned}
&\mathbf{1}_A - \mathbf{1}_B - \mathbf{1}_{A_i} \in D_{\succsim}\ ,\ i = 1, \ldots, k, \quad \text{and} \\
&\mathbf{1}_A - \mathbf{1}_B - \mathbf{1}_{B_j} \in D_{\succsim}\ ,\ j = 1, \ldots, l, \quad \text{therefore} \\
&(k+l)(\mathbf{1}_A - \mathbf{1}_B) - \mathbf{1}_A - \mathbf{1}_{B^c} = (k+l-1)(\mathbf{1}_A - \mathbf{1}_B) - \mathbf{1}_S \in D_{\succsim} \quad \Rightarrow \\
&(\mathbf{1}_A - \mathbf{1}_B) - \frac{1}{k+l-1}\mathbf{1}_S \in D_{\succsim}.
\end{aligned}$$

It is next shown that there exists a neighborhood of $\mathbf{1}_A - \mathbf{1}_B$ in $D_{\succsim}$. Let $\varepsilon < \frac{1}{2(k+l-1)}$ and let $\varphi \in B_0(S, \Sigma)$ be such that $\|\mathbf{1}_A - \mathbf{1}_B - \varphi\| < \varepsilon$. For all $s \in S$, $\varphi(s) > \mathbf{1}_A(s) - \mathbf{1}_B(s) - \frac{1}{2(k+l-1)}$, therefore φ dominates $\mathbf{1}_A - \mathbf{1}_B - \frac{1}{k+l-1}\mathbf{1}_S$. It follows that $\varphi = \mathbf{1}_A - \mathbf{1}_B - \frac{1}{k+l-1}\mathbf{1}_S + \psi$ for $\psi \in D_{\succsim}$ (since ψ is nonnegative), hence $\varphi \in D_{\succsim}$ and $\mathbf{1}_A - \mathbf{1}_B$ is an internal point of $D_{\succsim}$. ■

Claim 5. *If $\mathbf{1}_A - \mathbf{1}_B$ is on the boundary of $D_{\succsim}$, then $A \succsim B$.*

Proof. Suppose on the contrary that for some events A and B, $\mathbf{1}_A - \mathbf{1}_B$ is on the boundary of $D_{\succsim}$, yet $\neg(A \succsim B)$. As $\mathbf{1}_A - \mathbf{1}_B$ is on the boundary of $D_{\succsim}$, there exists $\varepsilon' > 0$ and $\varphi \in D_{\succsim}$ such that $\mathbf{1}_A - \mathbf{1}_B + \delta\varphi \in D_{\succsim}$ for every $0 < \delta < \varepsilon'$.

On the other hand, employing Non-Atomicity, there exists an event $F \subseteq A^c$ such that $F \succ\succ \emptyset$ and $\mathbf{1}_A - \mathbf{1}_B + \mathbf{1}_F \notin D_{\succsim}$. The previous claim entails that $\mathbf{1}_F$ is an interior point of $D_{\succsim}$, hence there exists $\varepsilon_\varphi > 0$ such that $\mathbf{1}_F + \delta\varphi \in D_{\succsim}$ for all $|\delta| < \varepsilon_\varphi$. Let $0 < \delta < \min(\varepsilon_\varphi, \varepsilon')$, then $\mathbf{1}_A - \mathbf{1}_B + \delta\varphi + \mathbf{1}_F - \delta\varphi = \mathbf{1}_A - \mathbf{1}_B + \mathbf{1}_F$ is in $D_{\succsim}$, since it is a sum of two vectors in $D_{\succsim}$. Contradiction. ■

Conclusion 3.

$$A \succsim B \quad \Leftrightarrow \quad \mathbf{1}_A - \mathbf{1}_B \in D_{\succsim}.$$

Proof. The set $D_{\succsim}$ contains, by its definition, all indicator differences $\mathbf{1}_{A'} - \mathbf{1}_{B'}$ for $A' \succsim B'$ (thus also the zero vector), and their positive linear combinations. However, by the previous claims, if $\mathbf{1}_A - \mathbf{1}_B$ may be represented as a positive linear combination of indicator differences $\mathbf{1}_{A'} - \mathbf{1}_{B'}$ for which $A' \succsim B'$, or if $\mathbf{1}_A - \mathbf{1}_B$ is on the boundary of $D_{\succsim}$, then $A \succsim B$. That is, every indicator difference $\mathbf{1}_A - \mathbf{1}_B$ in the closed convex cone generated by indicator differences indicating preference also satisfies $A \succsim B$. ■

Denote by $B(S, \Sigma)$ the space of all Σ-measurable, bounded real functions over S, endowed with the supremum norm. Denote by $ba(\Sigma)$ the space of all bounded, additive functions from Σ to $\mathbb{R}$, endowed with the total variation norm. The space $ba(\Sigma)$ is isometrically isomorphic to the conjugate space of $B(S, \Sigma)$. Since $B_0(S, \Sigma)$ is dense in $B(S, \Sigma)$, $ba(\Sigma)$ is also isometrically isomorphic to the conjugate space of $B_0(S, \Sigma)$.

Consider an additional topology on $ba(\Sigma)$. For $\varphi \in B_0(S, \Sigma)$ and $m \in ba(S, \Sigma)$, let $\varphi(m) = \int_S \varphi dm$. Every φ defines a linear functional over $ba(S, \Sigma)$, and $B_0(S, \Sigma)$ is a total space of functionals on $ba(S, \Sigma)$.[13] The $B_0(S, \Sigma)$ topology of $ba(S, \Sigma)$, by its definition, makes a locally convex linear topological space, and the linear functionals on $ba(S, \Sigma)$ which are continuous in this topology are exactly the functionals defined by $\varphi \in B_0(S, \Sigma)$. Event-wise convergence of a bounded generalized sequence μ_α in $ba(S, \Sigma)$ to μ is identical to its convergence to μ in the following topologies: the $B_0(S, \Sigma)$ topology, the $B(S, \Sigma)$ topology, and the weak* topology (see Maccheroni and Marinacci, 2001). Hence, the notion closedness of bounded subsets of $ba(S, \Sigma)$ is identical in all three topologies.

Let

$$\mathcal{M} = \{m \in ba(\Sigma) \mid \int_S \varphi dm \geq 0 \text{ for all } \varphi \in D_{\succsim} \}. \tag{8}$$

The set $\mathcal{M}$ is a convex cone, and contains the zero function. For a generalized sequence $\{m_\tau\}$ in $\mathcal{M}$, which converges to m in the $B_0(S, \Sigma)$ topology, $m_\tau(\xi) \to m(\xi)$ for every $\xi \in B_0(S, \Sigma)$. Therefore, having $m_\tau(\varphi) \geq 0$ for every $\varphi \in D_{\succsim}$ and every τ, yields that $m \in \mathcal{M}$. The set $\mathcal{M}$ is thus closed in the $B_0(S, \Sigma)$ topology.[14]

Claim 6.

$$\varphi \in D_{\succsim} \quad \Leftrightarrow \quad \int_S \varphi dm \geq 0 \textit{ for every } m \in \mathcal{M}.$$

[13] That is, $\varphi(m) = 0$ for every $\varphi \in B_0(S, \Sigma)$ implies that $m = 0$.

[14] This part of the proof is very similar to a proof found in 'Ambiguity from the differential viewpoint', a previous version of Ghirardato, Maccheroni and Marinacci (2004).

Proof. According to the definition of $\mathcal{M}$, it follows that if $\varphi \in D_{\succsim}$ then $\int_S \varphi dm \geq 0$ for every $m \in \mathcal{M}$. Now suppose that $\psi \notin D_{\succsim}$. Since $D_{\succsim}$ is a closed convex cone, and $B_0(S, \Sigma)$, endowed with the supnorm, is locally convex, then by a Separation Theorem (see Dunford and Schwartz (1957), Corollary V.2.12) there exists a non-zero, continuous linear functional separating $D_{\succsim}$ and ψ. Hence, since $0 \in D_{\succsim}$ and $a\psi \notin D_{\succsim}$ for all $a > 0$, there exists $m' \in ba(\Sigma)$ such that $\int_S \varphi dm' \geq 0 > \int_S \psi dm'$, for every $\varphi \in D_{\succsim}$. It follows that $m' \in \mathcal{M}$, and the proof is completed. ■

Conclusion 4.

$$A \succsim B \quad \Leftrightarrow \quad \int_S (\mathbf{1}_A - \mathbf{1}_B) dm \geq 0 \ \ \textit{for every } m \in \mathcal{M}.$$

Proof. Follows from Conclusion 3 and the previous claim. ■

According to the Non-Triviality assumption, $\mathcal{M} \neq \{0\}$. By Positivity, $\mathbf{1}_A \in D_{\succsim}$ for all $A \in \Sigma$, therefore $\int_S \mathbf{1}_A dm \geq 0$ for every $m \in \mathcal{M}$. It follows that the set $\mathcal{P} = \{\pi = m/m(S) \mid m \in \mathcal{M} \setminus \{0\}\}$ is a nonempty, $B_0(S, \Sigma)$-closed and convex set of additive probability measures over Σ, such that:

$$A \succsim B \quad \Leftrightarrow \quad \pi(A) \geq \pi(B) \text{ for every } \pi \in \mathcal{P}.$$

Observation 2. The set $\mathcal{P}$ is bounded (in the total variation norm), hence it is compact in the $B(S, \Sigma)$ topology, thus in the $B_0(S, \Sigma)$ topology (see Corollary V.4.3 in Dunford and Schwartz). Boundedness of $\mathcal{P}$ also implies that it is weak* closed.

The set $\mathcal{P}$, by its definition, is maximal w.r.t. inclusion (any $\pi' \notin \mathcal{P}$ yields $\int_S \varphi d\pi' < 0$ for some $\varphi \in D_{\succsim}$, hence $\pi'(A) < \pi'(B)$ for some pair of events that satisfy $A \succsim B$), $B_0(S, \Sigma)$-closed and convex.

Claim 7.

$$A \succ\succ B \quad \Rightarrow \quad \pi(A) - \pi(B) > \delta > 0 \ \textit{for every } \pi \in \mathcal{P}.$$

Proof. By definition of strong preference, $A \succ\succ B$ if and only if there exists a partition $\{G_1, \ldots, G_r\}$ of S, such that $A \setminus G_i \succsim B \cup G_j$ for all i, j. This means that there are partitions $\{A_1, \ldots, A_k\}$ of A, and $\{B'_1, \ldots, B'_l\}$ of B^c, such that $\pi(A) - \pi(A_i) \geq \pi(B) + \pi(B'_j)$ for all $\pi \in \mathcal{P}$ and all i, j. It cannot be that $\pi(A) - \pi(A_i) = 0$ or

$\pi(B) + \pi(B'_j) = 1$, since $\pi(B \cup B'_j) > 0$ for some j and $\pi(A \setminus A_i) < 1$ for some i. Hence, $k \geq 2$ and $l \geq 2$, and, for all $\pi \in \mathcal{P}$,

$$\begin{aligned} &\pi(A) - \pi(B) \geq \pi(A_i) + \pi(B'_j), \quad \text{for all } i, j \quad \Rightarrow \\ &(k + l - 1)(\pi(A) - \pi(B)) \geq 1 \end{aligned}$$

and the proof is completed with $\delta = 1/(k + l)$, for instance. ■

In particular, it follows that $\pi(A) > \pi(B)$ for every $\pi \in \mathcal{P}$, whenever $A \succ\succ B$.

Lemma 1. *The probability measures in $\mathcal{P}$ are uniformly absolutely continuous.*

Proof. Let B be an event, and suppose that $\mu'(B) > 0$ for some $\mu' \in \mathcal{P}$. The inequality implies that $\neg(B^c \succsim S)$, therefore by Non-Atomicity and claim 7 there exists an event $E \subseteq B$ that satisfies $\mu(E) > 0$ for every $\mu \in \mathcal{P}$. It follows that $\mu(B) > 0$ for every $\mu \in \mathcal{P}$, hence $\mu(B) > 0 \Leftrightarrow \mu'(B) > 0$, for every event B and any $\mu, \mu' \in \mathcal{P}$.

Non-Atomicity further implies that $\neg(B^c \cup E \succsim S)$, which implies that $\mu(B \setminus E) > 0$ for some, hence for all, $\mu \in \mathcal{P}$. Therefore, if $\mu(B) > 0$ there exists $E \subset B$ with $\mu(B) > \mu(E) > 0$. All probability measures in $\mathcal{P}$ are thus non-atomic.

According to the above arguments, there exists an event F_1 such that $0 < \mu(F_1) < 1$ for all $\mu \in \mathcal{P}$. As this implies $\neg(F_1^c \succsim S)$, it follows from Non-Atomicity that there exists a partition of F_1, $\{E_1, \ldots, E_m\}$, such that:

$$E_i \succ\succ \emptyset \quad \text{and} \quad \neg(\emptyset \succsim F_1 \setminus E_i) \ \text{ for } i = 1, \ldots, m.$$

For a fixed i, the preference $E_i \succ\succ \emptyset$ entails that there exists a partition of S, $\{G_1, \ldots, G_{r_i}\}$, that satisfies $E_i \succsim G_j$, $j = 1, \ldots, r_i$. Taking the refinement of the partitions for each i, there exists a partition $\{G_1, \ldots, G_r\}$ such that $E_i \succsim G_j$ for all $i = 1, \ldots, m$ and $j = 1, \ldots, r$. It follows that for each i, j, $\mu(G_j) \leq \mu(E_i)$ for every $\mu \in \mathcal{P}$. As $\emptyset \succsim \emptyset$, the partition $\{E_1, \ldots, E_m\}$ must consist of at least two atoms. Hence, for each j,

$$\mu(G_j) \leq \frac{1}{m} \sum_{i=1}^{m} \mu(E_i) \leq \frac{1}{2}\mu(F_1) < \frac{1}{2}, \ \text{for all } \mu \in \mathcal{P}.$$

Let $F_2 = G_j$ for G_j such that $\mu(G_j) > 0$ (there must exist such j since the G_j's partition S). Again $\neg(\emptyset \succsim F_2)$, and by Non-Atomicity there exists a partition of F_2,

$\{E'_1, \ldots, E'_l\}$, such that:

$$E'_i \succ\succ \emptyset \ \text{ and } \ \neg(\emptyset \succsim F_2 \setminus E'_i) \ \text{ for } i = 1, \ldots, l.$$

As in the previous step, it follows that $l \geq 2$ and there exists a partition $\{G'_1, \ldots, G'_k\}$ with $\mu(G'_j) \leq \mu(E'_i)$ for all indices i, j and all probabilities $\mu \in \mathcal{P}$. Thus, for all j,

$$\mu(G'_j) \leq \frac{1}{l} \sum_{i=1}^{l} \mu(E'_i) \leq \frac{1}{2}\mu(F_2) < \frac{1}{4}, \text{ for all } \mu \in \mathcal{P}.$$

In the same manner, for all $n \in \mathbb{N}$ there exists a partition $\{G_1, \ldots, G_r\}$ such that for all j, $\mu(G_j) < \frac{1}{2^n}$ for all $\mu \in \mathcal{P}$. It follows that the probabilities in $\mathcal{P}$ are uniformly absolutely continuous. ■

The proof of the direction $(i) \Rightarrow (ii)$ is in fact completed. One more corollary is added, that will prove useful in the sequel. This is the opposite direction to Claim 7.

Corollary 2. *For any events A and B, if there exists $\delta > 0$ such that $\mu(A) - \mu(B) > \delta$ for every $\mu \in \mathcal{P}$, then $A \succ\succ B$.*

Proof. Let A and B be events such that $\mu(A) - \mu(B) > \delta > 0$. By uniform absolute continuity, there exists a partition $\{G_1, \ldots, G_r\}$ of S such that for all j, $\mu(G_j) < \delta/2$ for every $\mu \in \mathcal{P}$. It is so obtained that

$$\mu(A \setminus G_i) > \mu(A) - \delta/2 > \mu(B) + \delta/2 > \mu(B \cup G_j), \text{ for all } i, j.$$

■

Conclusion 5.

$$A \succ\succ B \quad \Leftrightarrow \quad \mu(A) - \mu(B) > \delta > 0 \ , \ \textit{for every } \mu \in \mathcal{P}.$$

5.2.2 Proof of the direction (ii)⇒(i)

Suppose that for every $A, B \in \Sigma$, $A \succsim B$ if and only if $\pi(A) \succsim \pi(B)$ for every probability measure π in a $\succsim$-maximal set $\mathcal{P}$, and that $\mathcal{P}$ is uniformly absolutely continuous, and

all probabilities in the set are non-atomic. Assumptions P1 through P5 are shown to hold.

P1 Reflexivity and P2 Positivity.
For every $A \in \Sigma$ and every $\pi \in \mathcal{P}$, $\pi(A) \geq \pi(A)$ and $\pi(A) \geq 0$, hence $A \succsim A$ and $A \succsim \emptyset$.

P3 Non-Triviality. The $\succsim$-maximal set $\mathcal{P}$ is nonempty, thus $\pi(B) > \pi(A)$ for some $A, B \in \Sigma$ and $\pi \in \mathcal{P}$, implying $\neg(A \succsim B)$.

P4 Generalized Finite Cancellation.
Let $(A_i)_{i=1}^n$ and $(B_i)_{i=1}^n$ be two collections of events in Σ, such that $A_i \succsim B_i$ for all i, and $\sum_{i=1}^n (\mathbf{1}_{A_i}(s) - \mathbf{1}_{B_i}(s)) \leq k\,(\mathbf{1}_A(s) - \mathbf{1}_B(s))$ for all $s \in S$, for some $k \in \mathbb{N}$ and events $A, B \in \Sigma$. Then for every π in $\mathcal{P}$, $k\mathbb{E}_\pi\,(\mathbf{1}_A - \mathbf{1}_B) \geq \sum_{i=1}^n \mathbb{E}_\pi(\mathbf{1}_{A_i} - \mathbf{1}_{B_i}) \geq 0$. It follows that $\pi(A) \geq \pi(B)$ for every $\pi \in \mathcal{P}$, hence $A \succsim B$.

Claim 8. *If $\mu(F) > 0$ for an event F and probability measure $\mu \in \mathcal{P}$, then $F \succ\succ \emptyset$.*

Proof. Suppose that F is an event with $\mu(F) > 0$. By uniform absolute continuity of $\mathcal{P}$, there exists a partition $\{E_1, \ldots, E_m\}$ of S, such that $\mu(F) > \mu(E_i)$ for all i and all $\mu \in \mathcal{P}$. In the same manner, for any $F \setminus E_i$ there exists a partition $\{G_1^i, \ldots, G_{k_i}^i\}$ such that $\mu(F \setminus E_i) > \mu(G_j^i)$ for all j and all $\mu \in \mathcal{P}$. Let $\{G_1, \ldots, G_r\}$ be the refinement of $\{E_1, \ldots, E_m\}$ and the partitions for each i, then $\mu(F \setminus G_i) > \mu(G_j)$ for all i, j and all $\mu \in \mathcal{P}$. By definition, $F \succ\succ \emptyset$. ■

P5 Non-Atomicity.
Suppose that $\neg(A \succsim B)$. By the representation assumption, $\mu'(B) > \mu'(A)$ for some $\mu' \in \mathcal{P}$. Note that necessarily $\mu'(A^c) > 0$. It is required to show that there exists a partition $\{A_1', \ldots, A_k'\}$ of A^c such that for all i, $A_i' \succ\succ \emptyset$ and $\neg(A \cup A_i' \succsim B)$.

Uniform absolute continuity of the set $\mathcal{P}$ implies that there exists a partition $\{G_1, \ldots, G_r\}$ of S, such that for all j, $\mu(G_j) < \mu'(B) - \mu'(A)$ for all μ, thus specifically for μ'. The partition $\{G_1, \ldots, G_r\}$ induces a partition $\{A_1', \ldots, A_k'\}$ of A^c such that $\mu'(A_i') > 0$ and $\mu'(A \cup A_i') < \mu'(B)$. By the representation and the previous claim, $A_i' \succ\succ \emptyset$ and $\neg(A \cup A_i' \succsim B)$ for all i.

5.3 Proof of Theorem 3

First assume that $\succsim^l$ satisfies P1-P4, and P5 if S is infinite. Then, according to theorems 1 and 2, there exists a set $\mathcal{P}$ of probability measures such that $A \succsim^l B$ if and only if $\mu(A) \geq \mu(B)$ for every $\mu \in \mathcal{P}$.

Assume further that $\succsim$ and $\succsim^l$ satisfy Cumulative Dominance. Let f and g be acts such that $L(f, \mu)$ first order stochastically dominates $L(g, \mu)$ for every $\mu \in \mathcal{P}$. By definition, $L(f, \mu)(\{y \in X \mid y \succsim x\}) \geq L(g, \mu)(\{y \in X \mid y \succsim x\})$, namely $\mu(f \succsim x) \geq \mu(g \succsim x)$, for every $x \in X$ and every $\mu \in \mathcal{P}$. Employing the subjective multi-prior representation of $\succsim^l$ obtains that $\{f \succsim x\} \succsim^l \{g \succsim x\}$ for every $x \in X$, hence by Cumulative Dominance, $f \succsim g$.

Suppose now that $\succsim$ is multi-prior probabilistically sophisticated. In other words, suppose there exists a set $\mathcal{P}$ of probability measures over Σ, such that $\succsim^l$ admits a subjective multi-prior representation with $\mathcal{P}$, and for acts f and g, $f \succsim g$ whenever $L(f, \mu)$ first order stochastically dominates $L(g, \mu)$ for every $\mu \in \mathcal{P}$. If S is infinite, assume also that the set $\mathcal{P}$ is uniformly absolutely continuous. According to theorems 1 and 2, $\succsim^l$ satisfies axioms P1-P4, and also P5 if S is infinite. For Cumulative Dominance, let f and g be acts such that $\{f \succsim x\} \succsim^l \{g \succsim x\}$ for all $x \in X$. The multi-prior representation of $\succsim^l$ implies that $\mu(f \succsim x) \geq \mu(g \succsim x)$ for every $x \in X$ and every $\mu \in \mathcal{P}$, hence $L(f, \mu)$ first order stochastically dominates $L(g, \mu)$ for every $\mu \in \mathcal{P}$, yielding that $f \succsim g$ as required.

5.4 Proof of Proposition 1

If $\mathcal{P}_2 \subseteq \mathcal{P}_1$ then obviously $\pi(A) \geq \pi(B)$ for every $\pi \in \mathcal{P}_2$, whenever $\pi(A) \geq \pi(B)$ for every $\pi \in \mathcal{P}_1$. In the other direction, let $\succsim_1$ and $\succsim_2$ be two binary relations over Σ with a subjective multi-prior probability representation. Denote by $\mathcal{P}_1$ the $\succsim_1$-maximal set and by $\mathcal{P}_2$ the $\succsim_2$-maximal set. Assume that for every pair of events A and B, $A \succsim_1 B$ implies $A \succsim_2 B$. By the representation, it follows that $\pi(A) \geq \pi(B)$ for every $\pi \in \mathcal{P}_1$ implies $\pi(A) \geq \pi(B)$ for every $\pi \in \mathcal{P}_2$. The proof that $\mathcal{P}_2 \subseteq \mathcal{P}_1$ is given separately for the finite and the infinite cases.

5.4.1 Proof of (i)⇒(ii) when S is finite

Recall the construction of $\mathcal{V}$ and the derived $\succsim$-maximal set of probabilities from the proof of Theorem 1. Let $\mathcal{V}_1 = \{a\pi \mid a \geq 0, \pi \in \mathcal{P}_1\}$, $\mathcal{V}_2 = \{a\pi \mid a \geq 0, \pi \in \mathcal{P}_2\}$ be the closed convex cones associated with $\mathcal{P}_1$ and $\mathcal{P}_2$, respectively. Suppose on the contrary

that there exists $\pi' \in \mathcal{P}_2 \setminus \mathcal{P}_1$. As $\pi' \notin \mathcal{P}_1$, then also $a\pi' \notin \mathcal{V}_1$ for every $a \geq 0$, and since $\mathcal{V}_1$ is a cone, a separation theorem yields that there exists a non-zero $\varphi \in \mathbb{R}^S$ for which $\varphi \cdot v \geq 0 > \varphi \cdot$, for every $v \in \mathcal{V}_1$.

Let $D_{\succsim_1}$, $D_{\succsim_2}$ denote the closed convex cones generated by indicator differences $\mathbf{1}_A - \mathbf{1}_B$ for $A \succsim B$ (see the definition of $D_{\succsim}$ in the proof of Theorem 1). Employing Claim 3, the separating φ satisfies $\varphi \in D_{\succsim_1} \setminus D_{\succsim_2}$. That is, $\varphi = \sum_{i=1}^n \alpha_i(\mathbf{1}_{A_i} - \mathbf{1}_{B_i})$ for $A_i \succsim_1 B_i$, but $\sum_{i=1}^n \alpha_i(\mathbf{1}_{A_i} - \mathbf{1}_{B_i}) \notin D_{\succsim_2}$. It follows that there exist events A, B such that $A \succsim_1 B$ while not $A \succsim_2 B$. Contradiction.

5.4.2 Proof of (i)⇒(ii) when S is infinite

For $\varphi \in B_0(S, \Sigma)$ and $m \in ba(S, \Sigma)$, let $\varphi(m) = \int_S \varphi dm$. Every φ defines a linear functional over $ba(S, \Sigma)$, and $B_0(S, \Sigma)$ is a total space of functionals on $ba(S, \Sigma)$.[15] Consider the $B_0(S, \Sigma)$ topology of $ba(S, \Sigma)$. By its definition, in this topology $ba(S, \Sigma)$ is a locally convex linear topological space, and the linear functionals on $ba(S, \Sigma)$ which are continuous in the $B_0(S, \Sigma)$ topology are exactly the functionals defined by $\varphi \in B_0(S, \Sigma)$.

Recall the construction of $\mathcal{M}$ and the derived $\succsim$-maximal set of probabilities from the proof of Theorem 2. Let $\mathcal{M}_1 = \{a\pi \mid a \geq 0, \pi \in \mathcal{P}_1\}$, $\mathcal{M}_2 = \{a\pi \mid a \geq 0, \pi \in \mathcal{P}_2\}$ be the closed convex cones associated with $\mathcal{P}_1$ and $\mathcal{P}_2$, respectively. Suppose on the contrary that there exists $\pi' \in \mathcal{P}_2 \setminus \mathcal{P}_1$. As $\pi' \notin \mathcal{P}_1$, then also $a\pi' \notin \mathcal{M}_1$ for every $a \geq 0$, and a Separation Theorem, employed for the $B_0(S, \Sigma)$ topology of $ba(S, \Sigma)$, yields that there exists a non-zero $\varphi \in B_0(S, \Sigma)$, separating $\mathcal{M}_1$ and π'. Since $\mathcal{M}_1$ is a cone, there exists φ for which $\int_S \varphi dm \geq 0 > \int_S \varphi d\pi'$, for every $m \in \mathcal{M}_1$.

Let $D_{\succsim_1}$, $D_{\succsim_2}$ denote the closed convex cones generated by indicator differences $\mathbf{1}_A - \mathbf{1}_B$ for $A \succsim B$ (see Claim 5, the definition of $D_{\succsim}$ in the proof of Theorem 2 and Observation 1). Employing Claim 6, the separating φ satisfies $\varphi \in D_{\succsim_1} \setminus D_{\succsim_2}$. That is, $\varphi = \sum_{i=1}^n \alpha_i(\mathbf{1}_{A_i} - \mathbf{1}_{B_i})$ for $A_i \succsim_1 B_i$, but $\sum_{i=1}^n \alpha_i(\mathbf{1}_{A_i} - \mathbf{1}_{B_i}) \notin D_{\succsim_2}$. It follows that there exist events A, B such that $A \succsim_1 B$ while not $A \succsim_2 B$. Contradiction.

5.5 Proof of Proposition 2

Assume that (i) of the proposition holds. By Theorem 2, for any two events A and B, $A \succsim B$ if and only if $\mu(A) \geq \mu(B)$ for every probability measure μ in a uniformly

[15] That is, $\varphi(m) = 0$ for every $\varphi \in B_0(S, \Sigma)$ implies that $m = 0$.

absolutely continuous set $\mathcal{P}$. According to Savage **?**, $\succsim$ on $\mathcal{B}$ is represented by a unique probability, hence all probabilities in $\mathcal{P}$ coincide on $\mathcal{B}$. Savage's theorem also implies that the common part on $\mathcal{B}$ is convex-ranged. Denote the restriction of $\mathcal{P}$ to $\mathcal{B}$ by π. As $\succsim$ on $\mathcal{B}$ is complete, $\succsim$ and $\succsim'$ coincide there. Note that, according to Observation 2, $\mathcal{P}$ is compact, thus it attains its infimum (see Lemma I.5.10 in Dunford and Schwartz **?**).

Let $E \in \Sigma$ and $B \in \mathcal{B}$ be events, and suppose that $\pi(B) > \min_{\mu\in\mathcal{P}} \mu(E)$. Then $\pi(B) > \mu'(E)$ for some $\mu' \in \mathcal{P}$, and as $\pi(B) = \mu'(B)$, then by the representation of $\succsim$ it follows that $\neg(E \succsim B)$. Employing Ambiguity Aversion yields the preference $B \succ' E$. In the other direction, if $\min_{\mu\in\mathcal{P}} \mu(E) \geq \pi(B)$, then $\mu(E) \geq \mu(B)$ for all $\mu \in \mathcal{P}$, therefore $E \succsim B$ and by Consistency also $E \succsim' B$. Thus, summing the two implications,

$$E \succsim' B \quad \Leftrightarrow \quad \min_{\mu\in\mathcal{P}} \mu(E) \geq \pi(B). \tag{9}$$

Recall that π over $\mathcal{B}$ is convex-ranged. Hence, for any $E \in \Sigma$ there exists $B \in \mathcal{B}$ with $\pi(B) = \min_{\mu\in\mathcal{P}} \mu(E)$. Next it is shown that $E \sim' B$ must hold. First, by (9), $E \succsim' B$. Suppose on the contrary that $E \succ' B$. By Unambiguous Non-Atomicity, there exists an unambiguous partition of S, $\{B_1, \ldots, B_m\}$ with $B_i \in \mathcal{B}$ for all i, such that $E \succ' B \cup B_i$. According to the above arguments, $\min_{\mu\in\mathcal{P}} \mu(E) \geq \pi(B \cup B_i)$ for all i, hence $\min_{\mu\in\mathcal{P}} \mu(E) > \pi(B)$. Contradiction. Therefore $E \sim' B$.

Let E and F be events, and B and C events in $\mathcal{B}$, such that $\pi(B) = \min_{\mu\in\mathcal{P}} \mu(E)$ and $\pi(C) = \min_{\mu\in\mathcal{P}} \mu(F)$. Then, employing Transitivity (P1'),

$$E \succsim' F \quad \Leftrightarrow \quad B \succsim' C \quad \Leftrightarrow \quad B \succsim C \quad \Leftrightarrow \quad \min_{\mu\in\mathcal{P}} \mu(E) \geq \min_{\mu\in\mathcal{P}} \mu(F) \tag{10}$$

REFERENCES

Anscombe, F.J. and R.J. Aumann (1963): "A Definition of Subjective Probability", *The Annals of Mathematical Statistics*, 34, 199-205.

Aumann, R.J. (1962): "Utility Theory without the Completeness Axiom", *Econometrica*, Vol. 30, Issue 3, 445-462.

Berger, J. (1994): "An overview of robust Bayesian analysis", *Test*, 3 (1994), 559.

Bewley, T.F. (2002): "Knightian decision theory. Part I", *Decisions in Economics and Finance*, Vol. 25, Issue 2, 79-110.

Blume, L.E., D. Easley, and J.Y. Halpern (2009): "Constructive Decision Theory", *Economics Series, Institute for Advanced Studies*, No. 246.

de Finetti, B. (1931): "Sul significato soggettivo della probabilita ", *Fund. Math.*, Vol. 17, 298-329.

de Finetti, B. (1937): "La Provision: Ses Lois Logiques, ses Sources Subjectives," *Annales de l'Institut Henri Poincare* 7, 1-68. Translated into English by Henry E. Kyburg Jr., "Foresight: Its Logical Laws, its Subjective Sources," in Henry E. Kyburg Jr. & Howard E. Smokler (1964, Eds), Studies in Subjective Probability, 93-158, Wiley, New York; 2nd edition 1980, 53-118, Krieger, New York.

de Finetti, B. (1974): Theory of Probability, Vol. 1. New York: John Wiley and Sons.

Dubra J., F. Maccheroni, and E.A. Ok (2004): "Expected utility theory without the completeness axiom",*Journal of Economic Theory*, Vol. 115, No. 1, 118-133.

Dunford N., and J.T Schwartz (1957): Linear Operators, Part I. Interscience, New York.

Ellsberg D. (1961): "Risk, Ambiguity, and the Savage Axioms", *The Quarterly Journal of Economics*, Vol. 75, No. 4, 643-669.

Epstein, L.G. (1999): "A Definition of Uncertainty Aversion", *The Review of Economic Studies*, Vol. 66, No. 3, 579-608.

Epstein, L.G., and J. Zhang (2001): "Subjective Probabilities on Subjectively Unambiguous Events", *Econometrica*, Vol. 69, No. 2, 265-306.

Fishburn, Peter C. (1970), "Utility Theory for Decision Making." Wiley, New York.

Fishburn, P.C. (1986), "The Axioms of Subjective Probability", *Statistical Science*, Vol. 1, No. 3, 335-345.

Galaabaatar, T., and E. Karni (2011) "Subjective Expected Utility with Incomplete Preferences", mimeo.

Ghirardato, P., F. Maccheroni, and M. Marinacci (2004), "Differentiating Ambiguity and Ambiguity Attitude", *Journal of Economic Theory*, 118, 133173.

Ghirardato, P., F. Maccheroni, M. Marinacci, and M. Siniscalchi (2003), "A subjective spin on roulette wheels", *Econometrica*, 71, 1897-1908.

Gilboa, I., F. Maccheroni, M. Marinacci and D. Schmeidler (2010), "Objective and Subjective Rationality in a Multiple Prior Model", *Econometrica*, Vol. 78, No. 2, 755770.

Giron, F., and S. Rios (1980) "Quasi-Bayesian Behaviour: A more realistic approach to decision making?", *Trabajos de Estadstica y de Investigacin Operativa*, Vol. 31, No. 1, 17-38.

Kopylov, I. (2007), "Subjective probabilities on 'small' domains", *Journal of Economic Theory*, 133, 236-265.

Kraft, C.H., J.W. Pratt, and A. Seidenberg (1959), "Intuitive Probability on Finite Sets", *The Annals of Mathematical Statistics*, Vol. 30, No. 2, 408-419.

Krantz, D.H., R.D. Luce, P. Suppes, and A. Tversky (1971), *Foundations of Measurement*, New York: Academic Press.

Lehrer, E., and D. Schmeidler (2005), "Savage's Theorem Revisited ", mimeo.

Maccheroni F., and M. Marinacci (2001), "A Heine-Borel theorem for $ba(\Sigma)$", RISEC.

Machina, M.J., and D. Schmeidler (1992), "A More Robust Definition of Subjective Probability", *Econometrica*, Vol. 60, No. 4, 745780.

Narens, L. (1974), "Minimal conditions for additive conjoint measurement and qualitative probability", *Journal of Mathematical Psychology*, Vol. 11, Issue 4, 404-430.

Nehring, K. (2009), "Imprecise probabilistic beliefs as a context for decision-making under ambiguity", *Journal of Economic Theory*, 144, 10541091.

Ok, E.A., P. Ortoleva and G. Riella (2008), "Incomplete Preferences under Uncertainty: Indecisiveness in Beliefs vs. Tastes", Manuscript.

Ramsey, F.P. (1931): "Truth and Probability", The Foundation of Mathematics and Other Logical Essays. New York: Harcourt Brace.

Sarin, R., and P. Wakker (1992), "A Simple Axiomatization of Nonadditive Expected Utility", *Econometrica*, Vol. 60, No. 6, 1255-1272.

Sarin, R., and P. Wakker (2000), "Cumulative dominance and probabilistic sophistication", *Mathematical Social Sciences*, Vol. 40, 191196.

Savage, L.J. (1952, 1954), The Foundations of Statistics, Wiley, New York (2nd ed. 1972, Dover, New York).

Scott, D. (1964), "Measurement Structures and Linear Inequalities", *Journal of Mathematical Psychology*, Vol. 1, 233-247.

von Neumann, J. and O. Morgenstern (1944), Theory of Games and Economic Behavior, Princeton University Press, Princeton, 1944.

Wakker, P.P. (1981), "Agreeing Probability Measures for Comparative Probability Structures", *The Annals of Statistics*, Vol. 9, No. 3, 658-662.

Wakker, P.P. (2010) "Prospect Theory for Risk and Ambiguity.", Cambridge University Press, Cambridge, UK.

Purely Subjective Maxmin Expected Utility[*][†]

Shiri Alon [‡]and David Schmeidler[§]

This version, May 4, 2013

Abstract

The Maxmin Expected Utility decision rule suggests that the decision maker can be characterized by a utility function and a set of prior probabilities, such that the chosen act maximizes the minimal expected utility, where the minimum is taken over the priors in the set. Gilboa and Schmeidler axiomatized the maxmin decision rule in an environment where acts map states of nature into simple lotteries over a set of consequences. This approach presumes that objective probabilities exist, and, furthermore, that the decision maker is an expected utility maximizer when faced with risky choices (involving only objective probabilities). This paper presents axioms for a derivation of the maxmin decision rule in a purely subjective setting, where acts map states to points in a connected topological space. This derivation does not rely on a pre-existing notion of probabilities, and, importantly, does not assume the von Neuman & Morgenstern (vNM) expected utility model for decision under risk. The axioms employed are simple and each refers to a bounded number of variables.

[*]Acknowledgements: We are thankful for helpful comments to Edi Karni, Peter Klibanoff, Fabio Maccheroni and an anonymous referee. We are especially grateful to Itzhak Gilboa and Peter Wakker for very detailed remarks. Finally, thanks are due to the Israeli Science Foundation for a partial financial support.

[†]This is an extended and revised version of the paper presented under the same title at RUD 2008 in Oxford UK.

[‡]Bar-Ilan University, shiri.aloneron@gmail.com

[§]Interdisciplinary Center Herzliya, Tel Aviv University, and Ohio State University, davidschmeidler@gmail.com

Purely Subjective Maxmin Expected Utility

Contents

Purely Subjective Maxmin Expected Utility

last revised 4-5-2013

1 Introduction

There is a respectable body of literature dealing with axiomatic foundations of decision theory, and specifically with the non-Bayesian (or extended Bayesian) branch of it. We will mention a part of this literature to put our paper in context.

Building on the works of Ramsey (1931), de Finetti (1937), and von Neumann and Morgenstern (1944, 1953), Savage (1952, 1954) provided an axiomatic model of purely subjective expected utility maximization. The descriptive validity of this model was put in doubt long ago. Today many think that Savage's postulates do not constitute a sufficient condition for rationality, and some doubt that they are all necessary conditions for it. However, almost all agree that his work is by far the most beautiful and important axiomatization ever written in the social or behavioral sciences. "The crowning glory", as Kreps (1998, p. 120) put it. Savage's work has had a tremendous influence on economic modeling, convincing many theorists that the only rational way to make decisions is to maximize expected utility with respect to a subjective probability. Importantly, due to Savage's axioms, many believe that any uncertainty can and should be reduced to risk, and that this is the only reasonable model of decision making on which economic applications should be based. About a decade after Savage's seminal work Anscombe and Aumann (1963) (AA for short) suggested another axiomatic derivation of subjective probability, coupled with expected utility maximization. As in Savage's model, acts in AA's model map states of nature to a set of consequences. However, in Savage's model the set of consequences has no structure, and it may consist of merely two elements, whereas AA assume that the consequences are lotteries as in vNM's model, namely, distributions over a set of outcomes, whose support is finite. Moreover, AA impose the axioms of vNM-'s- theory (including the independence axiom) on preferences over acts, which imply that the decision maker maximizes expected utility in the domain of risk. On the other hand, AA's model can deal with a finite state space, whereas Savage's axioms imply that there are infinitely many states, and, moreover, that none of them is an "atom". Since in many economic applications there are only finitely many states, one cannot invoke Savage's theorem to justify the expected utility hypothesis in such models.

A viable alternative to the approaches of Savage and of AA is the assumption that the set of consequences is a connected topological space. (See Fishburn (1970) and Kranz, Luce, Suppes, and Tversky (1971) and the references therein.) These spaces are "rich" and therefore more restrictive than Savage's abstract set of consequences. On the other hand, such spaces are natural in many applications. In particular, considering a consumer problem under uncertainty the consequences are commodity bundles which, in the tradition of neoclassical consumer theory, constitute a convex subset of an n-dimensional Euclidean space, and thus a connected topological space. As opposed to AA's model, the richness of the space is not necessarily derived from mixture operations on a space of lotteries. Thus, no notion of probability is presupposed, and no restrictions are imposed on the decision maker's behavior under risk.

Despite the appeal of Savage's axioms, the Bayesian approach has come under attack on descriptive and normative grounds alike. In accord with the view held by Keynes (1921), Knight (1921), and others, Ellsberg (1961) showed that Savage's axioms are not necessarily a good description of how people behave, because people tend to prefer known to unknown probabilities. Moreover, some researchers argue that such preferences are not irrational. This was also the view of Schmeidler (1989), who suggested the first axiomatically based, general-purpose model of decision making under uncertainty, allowing for a non-Bayesian approach and a not necessarily neutral attitude to uncertainty. Schmeidler axiomatized expected utility maximization with a non-additive probability measure (also known as capacity), where the operation of integration is done as suggested by Choquet (1953-4). Schmeidler (1989) employed the AA model, thereby using objective probabilities and restricting attention to expected utility maximization under risk. Following his work, Gilboa (1987) and Wakker (1989) axiomatized Choquet Expected Utility theory (CEU) in purely subjective models; the former used Savage's framework, whereas the latter employed connected topological spaces. Thus, when applying CEU, one can have a rather clear idea of what the model implies even if the state space is finite, as long as the consequence space is rich (or vice versa).

Gilboa and Schmeidler (1989) (GS hereafter) suggested the theory of Maxmin Expected Utility (MEU), according to which beliefs are given by a set of probabilities, and decisions are aimed to maximize the minimal expected utility of the act chosen (see also Chateauneuf (1991)). This model has an overlap with CEU: when the

non-additive probability is convex, CEU can be described as MEU with the set of probabilities being the core of the non-additive probability. More generally, CEU can capture modes of behavior that are incompatible with MEU, including uncertainty-liking behavior. On the other hand, there are many MEU models that are not CEU. Indeed, even with finitely many states, where the dimension of the non-additive measures is finite, the dimension of closed and convex sets of measures is infinite (if there are at least three states).[1] Moreover, there are many applications in which a set of priors can be easily specified even if the state space is not explicitly given. As a result, there is an interest in MEU models that are not necessarily CEU. GS axiomatized MEU in AA's framework, paralleling Schmeidler's original derivation of CEU.

There are several reasons to axiomatize the MEU model without using objective probabilities. The raison d'être of the CEU and MEU models (and of several more recent models) is the assumption that, contrary to Savage's claim, Knightian uncertainty cannot be reduced to risk. Thus to require the existence of exogenously given additive probabilities while modeling Knightian uncertainty appear to be in a conceptual dissonance. Another drawback of the AA framework is the assumption that it is immaterial whether objective or subjective uncertainty is resolved first. While this assumption is natural under subjective expected utility, it is questionable under non-expected utility models. In addition, alternatives in the AA framework are two-stage acts which are remote from individuals' experience and economic models.

Most of the applications in economics assume consequences lie in a convex subset of, say, a Euclidean space, but do not assume linearity of the utility function. It is therefore desirable to have an axiomatic derivation of the MEU decision rule in this framework that is applicable to a finite state space and that does not rely on objective probabilities. Such a derivation can help us see more clearly what the exact implication of the rule is in many applications, without restricting the model in terms of decision making under risk or basing it on vNM utility theory.

Our goal here is to suggest a set of axioms delivering MEU representation, namely

[1] The space of all convex compact subsets of the plane with the Hausdorff metric can be embedded as a non negative cone of a linear topological space of infinite dimension. In this cone the set of all convex closed subsets of some non degenerate triangle includes an open set of the linear topological space. Such a triangle can represent all probability distributions over some state space with three states.

the existence of a utility function u on the set of consequences X, and a non empty, compact and convex set C of finitely additive probability measures on (S, Σ) (the states set, the events set), such that for any two acts f and g:

$$f \succsim g \quad \Leftrightarrow \quad \min_{P \in C} \int_S u(f(\cdot))dP \geq \min_{P \in C} \int_S u(g(\cdot))dP \ .$$

The first axiomatic derivation of MEU of this type is by Casadesus-Masanell, Klibanoff and Ozdenoren (2000). Another one is by Ghirardato, Maccheroni, Marinacci and Siniscalchi (2003). These are discussed in subsections 2.2 and 3.2, correspondingly. To compare these works with ours, we continue the introduction with a comment on axiomatization of the individual decision making under (Knightian) uncertainty. Such an axiomatization consists of a set of restrictions on preferences. This set usually constitutes a necessary and sufficient condition for the numerical representation of the preferences.

Axiomatizations may have different goals and uses, be distinguished across more than one normative or positive interpretations and be evaluated according to different tests. Here we emphasize and delineate simplicity and transparency as a necessary condition for an axiom. Let us start with a normative interpretation of the axioms.

In many decision problems, the decision maker does not have well-defined preferences that are accessible to him or her by introspection, and has to invest time and effort to evaluate his or her options. In accordance with the models discussed in this paper we assume that the decision maker constructs a states space and conceives the options as maps from states to consequences. The role of axioms in this situation is twofold. First, the axioms may help the decision maker to construct his or her mostly unknown preferences: the axioms may be used as "inference rules", using some known instances of pairwise preferences to derive others.[2] Second, the set of axioms can be used as a general rule that justifies a certain decision procedure. To the extent that the decision maker can understand the axioms and finds them agreeable, the representation theorem might help the decision maker to choose a decision procedure, thereby reducing the problem to the evaluation of some parameters. It is often the case that the mathematical representation of the decision rule also makes the evaluation of the parameters a simple task. For example, in the case of EU theory, one can use simple trade-off questions to evaluate one's utility function and one's subjective probability, and then use the theory to put these together in a unique way that

[2]On this see Gilboa et al. (2010).

satisfies the axioms.

Both normative interpretations of the axioms call for clarity and simplicity. In the first, instance-by-instance interpretation, the axioms are used by the decision maker to derive conclusions such as $f \succsim h$ from premises such as $f \succsim g$ and $g \succsim h$. Only simple and transparent axioms can be used by actual decision makers to construct their preferences. In the second interpretation, the axioms should be accepted by the decision maker as universal statements to which the decision maker is willing to commit. If the axioms fail to have a clear meaning, a critical decision maker will hesitate to accept the axioms and the conclusions that follow from them.

Finally, we briefly mention that descriptive interpretations of the axioms also favor simple over complex ones. One descriptive interpretation is the literal one, suggesting testing the axioms experimentally. The simpler the axioms, the easier it is to test them. Another descriptive interpretation is rhetorical: the axioms are employed to convince a listener that a certain mode of behavior may be more prevalent than it might appear at first sight. Here again, simplicity of the axioms is crucial for such a rhetorical task.

But how does one measure the simplicity of an axiom? One can define a very coarse measure of opaqueness of an axiom by counting the number of variables in its formulation: acts, consequences, states or events, instances of preference relation, and logical quantifiers. For example, transitivity requires three variables, three preference instances, and one implication – seven in all. It should be pointed out that such a measure of opaqueness is a test, and not a criterion, for simplicity and transparency of an axiom. In the present work we apply this measure very roughly. By a low or satisfactory level of opaqueness we mean smaller than some fixed number, say, smaller than 100. If an axiom is stated using some notations (same as definitions) their opaqueness is included in the count. Our purpose in this work is to introduce a purely subjective axiomatization of the MEU model with axioms of satisfactory opaqueness. One exception is the axiom of continuity. (See the discussion in Subsection 2.1, after the presentation of our continuity assumption.) Actually, we are not going to count the variables in the axioms we present. Simply, an axiom of unsatisfactory opaqueness means that it is of unbounded opaqueness. The latter applies to the continuity axiom.

We conclude the Introduction with a short outline. Because our goal is to derive

the representation formula as stated in the title of the paper we are left with the statement of axioms and the proof. This is the order of presentation. But to state the axioms we need first a strategy for the proof. An obvious possibility is to follow GS. They accomplished it in several steps. First they constructed a vNM utility over consequences, and a numerical representation of the preferences over acts, where the latter coincides with the utility on constant acts. Next they reduced acts to functions from states to utiles, i.e., to elements of $\mathbb{R}^S$, and the functional from acts to functional on $\mathbb{R}^S$. Finally they translated the axioms on preferences to properties of the functional on $\mathbb{R}^S$, and showed that this functional can be represented via a set of priors.

One way to carry out the first step, in which a utility function is derived, is to take an "off the shelve" result based on a purely subjective axiomatization. There are two obvious candidates. The first one is the tradeoffs approach, or more precisely, a special case of a theorem in Kobberling and Wakker (2003), **KW henceforth**. The theorem deals with the CEU model, but both, MEU model and CEU model, share basic axioms in the purely subjective framework as they do in the AA framework. Moreover, an MEU representation on binary acts (i.e., acts with two values) is also its CEU representation. However there is still a distance between KW result we quote and the cardinal utility result we need. It is explained in Subsection 2.1, and the proof of the required result is provided in the beginning of the Appendix. The latter includes all the proofs. Our new axioms and the main result, expressed in tradeoffs language, are presented in Subsection 2.2.

Another possible way to derive a utility function over consequences is to use the biseparable preferences representation by Ghirardato and Marinacci (2001), **GM henceforth**. Their axioms and results are introduced in Subsection 3.1, and in Subsection 3.2 our new axioms are expressed in their language. For the latter, we also require a result from Ghirardato, Maccheroni, Marinacci and Siniscalchi (2003). Finally, Section 4 contains extensions, discussion of further possible extensions, and additional results.

2 Tradeoffs approach

The following notation is used throughout the entire paper. Let S denote the set of states, Σ a σ-algebra of events over S, X the set of consequences,

$\mathcal{F} = \{f : S \to X \mid f$ obtains finitely many values and is measurable w.r.t. $\Sigma\}$ the set of (simple) acts, and $\succsim \subset \mathcal{F} \times \mathcal{F}$ the decision maker's preferences over acts. As usual, $\sim$ and $\succ$ denote the symmetric and asymmetric components of $\succsim$. For $x, y \in X$ and an event E, xEy stands for the act which assigns x to the states in E and y otherwise. Let $\overline{x}$ denote the constant act xEx (for any $E \in \Sigma$). Without causing too much of a confusion, we also use the symbol $\succsim$ for a binary relation on X, defined by: $x \succsim y$ iff $\overline{x} \succsim \overline{y}$.

2.1 Basic axioms and tradeoff consistency

We start with the restrictions on the set of consequences.

A0. Structural assumption:
S is nonempty, X is a connected, topological space, and X^S is endowed with the product topology.

A0 is an essential restriction of our approach. It follows the neoclassical economic model and the literature on separability.[3] For its introduction into decision theory proper see Fishburn (1970) and the references there, and Kranz, Luce, Suppes, and Tversky (1971) and the references there.[4] We note that the topology does not have to be assumed. One can use the order topology on X whose base consists of open intervals of the form $\{ z \in X \mid x \succ z \succ y \}$, for some $x, y \in X$. The restriction then is the connectedness of the space. A0 can be further weakened by imposing the order topology and its connectedness on $X/\sim$, i.e., on the $\sim$-equivalence classes. The connectedness restriction excludes decision problems where there are just a few deterministic consequences like in medical decisions. On the other hand many medical consequences are measured by QALY, quality-adjusted life years, that is, in time units. These situations can be embedded in our model.

We state now the four basic axioms. The first one is a simple weak order assumption, which is clearly a necessary condition for an MEU representation.

A1. Weak Order:

(a) For all f and g in $\mathcal{F}$, $f \succsim g$ or $g \succsim f$ (completeness).

[3]It is still an open problem to obtain a purely subjective representation of MEU in a Savage-like model, where X may be finite but S is rich enough.

[4]Debreu (1959) is the earliest paper we know. The mathematical foundations go back to Blaschke in the period between the two world wars.

(b) For all f, g, and h in $\mathcal{F}$, if $f \succsim g$ and $g \succsim h$ then $f \succsim h$ (transitivity).

A2. Continuity:
The sets $\{f \in \mathcal{F} \mid f \succ g\}$ and $\{f \in \mathcal{F} \mid f \prec g\}$ are open for all g in $\mathcal{F}$.

A remark about the opaqueness of the axiom of continuity is required. All the main representation theorems require such an axiom in one form or another. But any version of this axiom has an unbounded or infinite opaqueness as specified above. Ours is akin to continuity in the neoclassical consumer theory. As such it cannot be tested. Moreover, the decision maker should be agnostic toward it. Because of its opaqueness, it is as difficult for the decision maker to accept it as to reject it.

A3. Essentiality:
There exist an event $E \in \Sigma$ and consequences x, y such that $\overline{x} \succ xEy \succ \overline{y}$.

This axiom simplifies and shortens the formal presentation. If there is no event and consequences as in A3, then we are either in a deterministic setting or in a degenerate setting.

The last basic axiom is the usual monotonicity condition. like A1, it is simple and necessary for an MEU representation.

A4. Monotonicity:
For any two acts f and g, $f \succsim g$ holds whenever $f(s) \succsim g(s)$ for all states s in S.

Let us consider once again the spacial case where $X = \mathbb{R}^l_+$, the standard neoclassical consumption set. Then X^S consists of state contingent consumption bundles, like in Chapter 7 of the classic 'Theory of value', Debreu (1959). Our A4 is "orthogonal" to the neoclassical monotonicity of preferences on $\mathbb{R}^l_+$. The latter says that increase in quantity of any commodity is desirable. A4 says that preferences between consequences do not depend on the state of nature that occurred. However if one assumes neoclassical monotonicity on $(\mathbb{R}^l_+)^S = X^S$, it does not imply A4, but can accommodate it. Chapter 7 preferences on $(\mathbb{R}^l_+)^S$ can be represented by a continuous utility function, say, $J : (\mathbb{R}^l_+)^S \to \mathbb{R}$. One then can define a utility where for $x \in X = \mathbb{R}^l_+$, $u(x) = J(\bar{x})$. If

also A4 is imposed, u represents $\succsim$ on X and is continuous.[5] To get such result under A0 assumption, a separability condition has to be added. However we are interested in a stronger representation where u, and hence J, are cardinal. The 'why' and 'how' are discussed in the next subsection.

We next present the specific concept of tradeoffs which is appropriate for this work. In order to do that we need to recall and restate several definitions.

Definition 1. *A set of acts is* ***comonotonic*** *if there are* ***no*** *two acts f and g in the set and states s and t, such that, $f(s) \succ f(t)$ and $g(t) \succ g(s)$. Acts in a comonotonic set are said to be* ***comonotonic****.*

Comonotonic acts induce essentially the same ranking of states according to the desirability of their consequences. Given any numeration of the states, say $\pi : S \to \{1, ..., |S|\}$, the set $\{ f \in X^S \mid f(\pi(1)) \succsim \ldots \succsim f(\pi(|S|) \}$ is comonotonic. It is a largest-by-inclusion comonotonic set of acts.

Definition 2. *Given a comonotonic set of acts $\mathcal{A}$, an event E is said to be* ***comonotonically nonnull on*** $\mathcal{A}$ *if there are consequences x, y and z such that $xEz \succ yEz$ and the set $\mathcal{A} \cup \{xEz, yEz\}$ is comonotonic.*

The definition of tradeoff indifference we use restricts attention to binary comonotonic acts, that is, comonotonic acts which obtain at most two consequences.

Definition 3. *Let a, b, c, d be consequences. We write $< a; b > \sim^* < c; d >$ if there exist consequences x, y and an event E such that,*

$$aEx \sim bEy \quad \text{and} \quad cEx \sim dEy \tag{1}$$

with all four acts comonotonic, and E comonotonically nonnull on this set of acts.

Given this notation we can express the cardinality of a utility function within our framework.

Definition 4. *We say that $u : X \longrightarrow R$ cardinally represents $\succsim$ on X (or for short, u is cardinal), if it represents $\succsim$ ordinally, and*

$$< a; b > \sim^* < c; d > \Rightarrow \; u(a) - u(b) = u(c) - u(d) \; .$$

[5]More precisely: the way $\succsim$ was defined u represents it on X with or without A4. However without A4 this representation does not make sense.

It will be shown in the sequel that given our axioms so far, and A5 below, such u exists, and that u and v cardinally represent $\succsim$ iff $v(x) = \alpha u(x) + \beta$ for some positive number α, and any number β.

A special case of Definition 3 is of the form, $< a; b > \sim^* < b; c >$. It essentially says that b is half the way between a and c. It will be used in the statement of axioms A6 and A7 below. To make the definition of tradeoffs indifference useful we need the following axiom.

A5. Binary Comonotonic Tradeoff Consistency (BCTC): For any eight consequences $a, b, c, d,\ x, y, v, w$, and events E, F,

$$aEx \sim bEy, \quad cEx \sim dEy, \quad aFv \sim bFw \ \Rightarrow \ cFv \sim dFw \tag{2}$$

whenever the sets of acts $\{\ aEx, bEy, cEx, dEy\ \}$ and $\{\ aFv, bFw, cFv, dFw\ \}$ are comonotonic, E is comonotonically nonnull on the first set, and F is comonotonically nonnull on the second.

This axiom guarantees that the relations $aEx \sim bEy$ and $cEx \sim dEy$, defining the $\sim^*$ equivalence relation above, do not depend on the choice of E, x and y. This axiom is a weakening of the one used in KW. Note that the axiom involves only a finite number of variables (even if the concepts of comonotonicity of acts, and events being comonotonically nonnull are replaced by their definitions).

2.2 New axioms and the main result

As mentioned in the Introduction, in an AA type model the set of consequences X consists of an exogenously given set of all lotteries over some set of deterministic consequences, say Z. As a result, for any two acts f and g, and $\theta \in [0, 1]$, the act $h = \theta f + (1 - \theta)g$ is well defined where $h(s) = \theta f(s) + (1 - \theta)g(s)$ for all $s \in S$. This mixture is used in the statement of two axioms central in the derivation of the MEU decision rule. One axiom, *uncertainty aversion*, states that $f \succsim g$ implies $\theta f + (1 - \theta)g \succsim g$ for any $\theta \in [0, 1]$. The other axiom, that of *certainty independence*, uses a mixture of acts with constant acts. Without lotteries the sets X and $\mathcal{F}$ are not linear spaces, and the two axioms have to be restated without availability of the mixture operation. Recall, however, that X is a connected topological space.

Given acts f, g, and h, we introduce two ways to express the intuition that g is half way between f and h. One is by tradeoffs and the $\sim^*$ notation, and the other within the biseparable preferences approach in Section 3. In any case, under the basic set of axioms in either approach, a utility function u is derived, and for consequences x, y and z, y being half way between x and z implies $u(y) = \frac{1}{2}(u(x) + u(z))$. An act g is considered to be half way between acts f and h whenever this holds state-wise, so that $g(s)$ is half way between $f(s)$ and $h(s)$.

Employing a definition of 'half way', one way to formalize certainty independence goes as follows: Let f, g, and h be acts with $f \succ g$, and h a constant act. Suppose that for some $k \geq 1$ and acts $f = f_0, f_1, f_2, \ldots, f_{k-1}, f_k = h$, and $g = g_0, g_1, g_2, \ldots, g_{k-1}, g_k = h$, f_i is half way between f_{i-1} and f_{i+1}, and g_i is half way between g_{i-1} and g_{i+1}, for $i = 1, \ldots, k-1$. Then $f_i \succ g_i$ for $i = 1, \ldots, k-1$. This essentially is the way that Casadesus-Masanell et al. (2000) stated the axiom.[6] Thus their certainty independence axiom is of unbounded opaqueness.

As explained we would like to avoid, whenever possible, axioms of the form: "for any positive integer n, and for any two n-lists, etc...". Our uncertainty aversion axiom states that if $f \succsim h$ and g is half way between f and h, then $g \succsim h$. Our certainty independence axiom is: Suppose that g is half way between f and a constant act $\bar{w}$, and for some constant acts $\bar{x}$ and $\bar{y}$, $\bar{y}$ is half way between $\bar{x}$ and $\bar{w}$. Then $f \sim \bar{x}$ iff $g \sim \bar{y}$.

However our version of the certainty independence axiom does not suffice to obtain the required representation. An axiom named *Certainty Covariance* is added. Certainty Covariance says that given acts f and g, and consequences x and y : If, for all states s, the strength of preference of $f(s)$ over $g(s)$ is the same as the strength of preference of x over y, then, $f \sim \bar{x}$ iff $g \sim \bar{y}$. In other words: when consequences are translated into utiles, the conditions of the axiom imply that the change from the vector of utiles $(u(f(s)))_{s \in S}$, to the vector of utiles $(u(g(s)))_{s \in S}$, is parallel to the move on the diagonal from the constant vector with coordinates $u(x)$ to the constant vector with coordinates $u(y)$. The axiom requires that indifference (between $\bar{x}$ and f) be preserved by movements parallel to the diagonal in the utiles space.

We now formally present the three axioms introduced above, using the $\sim^*$ relation.

[6]Such sequences are called standard sequences; See Krantz et al. (1971) and the references there.

A6. Uncertainty Aversion:
For any three acts f, g, and h, if $f \succsim g$, and for all states $s \in S$, $< f(s); h(s) > \sim^* < h(s); g(s) >$, then $h \succsim g$.

A7. Certainty Independence:
Suppose that two acts, f and g, and three consequences, x, y, and w, satisfy $< x; y > \sim^* < y; w >$, and for all states $s \in S$, $< f(s); g(s) > \sim^* < g(s); w >$. Then $g \sim \overline{y}$ iff $f \sim \overline{x}$.

A8. Certainty Covariance:
Let f, g be acts and x, y consequences such that for all states $s \in S$, $< f(s); g(s) > \sim^* < x; y >$. Then, $f \sim \overline{x}$ iff $g \sim \overline{y}$.

As mentioned earlier, the three axioms are stated in the language of tradeoffs indifference $\sim^*$ and not in the language of the preferences on X (or $\mathcal{F}$). Yet, even if all instances of $\sim^*$ are replaced by their definition, all three axioms require a bounded number of variables.

The axiom of uncertainty aversion, A6, is a tradeoffs version of the uncertainty aversion axiom introduced by Schmeidler (1989, preprint 1984). Essentially, it replaces utility mixtures with the use of tradeoffs terminology. It is analogous to the uncertainty aversion axiom in a model with exogenous lotteries for the case of $\theta = 1/2$. Applying the axiom consecutively, and then using continuity, A2, guarantees the conclusion for any mixture. A6 expresses uncertainty aversion in that it describes the decision maker's will to reduce the impact of not knowing which state will occur. The reduction is achieved by averaging the consequences of f and g in every state.

In a similar manner to A6, axiom A7 produces a utilities analogue of the certainty independence axiom phrased by GS in their axiomatization of MEU, for the case $\theta = 1/2$. Consecutive application of A7 will yield the analogue for $\theta = 1/2^m$, where θ is the coefficient of the non-constant act. To get the analogue for all dyadic mixtures we have to supplement it with A8.

Certainty Covariance is a subjective equivalent of an axiom Grant and Polak

(2011) call constant absolute uncertainty aversion.[7] Analogously to the notion of constant absolute risk aversion, this axiom states that preference is maintained if alternatives are altered by a constant shift across all states. Grant and Polak use mixtures with exogenous probabilities to phrase the axiom in an AA setting, while we use tradeoffs to express the idea of a constant shift. Our Certainty Covariance axiom thus asserts that whenever an act f is indifferent to a constant act $\overline{x}$, then shifting both f and $\overline{x}$ by a constant shift across all states, the resulting acts, g and the constant $\overline{y}$, are still indifferent. Consequently, indifference curves for the relation are parallel.

When acts are represented in the utiles space, one can see that axiom A8 is some version of a parallelogram where one side is an interval on the diagonal. This is the order interval $[x, y]$. The parallel side is $[f, g] \approx [f(s), g(s)]_{s \in S}$, which should be thought of as an off diagonal "order" interval in X^S. The other two sides of the parallelogram are delineated by equivalences: $f \sim \overline{x}$ iff $g \sim \overline{y}$.

Having stated all axioms, we can formulate our main result.

Theorem 1. *Suppose that a binary relation $\succsim$ on $\mathcal{F}$ is given, and the structural assumption A0 holds. Then the following two statements are equivalent:*

(1) $\succsim$ satisfies

- *(A1) Weak Order*
- *(A2) Continuity*
- *(A3) Essentiality*
- *(A4) Monotonicity*
- *(A5) Binary Comonotonic Tradeoff Consistency*
- *(A6) Uncertainty Aversion*
- *(A7) Certainty Independence*
- *(A8) Certainty Covariance*

(2) There exist a continuous utility function $u : X \to \mathbb{R}$ and a non-empty, closed and convex set C of additive probability measures on Σ, such that, for all

[7] The two axioms were formulated independently, and presented at RUD 2008 in Oxford.

$f, g \in \mathcal{F}$,

$$f \succsim g \quad \Leftrightarrow \quad \min_{P \in C} \int_S u(f(\cdot))dP \geq \min_{P \in C} \int_S u(g(\cdot))dP \ .$$

Furthermore, the utility function u is unique up to an increasing linear transformation, the set C is unique, and for some event E, $0 < \min_{P \in C} P(E) < 1$.

The requirement that for some event E, $0 < \min_{P \in C} P(E) < 1$, reflects the assumption that there is an essential event. As mentioned earlier, this assumption is made merely for the sake of ease of presentation and may be dropped. All proofs appear in the appendix.

3 Biseparable approach

3.1 Basic axioms and act independence

First we recall the definition of a biseparable relation from GM.

Definition 5. *A functional $J : \mathcal{F} \to \mathbb{R}$ represents the relation $\succsim$ whenever: $J(f) \geq J(g)$ if and only if $f \succsim g$. The representation is monotonic if for any two acts f and g, $J(f) \geq J(g)$ whenever $f(s) \succsim g(s)$ for all states s.*

Definition 6. *An event E is essential if for some consequences x and y, $x \succ xEy \succ y$.*

Definition 7. *A preference relation over acts, $\succsim$, is said to be biseparable if it admits a nontrivial, monotonic representation J, and there exists a set function η on events such that for binary acts with $x \succsim y$,*

$$J(xEy) = u(x)\eta(E) + u(y)(1 - \eta(E)). \tag{3}$$

The function u is defined on X by, $u(x) = J(\overline{x})$, and it represents $\succsim$ on X. The set function η, when normalized s.t. $\eta(S) = 1$, is unique, and u and J are unique up to a positive multiplicative constant and an additive constant whenever there exists an essential event.

GM showed that this representation generalizes CEU and MEU. They characterized biseparable preferences using several axioms, among which are A0, A3 and A4 above. For the sake of uniformity, we will continue to assume our continuity axiom A2, which is somewhat stronger than the continuity axiom assumed by GM. The new axioms required for biseparability of the preferences will be denoted with *. First of all an

additional structural assumption, used by GM, is stated:

A0*. Separability:
The topology on X is separable.

Essentiality alone is not enough in this framework, and an additional axiom is required to guarantee that nonnull events are 'always' nonnull.

A3*. Consistent Essentiality:
For an event E, if for some $x \succ y$, $xEy \succ y$, (resp. $x \succ xEy$), then for all $a \succ b \succsim c$, $aEc \succ bEc$ (resp. for all $c \succsim a \succ b$, $cEa \succ cEb$).

To state the last axiom required for biseparability of preferences we first recall that for an act f, $x \in X$ is its **certainty equivalent** if $f \sim \bar{x}$. In the following definition and in the axiom the concept of certainty equivalence is employed.

Definition 8. *Let there be given two acts, f and g, and an essential event G. An act h is termed a G-mixture of f and g, if for all $s \in S$, $h(s) \sim f(s)Gg(s)$.*

Given axioms A0, A1, A2 and A4, it is obvious that a certainty equivalent of each act, and consequently event mixtures, can easily be proved to exist. However when stating the next axiom these proofs are not assumed.

A5*. Binary Comonotonic Act Independence:
Let two essential events, D and E, and three pairwise comonotonic, binary acts, aEb, cEd, and xEy be given. Suppose also that either xEy weakly dominates *aEb and cEd, or is weakly dominated by them. Then $aEb \succsim cEd$ implies that a D-mixture of aEb and xEy is weakly preferred to a D-mixture of cEd and xEy*, provided that both mixtures exist.

GM (2001) Theorem 11 says that assuming A0 and A0*, preferences are biseparable (with the uniqueness results) iff they satisfy A1, A2, A3, A3*, A4, and A5*.

The theorem implies that if $x \succ y \succ z$, and $u(y) = u(x)/2 + u(z)/2$, then y is half way between x and z. However we would like to express that y is half way between x and z in the language of preferences, and not by using an artificial construct like

utility. Ghirardato et al. (2003) did it in their Proposition 1. They showed that if the preferences over acts are biseparable, and $x \succ y \succ z$ in X, then:

$$u(y) = u(x)/2 + u(z)/2 \;\; iff$$

$$\exists\, a, b \in X, \textit{ and an essential event } E \textit{ s.t. } \bar{a} \sim xEy,\; \bar{b} \sim yEz, \textit{ and } xEz \sim aEb. \quad (4)$$

The line above is a behavioral definition of "y is half way between x and z." We denote it by $y \in \mathcal{H}(x, z)$. [8] We will use this notation to state our new axioms.

Note that both A3* and A5* involve a bounded number of variables in their formulation. On the other hand, Separability (A0*) is as complex as continuity. But it comes for free if X is in a Euclidean space with a topology induced by the Euclidean distance.

3.2 New axioms and the main result

For formulation of the theorem using biseparable preferences we simply repeat the new axioms with the $\mathcal{H}(\cdot, \cdot)$ notation.

A6*. Uncertainty Aversion:
For any three acts f, g, and h with $f \succsim g$: if $\forall s \in S$, $h(s) \in \mathcal{H}(f(s), g(s))$, then $h \succsim g$.

A7*. Certainty Independence:
Suppose that two acts, f and g, and three consequences, x, y, and w, satisfy:
$y \in \mathcal{H}(x, w)$, and $\forall s \in S$, $g(s) \in \mathcal{H}(f(s), w)$. Then $g \sim \overline{y}$ iff $f \sim \overline{x}$.

A8*. Certainty Covariance:
Let f, g be acts and x, y consequences such that $\forall s \in S$, $\mathcal{H}(f(s), y) = \mathcal{H}(x, g(s))$. Then, $f \sim \overline{x}$ iff $g \sim \overline{y}$.

As opposed to our application of the 'half way' notation $\mathcal{H}(\cdot, \cdot)$ in the axioms, Ghirardato et al. (2003), in their MEU axiomatization, introduce an artificial construct, $\oplus$, coupled with a mixture set $(M, \hat{=}, \hat{+})$. Their C-independence axiom states that for all $\alpha \in [0, 1]$, $f \succsim g$ implies $\alpha f \oplus (1-\alpha)\bar{x} \succsim \alpha g \oplus (1-\alpha)\bar{x}$. The $\oplus$ notation stands for a consecutive application of the $\mathcal{H}(\cdot, \cdot)$ notation (for all $s \in S$), where the number of

[8] This is Definition 4 in Ghirardato et al. (2003).

times it is applied may be infinite. Their axiom therefore requires an infinite number of variables.

The relation between A6, A7 and A8, and their starred counterparts, is derived from the observation that under the rest of the axioms (in each of the approaches), if $< x; y > \sim^* < y; z >$ then $y \in \mathcal{H}(x, z)$. Note that the relation between A8 and A8* is also based on the Euclidean geometry theorem which says: A (convex) quadrangle is a parallelogram iff its diagonals bisect each other. The main result within the biseparable approach is stated below. It's proof appears in the appendix.

Theorem 2. *Suppose that a binary relation $\succsim$ on $\mathcal{F}$ is given, and the structural assumptions A0 and A0* hold. Then the following two statements are equivalent:*

(1) $\succsim$ satisfies

- *(A1) Weak Order*
- *(A2) Continuity*
- *(A3) Essentiality*
- *(A3*) Consistent Essentiality*
- *(A4) Monotonicity*
- *(A5*) Binary Comonotonic Act Independence*
- *(A6*) Uncertainty Aversion*
- *(A7*) Certainty Independence*
- *(A8*) Certainty Covariance*

(2) There exist a continuous utility function $u : X \to \mathbb{R}$ and a non-empty, closed and convex set C of additive probability measures on Σ, such that, for all $f, g \in \mathcal{F}$,

$$f \succsim g \quad \Leftrightarrow \quad \min_{P \in C} \int_S u(f(\cdot))dP \geq \min_{P \in C} \int_S u(g(\cdot))dP \ .$$

Furthermore, the utility function u is unique up to an increasing linear transformation, the set C is unique, and for some event E, $0 < \min_{P \in C} P(E) < 1$.

The basic axioms in the two approaches, either A0-A5 in the tradeoffs approach or A0-A4, A0*, A3* and A5* in the biseparable approach, both yield a biseparable

representation. KW, in Subsection 5.2, outline how a biseparable representation follows from A0-A5, and in Section 7 show that Binary Comonotonic Act Independence (A5*) implies Binary Comonotonic Tradeoff Consistency (A5) (under axioms A0-A4).[9]

After assuming the set of basic axioms, the notion of a consequence y being 'half-way' between consequences x and z, in the two approaches, amounts to having $u(y) = u(x)/2 + u(z)/2$. Axioms A6-A8 or A6*-A8* can then be seen to yield the same attributes in utiles space, each expressed in its corresponding language.

4 Extensions and Comments

4.1 Bounded acts

Until now we have dealt with finite valued and measurable acts. An obvious question arises whether the main results hold when the set of acts is extended to include all measurable bounded acts.

Definition 9. *Given a complete and transitive binary relation, $\succsim$, on X, an act $a : S \to X$ is said to be bounded if there are two consequences, $x, y \in X$ s.t. for all $s \in S : x \succsim a(s) \succsim y$.*

Thus, if $\succsim$ is a binary relation on $\mathcal{F}$ satisfying A1 (Weak order), bounded acts are well defined. Furthermore, assuming A0 and A1 (The structural assumption and weak order), the set,

$$\mathcal{F}^b = \{f : S \to X \mid f \text{ is bounded and measurable w.r.t. } \Sigma\}$$

is well defined. One can ask under what additional conditions the binary relation on $\mathcal{F}$ can uniquely be extended to a binary relation on $\mathcal{F}^b$. Assuming that such conditions are satisfied we denote this extension also by $\succsim \subset \mathcal{F}^b \times \mathcal{F}^{\,b}$.

Theorem 3. *Suppose that A0 holds and $\succsim \subset \mathcal{F} \times \mathcal{F}$ satisfies A1-A8. Then $\succsim$ has a unique extension to a binary relation on $\mathcal{F}^b$, that satisfies the same axioms, and the representation (2) of Theorem 1 (or Theorem 2) holds for it.*

[9] As a result, under our non-degeneracy assumptions which imply at least two nonnull states, Separability (A0*) is in fact not required. We continue to assume it nevertheless, in order to employ the representation theorem from GM. In addition, GM assume a somewhat weaker continuity assumption than our A2. Consequently their representing functional is subcontinuous rather than continuous.

In other words, it suffices to impose the axioms on finite valued acts to get the representation for bounded acts. The proof of this result derives from Lemma 20 in the appendix.

4.2 Preference relations admitting CEU and MEU representations

Theorem 4 below is a purely subjective counterpart of a proposition in Schmeidler (1989), characterizing preference relations representable by both CEU and MEU rules. Axiom A5 is replaced by a more restrictive axiom, A5.1, below. The latter axiom, together with Axioms A1-A4, yield a purely subjective CEU representation (Definition 2). Addition of uncertainty aversion (A6) implies that the preference relation can equivalently be represented by an MEU functional, where the set of additive probabilities is the core of the nonadditive probability.

A5.1 Comonotonic Tradeoff Consistency:
For any four consequences a, b, c, d, events D, E and acts f, g, f', g',

$$aDf \sim bDg,\ cDf \sim dDg,\ aEf' \sim bEg' \Rightarrow cEf' \sim dEg' \tag{5}$$

whenever the sets of acts $\{aDf, bDg, cDf, dDg\}$ and $\{aEf', bEg', cEf', dEg'\}$ are comonotonic, D is comonotonically nonnull on the first set, and E on the second.

This definition originates in Wakker (1986,1989).

Theorem 4. *Assume that the structural assumption A0 holds, and let $\succsim$ be a binary relation on $\mathcal{F}$. Then the following two statements are equivalent:*

(1) $\succsim$ satisfies:

- *(A1) Weak Order*
- *(A2) Continuity*
- *(A3) Essentiality*
- *(A4) Monotonicity*
- *(A5.1) Comonotonic Tradeoff Consistency*
- *(A6) Uncertainty Aversion*

(2) There exist a continuous, non-constant, cardinal utility function $u : X \to \mathbb{R}$ *and a unique, convex, and nonadditive probability* η *on* Σ*, such that, for all* $f, g \in \mathcal{F}$*,*

$$f \succsim g \quad \Leftrightarrow \quad \int_S u \circ f d\eta \geq \int_S u \circ g d\eta$$

where the nonadditive probability η *satisfies*

$$\int_S u \circ f d\eta = \min\{\int_S u \circ f dP \mid P \in core(\eta)\} \tag{6}$$

Furthermore, for some event $E \in \Sigma$*,* $0 < \eta(E) < 1$*.*

The proof appears in Appendix A, subsection 5.5.

4.3 A generalization of Theorem 1

If the state space is assumed finite, then in the axioms applied to obtain Theorem 1, axiom A5, Binary Comonotonic Tradeoff Consistency, may be weakened so that consistency is required to hold only when the involved events, E and F, are either $\{s\}$ or $S \setminus \{s\}$, for any state $s \in S$. Essentiality (A3) should then be changed to state that for some state s and consequences x and y, $\overline{x} \succ x\{s\}y \succ \overline{y}$.

Under this weaker set of axioms a continuous utility function, which is unique up to an increasing linear transformation and respects tradeoff indifferences, may still be elicited. With A3 and A5 changed as explained, and A0-A2 and A4 stated as above, the resulting intermediate representation is weaker than the biseparable representation of GM (2001). The desired MEU representation however, for a finite state space S, may still be obtained.

When the state space is assumed finite, an even weaker set of assumptions may be used to derive a continuous utility function as required, which may in turn be applied as basis to obtain an MEU representation. The weaker substitutes for Essentiality (A3) and Binary Comonotonic Tradeoff Consistency (A5) are:

Consistent Comonotonic Essentiality:

(a) For all $s \in S$,

$$\begin{aligned} asx \succ bsx \quad &\Rightarrow \quad csy \succ dsy \ \text{ for all } c \succ d, \\ &\qquad \text{whenever } \{asx, bsx, csy, dsy\} \text{ are comonotonic.} \\ xsa \succ xsb \quad &\Rightarrow \quad ysc \succ ysd \ \text{ for all } c \succ d, \\ &\qquad \text{whenever } \{xsa, xsb, ysc, ysd\} \text{ are comonotonic.} \end{aligned}$$

(b) There exist distinct states s', s'' and consequences x, y such that $xs'y \succ \bar{y}$ and $\bar{x} \succ ys''x$.

Simple Binary Comonotonic Tradeoff Consistency (S-BCTC):

For any eight consequences a, b, c, d, x, y, z, w, and states s and t,

$$a\{s\}x \sim b\{s\}y, \quad c\{s\}x \sim d\{s\}y, \quad a\{t\}z \sim b\{t\}w \ \Rightarrow \ c\{t\}z \sim d\{t\}w$$

whenever the sets of acts $\{\ a\{s\}x, b\{s\}y, c\{s\}x, d\{s\}y\ \}$ and $\{\ a\{t\}z, b\{t\}w, c\{t\}z, d\{t\}w\ \}$ are comonotonic, $\{s\}$ is comonotonically nonnull on the first set and $\{t\}$ is comonotonically nonnull on the second set.

The resulting representation under these versions of A3 and A5, along with A0-A2 and A4, is weaker than the biseparable representation of GM (2001) and the representation mentioned above. For further details see Alon (2012). Nevertheless, when these axioms are supplemented with A6, A7 and A8, an MEU representation still follows.

4.4 Purely subjective Variational Preferences

An important generalization of the GS Maxmin Expected Utility model is the Variational Preferences model of Maccheroni, Marinacci and Rustichini (2006). In the latter, which also is placed in the Anscombe-Aumann setup, the axiom of Certainty Independence of GS, is replaced by the axiom of Weak Certainty Independence. Weak Certainty Independence states that if an $\alpha : 1 - \alpha$ mixture of an act f and a constant act $\overline{x}$ is preferred to an $\alpha : 1 - \alpha$ mixture of an act g and the same constant act $\overline{x}$, then, for any constant act $\overline{y}$, the $\alpha : 1 - \alpha$ mixture of f and $\overline{y}$ is preferred to the same mixture of g and $\overline{y}$. The representation of preferences which allow this seemingly slight weakening of the Certainty Independence of GS is a significant generalization of MEU representation. An obvious question arises: can we use the techniques of the

present paper to get a purely subjective model of Variational Preferences?

At first sight it seems possible. The Certainty Independence of GS is replaced by us with two axioms, Certainty Independence (A7) and Certainty Covariance (A8). As discussed in Subsection 2.2, Certainty Covariance is a purely subjective counterpart of an axiom phrased by Grant and Polak (2011) in an AA setting. Grant and Polak show that their axiom, under the basic set of axioms they assume, implies Weak Certainty Independence. The same holds under our set of basic axioms. However, our basic axioms, A1 - A5, imply that our preferences are biseparable. [10] However, Variational preferences are not biseparable. Moreover, even under weaker versions of tradeoff consistency, discussed in Subsection 4.3, the additive representation over binary acts that emerges is incompatible with the Variational model. Thus, using our method does not enable representation of purely subjective variational preferences. It still is an open problem.

5 Appendix: The proofs

We begin by listing two observations, which are standard in neoclassical consumer theory. These observations will be used in the sequel, sometimes without explicit reference.

Observation 5. *Weak order, Substantiality and Monotonicity imply that there are two consequences* $x^*, x_* \in X$ *such that* $x^* \succ x_*$.

Observation 6. *Weak order, Continuity and Monotonicity imply that each act* $f \in \mathcal{F}$ *has a certainty equivalent, i.e., a constant act* $\overline{x}$ *such that* $f \sim \overline{x}$.

5.1 Proof of Theorem 1: (1)⇒(2)

Our first step is to use A0 through A5 to derive a cardinal utility function which represents $\succsim$ on constant acts, that is, a utility function that satisfies that $< a; b > \sim^* < c; d >$ implies $u(a) - u(b) = u(c) - u(d)$. A representation on $\mathcal{F}$ is then defined using certainty equivalents.

Proposition 7. *Under axioms A0 to A5 the following is satisfied:*

[10] This has been conjectured in KW and the main steps of the proof have been outlined there. We did not prove it because it was not needed for our first representation theorem, Theorem 1.

(1) there exists a continuous function $u : X \to \mathbb{R}$ such that for all $x, y \in X$, $x \succsim y \iff u(x) \geq u(y)$, and $< a; b > \sim^* < c; d >$ implies $u(a) - u(b) = u(c) - u(d)$. Furthermore, u is unique up to a positive linear transformation.

(2) Given a function u as in (1) there exists a unique continuous $J : \mathcal{F} \to \mathbb{R}$, such that for all $x \in X$, $J(\overline{x}) = u(x)$, and for all $f, g \in \mathcal{F}$,

$$f \succsim g \iff J(f) \geq J(g) \ .$$

5.1.1 Proof of Proposition 7

Let E be an event such that $\overline{x} \succ xEy \succ \overline{y}$ for some consequences x, y (exists by Essentiality (A3)). The next claim shows that strict preferences then hold for all such consequences.

Claim 8. *Let D be an event and x, y, z consequences such that $xDz \succ yDz$ and xDz, yDz are comonotonic. Then $\alpha D\gamma \succ \beta D\gamma$ whenever α, β, γ are consequences for which $\alpha \succ \beta$ and $\{\alpha D\gamma, \beta D\gamma, xDz, yDz\}$ is a comonotonic set of acts.*

Proof. Assume w.l.o.g. that $y \succsim z$. Suppose on the contrary that $\alpha D\gamma \sim \beta D\gamma$. By assumption D is comonotonically nonnull on $\{\alpha D\gamma, \beta D\gamma\}$. Since obviously $\alpha D\gamma \sim \alpha D\gamma$ and $\alpha D\beta \sim \alpha D\beta$ (which belongs to the same comonotonic set of acts), then BCTC (A5) implies that also $\alpha D\beta \sim \beta D\beta = \bar{\beta}$. If D^c is comonotonically *null* on $\{\beta D^c\alpha\}$, then $\alpha D\beta = \beta D^c\alpha \sim \alpha D^c\alpha = \bar{\alpha}$. Otherwise, if D^c is comonotonically *nonnull* on $\{\beta D^c\alpha\}$, then from the three equivalences: $\alpha D\gamma \sim \beta D\gamma$, $\beta D\gamma \sim \beta D\gamma$ and $\beta D^c\alpha \sim \beta D^c\alpha$, applying BCTC, we conclude $\alpha D^c\alpha = \bar{\alpha} \sim \beta D^c\alpha = \alpha D\beta$. In any case, whether D^c is comonotonically null or nonnull as assumed, it is obtained that $\bar{\alpha} \sim \bar{\beta}$. Contradiction. If $z \succsim x$ then the roles of α and β are reversed and the proof remains the same. ■

We define a new decision problem. Define $\{E, E^c\}$ to be the new state space. Let $X^{\{E,E^c\}}$ be the set of acts, functions from the new state space to the set of consequences X. For consequences a and b, (a, b) denotes an act in $X^{\{E,E^c\}}$, assigning a to E and b to E^c. To avoid confusion, we use the notation (a, b) and (a, a) for acts in $X^{\{E,E^c\}}$ and reserve the notation aEb and $\bar{a}$ for acts in $\mathcal{F}$. Define a binary relation $\succsim_E$ on $X^{\{E,E^c\}}$ by:

$$\text{for all} \quad (a, b), (c, d) \in X^{\{E,E^c\}},$$
$$(a, b) \succsim_E (c, d) \iff aEb \succsim cEd, \quad aEb, cEd \in \mathcal{F}.$$

There exists a one-to-one correspondence between the sets $X^{\{E,E^c\}}$ and $\{\ aEb \mid a,b \in X\ \}$. Thus, the product topology on $X^{\{E,E^c\}}$ is equivalent to the original topology, restricted to the set $\{\ aEb \mid a,b \in X\ \}$. We summarize some of the attributes of $\succsim_E$ in the next lemma.

Lemma 9. *The binary relation $\succsim_E$ on $X^{\{E,E^c\}}$ satisfies Weak order, Continuity, Monotonicity, Non-nullity of the states E and E^c and (Binary) Comonotonic Tradeoff Consistency.*

Proof. Weak Order, Continuity, Monotonicity and Non-nullity follow from the definition of $\succsim_E$ and the attributes of the original relation $\succsim$, and from the equivalence between the topology on $X^{\{E,E^c\}}$ and the topology on X^S restricted to $\{\ aEb \mid a,b \in X\}$.

Note that when there are only two states, axioms A5 (Binary Comonotonic Tradeoff Consistency) and A5.1 (Comonotonic Tradeoff Consistency) coincide. It follows that $\succsim_E$ satisfies Comonotonic Tradeoff Consistency. ■

Additive representation of $\succsim_E$ on $X^{\{E,E^c\}}$ is obtained using Corollary 10 and Observation 9 of KW, applied to the case of two states, stated below. Uniqueness of u and ρ follows from both E and E^c being comonotonically nonnull on the set $\{aEb \mid a \succsim b\}$ (by choice of E). KW use a slightly different tradeoff consistency condition than ours, however assuming the rest of our axioms their condition follows (see proof in Subsection 5.6).

Lemma 10. *(Corollary 10 and Observation 9 of KW 2003:) Assume the conclusions of Lemma 9. Then there exists a nonadditive probability ρ on $2^{\{E,E^c\}}$ and a continuous utility function $u : X \to \mathbb{R}$, such that $\succsim_E$ is represented on $X^{\{E,E^c\}}$ by the following CEU functional U:*

$$U((a,b)) = \begin{cases} u(a)\rho(E) + u(b)[1-\rho(E)] & a \succsim b \\ u(b)\rho(E^c) + u(a)[1-\rho(E^c)] & \textit{otherwise} \end{cases} \tag{7}$$

Furthermore, u is unique up to an increasing linear transformation and ρ is unique.

Denote by U the CEU functional over acts of the form aEb, which existence is guaranteed by the lemma. Let u denote the corresponding utility function and ρ the corresponding nonadditive probability. The functional U represents $\succsim$ on acts of the form aEb and the utility function u represents $\succsim$ on X, and is unique up to an

increasing linear transformation. We proceed to show that u is cardinal, in the sense that tradeoff indifference implies equivalence of utility differences.

Let $F \neq E$ be an event. If F and F^c are comonotonically nonnull on some comonotonic set of the form aFb (either $a \succsim b$ or $b \succsim a$) then a CEU functional as in Lemma 10, representing $\succsim$ on acts of the form aFb, may be obtained. The link between this functional and U is elucidated in the following lemma.

Lemma 11. *Let $F \neq E$, and suppose that F and F^c are comonotonically nonnull on some comonotonic set of the form aFb. Let W be a CEU representation on $X^{\{F,F^c\}}$ (according to Lemma 10), and w the corresponding utility function. Then $w = \sigma u + \tau$, $\sigma > 0$.*

Proof. Denote by φ the nonadditive probability corresponding to W. Denote, for $x \in X$, $V_1(x) = u(x)\rho(E)$, and recall that $\rho(E) > 0$, as E is comonotonically nonnull on $\{aEb \mid a \succsim b\}$. Suppose that F and F^c are comonotonically nonnull on $\{aFb \mid a \succsim b\}$, and let $W_1(x) = w(x)\varphi(F)$. Binary Comonotonic Tradeoff Consistency allows us to follow the steps of Lemma VI.8.2 in Wakker 1989, to obtain $W_1 = \eta V_1 + \lambda,\ \eta > 0$, thus $w(x) = \sigma u + \tau$ (set $\sigma = \eta\rho(E)/\varphi(F) > 0$, where $\varphi(F) > 0$ follows from F being comonotonically nonnull as assumed).

Otherwise, if F and F^c are comonotonically nonnull on $\{aFb \mid a \precsim b\}$, the result may be obtained using a functional $W_2(x) = w(x)\varphi(F^c)$ (here again $\varphi(F^c) > 0$ because F^c is comonotonically nonnull as assumed). ■

Corollary 12. *For any four consequences a, b, c, d, $< a; b > \sim^* < c; d >$ implies $u(a) - u(b) = u(c) - u(d)$.*

Proof. Let $< a; b > \sim^* < c; d >$. Then there exist consequences x, y and an event F such that $aFx \sim bFy$ and $cFx \sim dFy$, with $\{aFx, bFy, cFx, dFy\}$ comonotonic and F comonotonically nonnull on this set. Assume first that F^c is comonotonically *null* on $\{aFx, bFy, cFx, dFy\}$. It is next proved that in that case it must be that $a \sim b$ and $c \sim d$.

Suppose on the contrary that $a \succ b$ (w.l.o.g.; similar arguments work for the case $b \succ a$). If $a \succsim x$ and $b \succsim y$, then aFy, bFy are comonotonic, and by nullity of F^c on the comonotonic set involved, $aFx \sim aFy \sim bFy$, and obviously also $aFy \sim aFy$, where F is assumed to be comonotonically nonnull on the comonotonic set of acts $\{aFy, bFy\}$. If $a \precsim x$ and $b \precsim y$, then similarly aFx, bFx are comonotonic, and by nullity of F^c on the comonotonic set involved, $aFx \sim bFy \sim bFx$, and also $aFx \sim aFx$, where F is comonotonically nonnull on the comonotonic set of acts $\{aFx, bFx\}$.

In any case, it is also true that $aEb \sim aEb$, where E is comonotonically nonnull on the set of acts containing aEb. By Binary Comonotonic Tradeoff Consistency it follows that $aEb \sim bEb = \bar{b}$, contradicting Claim 8. Therefore it must be that $a \sim b$. The same can be proved for the consequences c and d. Thus, if F^c is comonotonically null on the set of acts $\{aFx, bFy, cFx, dFy\}$, then the above indifference relations imply $a \sim b$ and $c \sim d$, resulting trivially $u(a) - u(b) = 0 = u(c) - u(d)$.

Suppose now that F^c is comonotonically nonnull on the relevant binary comonotonic set. According to Lemma 10 there exists a CEU representation over acts of the form aFb, with a utility function unique up to an increasing linear transformation, and a unique nonadditive probability. Denote the CEU functional by W, with a corresponding utility function w. The indifference relations $aFx \sim bFy$ and $cFx \sim dFy$ then imply $w(a) - w(b) = w(c) - w(d)$. If $F = E$ or $F = E^c$ then by the uniqueness result in Lemma 10, $w = \sigma u + \tau$ with $\sigma > 0$. If F is another event, then Lemma 11 yields this relation. In any case, $w(a) - w(b) = w(c) - w(d)$ implies $u(a) - u(b) = u(c) - u(d)$. ■

The next lemma establishes the uniqueness of u.

Lemma 13. *u is unique up to a positive linear transformation.*

Proof. Let $\hat{u}$ be some other representation of $\succsim$ on X, which satisfies that if $< a; b > \sim^* < c; d >$ then $\hat{u}(a) - \hat{u}(b) = \hat{u}(c) - \hat{u}(d)$. Both u and $\hat{u}$ represent $\succsim$ on X, therefore $\hat{u} = \psi \circ u$, where ψ is continuous and strictly increasing.

Non-triviality of $\succsim$, connectedness of X and continuity of u imply that $u(X)$ is a non-degenerate interval. Let ξ be an internal point of $u(X)$. There exists an interval R small enough around ξ such that, for all $u(a), u(c)$ and $u(b) = [u(a) + u(c)]/2$ in R, there are $x, y \in X$ for which aEx, bEy, bEx, cEy are comonotonic, with $a \succsim x$, $b \succsim y$, $b \succsim x$ and $c \succsim y$, and

$$u(a) - u(b) = (u(y) - u(x))\frac{1 - \rho(E)}{\rho(E)} = u(b) - u(c).$$

That is, $aEx \sim bEy$ and $bEx \sim cEy$, with $\{aEx, bEy, bEx, cEy\}$ comonotonic, which is precisely the definition of $< a; b > \sim^* < b; c >$. By the assumption on $\hat{u}$, $\hat{u}(a) - \hat{u}(b) = \hat{u}(b) - \hat{u}(c)$, which in fact implies that for all $\alpha, \gamma \in R$, ψ satisfies $\psi((\alpha + \gamma)/2) = [\psi(\alpha) + \psi(\gamma)]/2$. By Theorem 1 of section 2.1.4 of Aczel (1966), ψ must be a positive linear transformation. ■

Statement (1) of Proposition 7 is proved. Having a specific u, a unique representation J on $\mathcal{F}$ follows using certainty equivalents. By Observation 6 there exists, for any $f \in \mathcal{F}$, a constant act $CE(f)$ such that $f \sim CE(f)$. Set $J(f) = u(CE(f))$, then for all $f, g \in \mathcal{F}$,

$$f \succsim g \iff CE(f) \succsim CE(g) \iff J(f) = u(CE(f)) \geq u(CE(g)) = J(g)$$

That is, J represents $\succsim$ on $\mathcal{F}$. J is unique and continuous by its definition and by continuity of u. The proof of the proposition is completed.

5.2 Proof of the implication (1)⇒(2) of Theorem 1 - continued

Throughout this subsection E will denote an event that satisfies $\overline{x} \succ xEy \succ \overline{y}$ whenever $x \succ y$ (exists by Essentiality and Claim 8), and U will denote the CEU representation on binary acts of the form aEb (according to Lemma 10). u and ρ will notate the corresponding utility function on X and nonadditive probability on $2^{\{E,E^c\}}$, respectively. By choice of E, $0 < \rho(E) < 1$. As proved above, u satisfies (1) of Proposition 7, and we denote by J the representation of $\succsim$ over $\mathcal{F}$ obtained according to (2) of the proposition. By connectedness of X and continuity of u, $u(X)$ is an interval. Since u is unique up to a positive linear transformation, we fix it for the rest of this proof such that there are x^*, x_* for which $u(x_*) = -2$ and $u(x^*) = 2$. θ is a consequence such that $u(\theta) = 0$.

The proof is conducted in three logical steps. First, an MEU representation is obtained on a subset of $\mathcal{F}$. This is done by moving to work in utiles space, and applying tools from GS. Afterwards, the representation is extended to a 'stripe', in utiles space, around the main diagonal. The last step extends the representation to the entire space.

Claim 14. *There exists a subset $Y \subset X$ such that $u(Y)$ is a non-degenerate interval, $\theta \in Y$, and for all $a, b, c, d \in Y$, $u(a) - u(b) = u(c) - u(d)$ implies $< a; b > \sim^* < c; d >$.*

Proof. Employing the representation U and Continuity of u, there exists an interval $[-\tau, \tau]$ $(0 < \tau)$ small enough such that if $u(a), u(b), u(c), u(d) \in [-\tau, \tau]$ satisfy $u(a) - u(b) = u(c) - u(d)$, there are $x, y \in X$ for which $u(x), u(y) \leq -\tau$ and

$$u(a) - u(b) = \frac{1 - \rho(E)}{\rho(E)}[u(y) - u(x)] = u(c) - u(d).$$

It follows that $aEx \sim bEy$ and $cEx \sim dEy$. By choice of the consequences and the event E, the set $\{aEx, bEy, cEx, dEy\}$ is comonotonic and E is comonotonically nonnull on this set. Thus $< a; b > \sim^* < c; d >$.

Let $Y = \{y \in X \mid u(y) \in [-\tau, \tau]\}$. Y is a subset of X, $u(Y)$ is an interval, $\theta \in Y$, and for all $a, b, c, d \in Y$, $u(a) - u(b) = u(c) - u(d)$ implies $< a; b > \sim^* < c; d >$. ■

Throughout this subsection Y will denote a subset of X as specified in the above claim. Note that utility mixtures of acts in Y^S are also in Y^S (this fact is employed in the proof without further mention).

The following is a utilities analogue to Schmeidler's (1989) uncertainty aversion axiom.

A6u. Uncertainty Aversion in Utilities:
Let the preference relation $\succsim$ be represented on constant acts by a utility function u. If f, g, h are acts such that $f \succsim g$, and for all states s, $u(h(s)) = \frac{1}{2}u(f(s)) + \frac{1}{2}u(g(s))$, then $h \succsim g$.

Lemma 15. *If (1) of the theorem holds then Uncertainty Aversion in Utilities (A6u) is satisfied on $\mathcal{F} \cap Y^S$.*

Proof. Let $f, g, h \in \mathcal{F} \cap Y^S$ be acts satisfying $f \succsim g$, h such that in all states s, $u(h(s)) = \frac{1}{2}u(f(s)) + \frac{1}{2}u(g(s))$. Then for all states s, $u(f(s)) - u(h(s)) = u(h(s)) - u(g(s))$, therefore $< f(s), h(s) > \sim^* < h(s), g(s) >$. By A6 it follows that $h \succsim g$. ■

Comment 16. *Applying A6u consecutively implies that if $f, g, h \in \mathcal{F} \cap Y^S$ are acts satisfying $f \succsim g$, h an act such that for all states s, $u(h(s)) = \frac{k}{2^m}u(f(s)) + (1 - \frac{k}{2^m})u(g(s))$, ($\frac{k}{2^m} \in [0, 1]$) then $h \succsim g$. Continuity then yields that the same is true for any $\alpha : 1 - \alpha$ utilities mixture ($\alpha \in [0, 1]$).*

Claim 17. *Let $f \in \mathcal{F} \cap Y^S$ be an act, $\overline{x} \in \mathcal{F} \cap Y^S$ a constant act such that $f \sim \overline{x}$, and $\alpha \in (0, 1)$. If g is an act such that for all states s, $u(g(s)) = \alpha u(f(s)) + (1 - \alpha)u(x)$, then $g \sim \overline{x}$.*

Proof. We first prove the claim for $\alpha = \frac{1}{2^m}$, that is, when g satisfies, for all states s, $u(g(s)) = \frac{1}{2^m}u(f(s)) + (1 - \frac{1}{2^m})u(x)$. The proof is by induction on m. For $m = 1$, in all states s, $< f(s); g(s) > \sim^* < g(s); x >$, which by certainty independence (A7)

(with $\overline{w} = \overline{x}$) implies $g \sim \overline{x}$. Assume that for g that satisfies $u(g(s)) = \frac{1}{2^m}u(f(s)) + (1 - \frac{1}{2^m})u(x)$, $g \sim \overline{x}$. Let h be an act such that $u(h(s)) = \frac{1}{2^{m+1}}u(f(s)) + (1 - \frac{1}{2^{m+1}})u(x)$. The act h satisfies, for all states s, $u(h(s)) = \frac{1}{2}u(g(s)) + \frac{1}{2}u(x)$, and $g \sim \overline{x}$ by the induction assumption. Employing Certainty Independence again implies $h \sim \overline{x}$.

Now let $\alpha = \frac{k}{2^m}$ ($k \in \{2, \ldots, 2^m - 1\}$). For all states s, $u(g(s)) = \frac{k}{2^m}u(f(s)) + (1 - \frac{k}{2^m})u(x)$. By Uncertainty Aversion in Utilities (A6u), $g \succsim \overline{x}$. We assume $g \succ \overline{x}$ and derive a contradiction.

Let g' be such that $u(g'(s)) = \frac{1}{2}u(g(s)) + \frac{1}{2}u(x)$ for all states s, $\overline{y}$ a constant act such that $g \sim \overline{y}$. For all states s, $u(g(s)) \in u(Y)$, and $u(Y)$ is an interval, therefore by Monotonicity and Essentiality $u(y) \in u(Y)$. Employing Certainty Independence, $g' \sim \overline{z}$ for z such that $u(z) = \frac{u(x)+u(y)}{2} > u(x)$, so $g' \succ \overline{x}$. However, There exists an act h such that $u(h(s)) = \frac{j}{2^n}u(g(s)) + (1 - \frac{j}{2^n})u(g'(s))$ and also $u(h(s)) = \frac{1}{2^m}u(f(s)) + (1 - \frac{1}{2^m})u(x)$ for some $j, n, m \in \mathbb{N}$. By the previous paragraph, $h \sim \overline{x}$. But Uncertainty Aversion in Utilities requires that $h \succsim g'$, contradiction. Therefore if g satisfies $u(g(s)) = \frac{k}{2^m}u(f(s)) + (1 - \frac{k}{2^m})u(x)$ for all states s, then $g \sim \overline{x}$. By continuity it follows that the same is true for any $\alpha \in (0, 1)$. ■

Claim 18. *Let $f \in \mathcal{F} \cap Y^S$ be an act, $\overline{w}, \overline{x}, \overline{y} \in \mathcal{F} \cap Y^S$ constant acts and $\alpha \in (0, 1)$, such that $f \sim \overline{x}$ and $u(y) = \alpha u(x) + (1 - \alpha)u(w)$. If h is an act such that in all states s, $u(h(s)) = \alpha u(f(s)) + (1 - \alpha)u(w)$, then $h \sim \overline{y}$.*

Proof. Let g be an act such that for all states s, $u(g(s)) = \alpha u(f(s)) + (1 - \alpha)u(x)$, then $g \in \mathcal{F} \cap Y^S$. For all states s,

$$\begin{aligned}
&u(h(s)) = \alpha u(f(s)) + (1 - \alpha)u(w) = u(g(s)) + u(y) - u(x) \Longleftrightarrow \\
&u(h(s)) - u(g(s)) = u(y) - u(x) \quad \Longleftrightarrow \\
&< h(s); g(s) > \sim^* < y; x > .
\end{aligned}$$

According to Claim 17, $g \sim \overline{x}$, therefore by Certainty Covariance $h \sim \overline{y}$. ■

The following is an analogue to Lemma 3.3 from GS, and follows the lines of the proof given there. Denote by B_0 the set of real-valued functions on S which assume finitely many values, and by B the space of all bounded Σ-measurable real-valued functions on S. For any $\gamma \in \mathbb{R}$, $\overline{\gamma} \in B_0$ is the constant function γ^S.

Lemma 19. *There exists a functional $I : B_0 \to \mathbb{R}$ such that:*

(i) For $f \in \mathcal{F} \cap Y^S$, $I(u \circ f) = J(f)$.

(ii) I is monotonic (i.e., for $a, b \in B_0$, if $a \geq b$ then $I(a) \geq I(b)$).

(iii) I is superadditive and homogeneous of degree 1 (therefore $I(\overline{1}) = 1$).

(iv) I satisfies certainty independence: for any $a \in B_0$ and $\overline{\gamma}$ a constant function, $I(a + \overline{\gamma}) = I(a) + I(\overline{\gamma})$.

Proof. For $a \in B_0 \cap (u(Y))^S$ define $I(a)$ by (i). By monotonicity, if in all states s, $f(s) \sim g(s)$, then $f \sim g$. I is therefore well defined.

We show that I is homogeneous on $B_0 \cap (u(Y))^S$. Let $a, b \in B_0 \cap (u(Y))^S$, $b = \alpha a$. Let f, g be acts such that $u \circ f = a$, $u \circ g = b$, and $x, y \in X$ such that $f \sim \overline{x}$, $u(y) = \alpha u(x)$. Monotonicity and the fact that Y contains θ imply that $\overline{x}, \overline{y} \in \mathcal{F} \cap Y^S$. Applying Claim 18 with $\overline{w} = \bar{\theta}$ implies that $g \sim \overline{y}$. Thus $I(b) = J(g) = u(y) = \alpha u(x) = \alpha J(f) = \alpha I(a)$, and homogeneity is established.

The functional I may now be extended from $B_0 \cap (u(Y))^S$ to B_0 by homogeneity. By its definition, I is homogeneous on B_0, and by monotonicity of the preference relation I is also monotonic. Note also that if $y \in Y$ and $u(y) = \gamma$, then $I(\overline{\gamma}) = u(y) = \gamma$, and homogeneity implies that $I(\overline{1}) = 1$. Again by homogeneity it suffices to show that certainty independence and superadditivity of I hold on $B_0 \cap (u(Y))^S$.

For certainty independence, let $a, \overline{\gamma} \in B_0 \cap (u(Y))^S$. Let $f \in \mathcal{F} \cap Y^S$ be such that $u \circ f = a$, and denote $I(a) = \beta$. Suppose $\overline{x}, \overline{y}, \overline{w} \in \mathcal{F} \cap Y^S$ are constant acts such that $u(x) = \beta$, $u(w) = \gamma$, $u(y) = (u(x) + u(w))/2$ and h an act such that in all states s, $u(h(s)) = \frac{1}{2}(u(f(s)) + u(w))$. Then $h \in \mathcal{F} \cap Y^S$, $f \sim \overline{x}$, and by Claim 18 it follows that $h \sim \overline{y}$, which implies

$$J(h) = I(u \circ h) = I(\frac{1}{2}(u \circ f + \overline{\gamma})) = u(y) = \frac{1}{2}(\beta + \gamma) = \frac{1}{2}(I(a) + \gamma).$$

Hence, using homogeneity to extend the result to B_0, for all $a, \overline{\gamma} \in B_0$, $I(a + \overline{\gamma}) = I(a) + I(\overline{\gamma})$.

Finally we show that I is superadditive on $B_0 \cap (u(Y))^S$. Let $a, b \in B_0 \cap (u(Y))^S$ and $f, g \in \mathcal{F} \cap Y^S$ such that $u \circ f = a,\ u \circ g = b$. Assume first that $I(a) = I(b)$. By Lemma 15, if h is an act such that $u \circ h = \frac{1}{2}(u \circ f + u \circ g)$, then $h \succsim g$. Thus,

$$J(h) = I(u \circ h) = I(\frac{1}{2}(u \circ f + u \circ g)) \geq J(g) = \frac{1}{2}(I(a) + I(b)).$$

so by homogeneity, if $a, b \in B_0$ satisfy $I(a) = I(b)$ then $I(a + b) \geq I(a) + I(b)$.

Now assume that $I(a) > I(b)$ and let $\gamma = I(a) - I(b),\ c = b + \overline{\gamma}$. Certainty independence of I implies $I(c) = I(b) + \gamma = I(a)$. Again by certainty independence and by superadditivity for the case $I(a) = I(b)$, it follows that

$$I(\frac{1}{2}(a + b)) + \frac{1}{2}\gamma = I(\frac{1}{2}(a + c)) \geq \frac{1}{2}(I(a) + I(c)) = \frac{1}{2}(I(a) + I(b)) + \frac{1}{2}\gamma.$$

and superadditivity is proved for all $a, b \in B_0 \cap (u(Y))^S$, and may be extended to B_0 using homogeneity. ■

Lemmas 3.4 and 3.5 from GS may now be applied to obtain an MEU representation of $\succsim$ on $\mathcal{F} \cap Y^S$.

Lemma 20. *(Lemma 3.4 from GS):*
There exists a unique continuous extension of I to B, and this extension is monotonic, superlinear and C-independent.

Lemma 21. *(Lemma 3.5 and an implied uniqueness result from GS):*
If I is a monotonic superlinear certainty-independent functional on B, satisfying $I(\overline{1}) = 1$, there exists a closed and convex set C of finitely additive probability measures on Σ such that for all $b \in B$, $I(b) = \min\{\int b dP \mid P \in C\}$. The set C is unique.

Corollary 22. *$\succsim$ is represented on $\mathcal{F} \cap Y^S$ by an MEU functional.*

Let C denote the set of additive probability measures on Σ, involved in the MEU representation of $\succsim$ on $\mathcal{F} \cap Y^S$. For all $f \in \mathcal{F} \cap Y^S$, $J(f) = I(u \circ f) = \min\{\int u \circ f dP \mid P \in C\}$. Outside of $B_0 \cap (u(Y))^S$, I is extended by homogeneity. By the above lemmas, I is monotonic, superadditive, homogeneous of degree 1 and satisfies certainty independence.

It is now required to show that the homogeneous extension of I outside of $B_0 \cap (u(Y))^S$ is consistent with the preference relation. That is, that for all $f \in \mathcal{F}$,

$J(f) = I(u \circ f) = \min\{\int u \circ f dP \mid P \in C\}$. The proof consists of two steps. First, the representation is extended to a 'stripe' (in utiles space) parallel to the main diagonal, obtained by 'sliding' $B_0 \cap (u(Y))^S$ along the diagonal. Then the representation is further extended from that stripe to the entire space.

For an act $g \in \mathcal{F}$, let $diag(g) = \{f \in \mathcal{F} \mid u \circ f = u \circ g + \overline{\gamma},\ \gamma \in \mathbb{R}\}$. That is, $diag(g)$ contains all acts that can be obtained from g by constant shifts (translations parallel to the main diagonal, in utiles space). Note that $diag(g)$ is convex w.r.t. utility mixtures.

Claim 23. *Let $g \in \mathcal{F} \cap Y^S$. Then for all acts $f \in diag(g)$, $J(f) = I(u \circ f)$.*

Proof. Let f, g be acts that satisfy $g \in \mathcal{F} \cap Y^S$ and $u \circ f = u \circ g + \overline{\varepsilon}$, $\varepsilon > 0$, so $f \in diag(g)$. Suppose that x, y are consequences such that $g \sim \overline{y}$ and $u(x) - u(y) = \varepsilon$. It follows that for all states t, $u(f(t)) - u(g(t)) = u(x) - u(y)$.

Similarly to the proof of Claim 14, if ε is small enough then there are z_1, z_2 such that $u(z_1), u(z_2) < \min_s u(g(s))$, and (recall that $\rho(E) > 0$)

$$\varepsilon = u(f(s)) - u(g(s)) = \frac{1 - \rho(E)}{\rho(E)}[u(z_2) - u(z_1)] = u(x) - u(y)$$

It follows that for all consequences $f(s)$ and $g(s)$, $f(s)Ez_1 \sim g(s)Ez_2$ and $xEz_1 \sim yEz_2$, with $\{f(s)Ez_1, g(s)Ez_2, xEz_1, yEz_2\}$ comonotonic, and E comonotonically nonnull on this set. Thus $< f(s); g(s) > \sim^* < x; y >$ for all states s. Certainty Covariance implies $f \sim \overline{x}$. Applying certainty independence of I,

$$I(u \circ f + u \circ \overline{y}) = I(u \circ f) + u(y) = I(u \circ g + u \circ \overline{x}) = J(g) + u(x),$$

and $I(u \circ f) = u(x) = J(f)$. The same procedure may be repeated (in small 'tradeoff steps') for all $f \in diag(g)$ that satisfy $u \circ f = u \circ g + \overline{\gamma}$, $\gamma \geq 0$, to obtain $J(f) = I(u \circ f)$.

In order to extend the representation to the lower part of the diagonal, let f, g be acts that satisfy $g \in \mathcal{F} \cap Y^S$ and $u \circ f = u \circ g - \overline{\varepsilon}$, $\varepsilon > 0$. Let x, y be consequences such that $g \sim \overline{y}$ and $u(y) - u(x) = \varepsilon$. Similarly to the previous case, if ε is small enough then there are z_3, z_4 such that $u(z_3), u(z_4) > \max u(g(s))$ and (recall that $\rho(E) < 1$)

$$\varepsilon = u(g(s)) - u(f(s)) = \frac{\rho(E)}{1 - \rho(E)}[u(z_3) - u(z_4)] = u(y) - u(x)$$

It follows that for all consequences $f(s)$ and $g(s)$, $f(s)E^c z_3 \sim g(s)E^c z_4$ and $xE^c z_3 \sim yE^c z_4$, with the set $\{f(s)E^c z_3, g(s)E^c z_4, xE^c z_3, yE^c z_4\}$ comonotonic, and E^c comonotonically nonnull on this set. Therefore $< f(s); g(s) > \sim^* < x; y >$ for all states s. By Certainty Covariance it follows that $f \sim \overline{x}$, and certainty independence of I again implies that $J(f) = I(u \circ f)$. Subsequent applications yield the result for all acts f such that $u \circ f = u \circ g - \overline{\gamma}$, $\gamma \geq 0$, and the proof is completed. ■

Claim 22 implies that $J(f) = \min\{\int u \circ f dP \mid P \in C\}$ for all $f \in \mathcal{F}$ that satisfy $u \circ f = u \circ g + \overline{\gamma}$ for some $g \in \mathcal{F} \cap Y^S$ and $\gamma \in \mathbb{R}$. That is, $\succsim$ is represented by an MEU functional in a 'stripe' containing all acts with utilities in $[-\tau, \tau]$, and those obtained from them by translations parallel to the main diagonal (in utiles space).

To complete the proof and show that the MEU representation holds for all acts, an auxiliary result is proved. It consists of showing that for any subset of consequences there exists a consequence ξ that has a tradeoff-midpoint with every intermediate consequence (including the extreme ones, if those exist).

Claim 24. *Let $m, M \in u(X)$ be such that $m < M$. Then there exists a consequence $\xi \in X$, $m < u(\xi) < M$, such that for all $x \in X$ with $m \leq u(x) \leq M$, there is $y \in X$ that satisfies*

$$< x; y > \sim^* < y; \xi >$$

Proof. Choose a consequence $\xi \in X$ such that $u(\xi) = \rho(E)M + (1 - \rho(E))m$. By continuity of u and connectedness of X, the definition of M and m and the fact that $0 < \rho(E) < 1$, such a consequence exists, and its utility is strictly between M and m. To show that every $x \in X$ with utility (weakly) between M and m has a tradeoff-midpoint with ξ, the cases $u(x) \geq u(\xi)$ and $u(\xi) > u(x)$ are examined separately.

Suppose first that $u(\xi) \leq u(x) \leq M$. In order to show that there exists y that satisfies $< x; y > \sim^* < y; \xi >$ it suffices to show that there are consequences a, b which satisfy $u(\xi) \geq u(b) \geq u(a) > m$, and $xEa \sim yEb$, $yEa \sim \xi Eb$. For that, using the CEU representation over binary acts contingent on E, it suffices to prove that there are $a, b \in X$ for which,

$$u(x)\rho(E) + u(a)[1 - \rho(E)] = \frac{u(x) + \rho(E)M + (1 - \rho(E))m}{2}\rho(E) + u(b)[1 - \rho(E)] \ ,$$

where $M \geq u(x) \geq u(\xi) \geq u(b) \geq u(a) \geq m$, or, rearranging the above expression,

$$u(b) - u(a) = \frac{\rho(E)}{2(1-\rho(E))}[u(x) - \rho(E)M - (1-\rho(E))m]$$

with $M \geq u(x) \geq u(\xi) \geq u(b) \geq u(a) > m$. By continuity of u and connectedness of X, such $u(b) - u(a)$ obtains all values from zero up to $u(\xi) - m = \rho(E)(M-m)$. Thus, it is enough to prove,

$$\begin{aligned} \frac{\rho(E)}{2(1-\rho(E))}[u(x) - \rho(E)M - (1-\rho(E))m] \leq \rho(E)(M-m) \quad &\Longleftrightarrow \\ u(x) - \rho(E)M - (1-\rho(E))m \leq 2(1-\rho(E))M - 2(1-\rho(E))m \quad &\Longleftrightarrow \\ u(x) \leq \rho(E)M + 2(1-\rho(E))M - (1-\rho(E))m = M + (1-\rho(E))(M-m) \ . \end{aligned}$$

As by assumption $u(x) \leq M$ and $M > m$, this condition is satisfied, and a tradeoff-midpoint of ξ and any $x \in X$ with $u(\xi) \leq u(x) \leq M$ exists as required.

Next suppose that $x \in X$ satisfies $m \leq u(x) < u(\xi)$. Similarly to the previous case, it suffices to show that there are consequences c, d which satisfy $M \geq u(d) \geq u(c) \geq u(\xi)$, and $\xi E^c c \sim y E^c d$, $y E^c c \sim x E^c d$. Using once more the CEU representation over binary acts contingent on E, it suffices to prove that there are $c, d \in X$ for which

$$u(d)\rho(E) + u(x)[1-\rho(E)] = u(c)\rho(E) + \frac{u(x) + \rho(E)M + (1-\rho(E))m}{2}[1-\rho(E)]$$

where $M \geq u(d) \geq u(c) \geq u(\xi) > u(x) \geq m$, or, rearranging the above expression,

$$u(d) - u(c) = \frac{1-\rho(E)}{2\rho(E)}[\rho(E)M + (1-\rho(E))m - u(x)]$$

with $M \geq u(d) \geq u(c) \geq u(\xi) > u(x) \geq m$. By continuity of u and connectedness of X, $u(d) - u(c)$ obtains all values in the range $[0, M - u(\xi)] = [0, (1-\rho(E))(M-m)]$. Thus, it is enough to prove,

$$\begin{aligned} \frac{1-\rho(E)}{2\rho(E)}[\rho(E)M + (1-\rho(E))m - u(x)] \leq (1-\rho(E))(M-m) \quad &\Longleftrightarrow \\ \rho(E)M + (1-\rho(E))m - u(x) \leq 2\rho(E)M - 2\rho(E)m \quad &\Longleftrightarrow \\ u(x) \geq m - \rho(E)(M-m) \end{aligned}$$

which is again true based on the assumptions that $u(x) \geq m$ and $M > m$. ■

Corollary 25. *For all $f \in \mathcal{F}$, $J(f) = I(u \circ f)$.*

Proof. As stated above, by claims 22 and 23, $\succsim$ is represented by an MEU functional on a 'stripe' containing all acts with utilities in $[-\tau, \tau]$, and those obtained from them, in utiles space, by translations parallel to the main diagonal. It is left to show that $J(f) = I(u \circ f)$ for all acts outside that stripe as well.

Let $f, g \in \mathcal{F}$ be acts and x, y, ξ consequences such that: f is not a constant act, $J(g) = I(u \circ g)$, $< f(s); g(s) > \sim^* < g(s); \xi >$ for all states s, $g \sim \overline{y}$ and $< x; y > \sim^* < y; \xi >$. The existence of the required ξ is guaranteed by applying Lemma 24 to $M = \max_{s \in S} u(f(s))$ and $m = \min_{s \in S} u(f(s))$ (x such that $u(y) = (u(x) + u(\xi))/2$ must satisfy $m \leq u(x) \leq M$). Employing the axiom of Certainty Independence (A7) it may be asserted that $f \sim \overline{x}$. Thus, by homogeneity and certainty independence of I,

$$u(y) = I(u \circ g) = \frac{1}{2}\left(I(u \circ f) + u(\xi)\right) = \frac{1}{2}\left(I(u \circ f) + 2u(y) - u(x)\right)$$

and $I(u \circ f) = u(x) = J(f)$.

In that manner the MEU representation that applies in the stripe may be extended to any act in $\mathcal{F}$ which has a tradeoff-midpoint inside the stripe. By repeating this procedure, the MEU representation is seen to apply to the entire acts space. ■

It is concluded that for all $f \in \mathcal{F}$, $J(f) = \min\{\int u \circ f dP \mid P \in C\}$, that is, $\succsim$ on $\mathcal{F}$ is represented by an MEU functional, with a utility function unique up to an increasing linear transformation, and a unique set of prior probabilities C.

The fact that for some event E, $0 < \min_{P \in C} P(E) < 1$, follows from Essentiality: the event which existence is guaranteed by Essentiality should satisfy that $\overline{x} \succ xEy \succ \overline{y}$ for $x \succ y$. Translating to the MEU representation yields $u(x) > u(x) \min_{P \in C} P(E) + u(y)(1 - \min_{P \in C} P(E)) > u(y)$, hence $0 < \min_{P \in C} P(E) < 1$. The proof of the direction (1)⇒(2) of the main theorem is completed.

5.3 Proof of the implication (2)⇒(1) of Theorem 1

By definition of the representation on $\mathcal{F}$, $\succsim$ is a weak order, satisfying Monotonicity. Define a functional $I : B_0 \to \mathbb{R}$ by $I(b) = \min\{\int_S b dP \mid P \in C\}$. Hence the preference relation is represented by $J(f) = I(u \circ f)$. By its definition, I is continuous and superlinear, therefore $\succsim$ is continuous and satisfies Uncertainty Aversion in Utilities

(A6u). Uncertainty aversion (A6) results by observing that
$< f(s), h(s) > \sim^* < h(s), g(s) >$ implies $u(h(s)) = \frac{1}{2}[u(f(s)) + u(g(s))]$.

To see that Essentiality holds let E be an event such that
$0 < \min_{P \in C} P(E) < 1$, and take $x \succ y$, which must exist according to the uniqueness result. Then $u(x) = J(\overline{x}) > u(x) \min_{P \in C} P(E) + u(y)(1 - \min_{P \in C} P(E)) = J(xEy) > u(y) = J(\overline{y})$.

Next it is proved that Certainty Independence is satisfied. If $\gamma \in \mathbb{R}$ and $\overline{\gamma}$ is the constant function returning γ in every state, then for all $b \in B_0$,
$I(b + \bar{\gamma}) = I(b) + I(\bar{\gamma}) = I(b) + \gamma$. If $\alpha > 0$ then $I(\alpha b) = \alpha I(b)$. Let f, g be acts and ξ, x, y consequences that satisfy: for all states s, $< f(s), g(s) > \sim^* < g(s), \xi >$ and $< x; y > \sim^* < y; \xi >$. Then $u(g(s)) = \frac{1}{2}[u(f(s)) + u(\xi)]$ and $u(y) = \frac{1}{2}[u(x) + u(\xi)]$. Therefore,

$$J(g) = I(u \circ g) = \frac{1}{2} I(u \circ f) + \frac{1}{2} u(\xi) = \frac{1}{2} J(f) + \frac{1}{2} u(\xi)$$

and $J(f) = u(x) \;\Leftrightarrow\; J(g) = u(y)$, that is, $f \sim \overline{x} \;\Leftrightarrow\; g \sim \overline{y}$, as required.

To see that Certainty Covariance (A8) holds let f, g be acts and x, y consequences such that for all states s, $< f(s); g(s) > \sim^* < x; y >$. Then for all states s, $u(f(s)) - u(g(s)) = u(x) - u(y)$, which implies $I(u \circ f) + u(y) = I(u \circ g) + u(x)$ and thus $f \sim \overline{x} \;\Leftrightarrow\; g \sim \overline{y}$.

Finally, under MEU, sets of binary comonotonic acts are EU-sets, hence binary comonotonic tradeoff consistency is satisfied.

5.4 Proof of Theorem 2

By Theorem 11 in GM (2001), assuming A0 and A0*, the binary relation $\succsim$ satisfies axioms A1,A2,A3,A3*,A4 and A5* if and only if there exist a real valued, continuous, monotonic and nontrivial representation $J : \mathcal{F} \longrightarrow \mathbb{R}$ of $\succsim$, and a monotonic set function η on events, such that for binary acts with $x \succsim y$,

$$J(xEy) = u(x)\eta(E) + u(y)(1 - \eta(E)). \tag{8}$$

The function u is defined on X by, $u(x) = J(\overline{x})$, and it represents $\succsim$ on X. The set function η, when normalized s.t. $\eta(S) = 1$, is unique, and u and J are unique up to a positive multiplicative constant and an additive constant.

It remains to show that addition of axioms A6*,A7* and A8* yields an MEU representation as in (2) of the theorem. This is done simply by showing that when these axioms are translated into utiles space, they imply the exact same attributes as their counterpart axioms A6, A7 and A8. The key result is Proposition 1 of Ghirardato et al. (2003), which states that for biseparable preference relations, and $x \succ y \succ z$ in X, $y \in \mathcal{H}(x,z)$ if and only if $u(y) = u(x)/2 + u(z)/2$. Applying that result, we obtain equivalence between A6*, A7* and A8*, and their utilities-analogue axioms. First, axiom A6* is equivalent to Uncertainty Aversion in Utilities (A6u), as phrased above. Second, A7* is equivalent to the following utilities-based axiom:

A7u. Certainty Independence in Utilities:
Let the preference relation $\succsim$ be represented on constant acts by a utility function u. Suppose that two acts, f and g, and three consequences, x, y and w, satisfy: $u(y) = \frac{1}{2}u(x) + \frac{1}{2}u(w)$, and for all states s, $u(g(s)) = \frac{1}{2}u(f(s)) + \frac{1}{2}u(w)$. Then $f \sim x \;\Leftrightarrow\; g \sim y$.

Finally, A8* is equivalent to the utilities axiom:

A8u. Certainty Covariance in Utilities:
Let f, g be acts and x, y consequences such that $\forall s \in S$:
$\frac{1}{2}u(f(s)) + \frac{1}{2}u(y) = \frac{1}{2}u(g(s)) + \frac{1}{2}u(x)$. Then $f \sim \overline{x}$ iff $g \sim \overline{y}$.

Having these utilities axioms (which immediately apply to the entire acts space, as opposed to the tradeoffs case), the rest of the proof (both directions) is identical to the proof of Theorem 1.

5.5 Proof of Theorem 4

To obtain a CEU representation on simple acts we apply Corollary 10 from KW, stated below, supplemented with their subsection 5.1, which extends the theorem to infinite state spaces. As in the proof of Theorem 1, here as well KW use a slightly different tradeoff consistency condition than ours. However, assuming the rest of our axioms, their condition follows (see proof in Subsection 5.6).

Lemma 26. *(Corollary 10 of KW 2003, extended to an infinite state space in their subsection 5.1:) Given a binary relation $\succsim$ on $\mathcal{F}$, the following two statements are*

equivalent:

(1) $\succsim$ *satisfies:*

(A1) Weak Order

(A2) Continuity

(A4) Monotonicity

(A5.1) Comonotonic Tradeoff Consistency

(2) There exist a continuous utility function $u : X \to \mathbb{R}$ *and a nonadditive probability* v *on* Σ *such that, for all* $f, g \in \mathcal{F}$,

$$f \succsim g \quad \Leftrightarrow \quad \int_S u \circ f dv \geq \int_S u \circ g dv$$

Claim 27. *If, in addition to axioms A1,A2,A4 and A5.1, Essentiality is satisfied, then* u *is cardinal and* v *is unique and satisfies* $0 < v(E) < 1$ *for some event* E.

In the other direction, assuming that u *is cardinal, and that* v *is unique and satisfies* $0 < v(E) < 1$ *for some event* E*, Essentiality holds.*

Proof. Suppose that Essentiality is satisfied. By Observation 9 of KW, the utility function u is cardinal, and the nonadditive probability function v is unique. Applying the CEU representation obtains that there exists an event E for which $0 < v(E) < 1$.

Now assume that u is cardinal and v unique, satisfying that for some event E, $0 < v(E) < 1$. First, cardinality of u guarantees that there exist consequences x and y such that $x \succ y$. In addition, for the event E, $\overline{x} \succ xEy \succ \overline{y}$, thus Essentiality holds. ■

Let u be the utility function and v the non-additive probability obtained by Lemma 26. The proof of the theorem continues in subsections 5.5.1 and 5.5.2 below.

5.5.1 Proof of the implication (1)⇒(2) of Theorem 4.

It is left to prove that addition of Uncertainty Aversion yields the specific CEU form in (6) (along with the detailed attributes of v).

Lemma 28. *Let* $\succsim$ *satisfy (1) of Lemma 26. Then Uncertainty Aversion (A6) implies Uncertainty Aversion in Utilities (A6u).*

Proof. Let f, g, h be such that $f \succsim g$, and for all states t, $u(h(t)) = \frac{1}{2}u(f(t)) + \frac{1}{2}u(g(t))$. It is proved that $h \succsim g$.

Let x be an internal consequence, that is, $x^* \succ x \succ x_*$ for some x^*, x_*. Since X is connected and u continuous, there exist acts f', g', h' such that, in all states s,

$$\begin{aligned} u(f'(t)) &= \frac{1}{2^n}u(f(t)) + (1 - \frac{1}{2^n})u(x) \\ u(g'(t)) &= \frac{1}{2^n}u(g(t)) + (1 - \frac{1}{2^n})u(x) \\ u(h'(t)) &= \frac{1}{2^n}u(h(t)) + (1 - \frac{1}{2^n})u(x) = \frac{1}{2}u(f'(t)) + \frac{1}{2}u(g'(t)) \end{aligned}$$

By definition of the CEU functional, $f \succsim g \Leftrightarrow f' \succsim g'$ and $h \succsim g \Leftrightarrow h' \succsim g'$. Therefore it suffices to prove that $h' \succsim g'$.

In order to use Uncertainty Aversion we need to show that for every state t, $< f'(t); h'(t) > \sim^* < h'(t); g'(t) >$. That is, it should be proved that for every state t there are consequences y, z and an event E such that $f'(t)Ey \sim h'(t)Ez$ and $h'(t)Ey \sim g'(t)Ez$. It is next shown that these indifference relations may be satisfied for an event E such that $\overline{x} \succ xEy \succ \overline{y}$ whenever $x \succ y$ (exists by Essentiality and Claim 8), and f', g', h' are made 'close enough' to the internal consequence x.

Let E be an event such that $\overline{x} \succ xEy \succ \overline{y}$ whenever $x \succ y$. By the CEU representation it follows that $0 < v(E) < 1$. Let $z_1 \succ z_2 \succ z_3$ be internal consequences that satisfy

$$z_1 \succ x \succ z_2 \text{ and } u(z_1) - u(z_2) < \frac{2(1 - v(E))}{v(E)}(u(z_2) - u(z_3))$$

For any two consequences $a \succ c$ between z_1 and z_2, $u(a) - u(c) \leq u(z_1) - u(z_2)$. Thus, applying continuity of u, there are consequences y, z such that

$$z_2 \succ z \succ y \succ z_3 \text{ and } u(a) - u(c) = \frac{2[1 - v(E)]}{v(E)}(u(z) - u(y)).$$

Rearranging the expression and letting b be a consequence with $u(b) = \frac{u(a)+u(c)}{2}$, yields $aEy \sim bEz$ and $bEy \sim cEz$. Since all acts are comonotonic and E is comonotonically nonnull on the set containing them, we conclude that for every a, c

such that $z_1 \succ a \succ c \succ z_2$ there is a consequence b that satisfies $u(b) = \frac{u(a)+u(c)}{2}$ and $< a; b > \sim^* < b; c >$.

Taking a large enough n, for each t it may be asserted that $z_1 \succ f'(t), g'(t) \succ z_2$. As all acts obtain finitely many values there is a number n such that the preferences hold for all states simultaneously, thus $< f'(t); h'(t) > \sim^* < h'(t); g'(t) >$ for all states t. By Uncertainty Aversion $h' \succsim g'$ and the proof is completed. ■

Denote by B_0 the set of real-valued functions on S which assume finitely many values. Let $J : \mathcal{F} \to \mathbb{R}$ denote the representation of $\succsim$ over acts, that is, for all $f \in \mathcal{F}$, $J(f) = \int_S u \circ f dv$, and define by I the corresponding function over B_0, so that $I(a) = \int_S a dv$ for $a \in B_0$.

Claim 29. *Let $a, b \in B_0$ be such that $I(a) = I(b)$, then $I(a + b) \geq I(a) + I(b)$.*

Proof. By its definition, I is homogeneous, so it suffices to show that the claim is satisfied for all $a, b \in B_0 \cap (u(X))^S$. Let $a, b \in B_0 \cap (u(X))^S$ be such that $I(a) = I(b)$. Let $f, g \in \mathcal{F}$ be such that $u \circ f = a,\ u \circ g = b$. Then $J(f) = I(a) = I(b) = J(g)$, and $f \sim g$. By A6u it follows that if h is an act such that in all states s, $u(h(s)) = \frac{1}{2}u(f(s)) + \frac{1}{2}u(g(s))$, then $h \succsim g$. Thus, applying homogeneity once more,

$$J(h) = I(u \circ h) = I(\frac{1}{2}u \circ f + \frac{1}{2}u \circ g) = \frac{1}{2}I(a + b) \geq J(g) = I(b) = \frac{1}{2}(I(a) + I(b))$$

■

To obtain the specific CEU form as minimum expected utility over the core of v, a result from Schmeidler (1989) is stated.

Lemma 30. *(part of a proposition from Schmeidler 1989): Suppose that $\succsim$ on $\mathcal{F}$ is represented by a CEU functional $J(f) = I(u \circ f) = \int_S u \circ f dv$. The following conditions are equivalent:*

(i) For all $a, b \in B_0$, if $I(a) = I(b)$ then $I(a + b) \geq I(a) + I(b)$.

(ii) v is convex.

(ii) For all $a \in B_0$, $I(a) = \min\{\int_S a dP \mid P \in core(v)\}$.

Thus, by Lemma 29, the representation in (2) obtains. Essentiality guarantees that $0 < v(E) < 1$ for some event E.

5.5.2 Proof of the implication (2)⇒(1) of Theorem 4.

By Lemma 26, axioms A1, A2, A4 and A5.1 are satisfied. By claim 27, Essentiality holds. It remains to show that when the integral w.r.t. v takes on the special form of minimum expectation over a set of priors, Uncertainty Aversion (A6) is satisfied. However since $< f(s); h(s) > \sim^* < h(s); g(s) >$ implies $u(h(s)) = \frac{1}{2}[u(f(s)) + u(g(s))]$, and the representing functional consists of taking a minimum, the required inequality is easily seen to hold.

5.6 On the equivalence of tradeoff consistency definitions

KW use the following condition in eliciting a CEU representation:

KW Comonotonic Tradeoff Consistency:

Improving any outcome in the $\sim^*$ relation breaks that relation.

Our formulation of Comonotonic Tradeoff Consistency (A5.1), as well as Binary Comonotonic Tradeoff Consistency (BCTC, A5) applied to a state space with two states, correspond to an axiom KW call *Comonotonic Strong Indifference-Tradeoff Consistency*:

KW Comonotonic Strong Indifference-Tradeoff Consistency:

For any four consequences a, b, c, d, four acts f, g, f', g' and events D, E,

$$aDf \sim bDg, \quad cDf \sim dDg, \quad aEf' \sim bEg' \;\Rightarrow\; cEf' \sim dEg'$$

whenever the sets of acts $\{ aDf, bDg, cDf, dDg \}$ and $\{ aEf', bEg', cEf', dEg' \}$ are comonotonic, D is comonotonically nonnull on the first set, and E is comonotonically nonnull on the second.

We show that assuming the rest of our axioms, Comonotonic Tradeoff Consistency (equivalently, BCTC for $|S| = 2$) implies KW Comonotonic Tradeoff Consistency. Thus, we may use their results to derive a 6, as is done in Appendix A. To show their axiom is implied, we require an additional condition of monotonicity.

Comonotonic Strong Monotonicity:

For any two comonotonic acts f and g, if $f(s) \succsim g(s)$ for all states s, and $f(t) \succ g(t)$ for a state t that is comonotonically nonnull on $\{f, g\}$, then $f \succ g$.

Lemma 31. *Assume that the binary relation $\succsim$ satisfies Weak Order, Continuity, Essentiality, Monotonicity and Comonotonic Tradeoff Consistency. Then $\succsim$ satisfies Comonotonic Strong Monotonicity.*

Proof. Let f, g be comonotonic acts such that $f(s) \succsim g(s)$ for all states $s \in S$, and $f(t) \succ g(t)$ for some state $t \in S$ which is comonotonically nonnull on $\{f, g\}$. Let $\{E_1, \ldots, E_n\}$ be a partition w.r.t. which both f and g are measurable, and such that $f(E_1) \succsim f(E_2) \succsim \ldots f(E_n)$.[11] As f and g are comonotonic the same ordering holds for g. Define acts $h_0, h_1, \ldots, h_n$ as follows: $h_0 = g,\ h_2 = fE_1g,\ h_3 = f(E_1 \cup E_2)g,\ \ldots,\ h_{n-1} = f(E_1 \cup \ldots \cup E_{n-1})g,\ h_n = f$. All these acts are comonotonic, and by Monotonicity $h_n \succsim \ldots \succsim h_1 \succsim h_0$, and h_i, h_{i-1} differ by at most one consequence. Thus, it suffices to prove that for any two comonotonic acts aDh, bDh, with $a \succ b$ and D comonotonically nonnull on $\{aDh, bDh\}$, $aDh \succ bDh$.

Let E be an event for which $\alpha \succ \alpha E \beta \succ \beta$ whenever $\alpha \succ \beta$ (exists by Essentiality and Claim 8). We assume that $aDh \sim bDh$ and derive a contradiction. Obviously $aDh \sim aDh$ and $aEb \sim aEb$, therefore by Comonotonic Tradeoff Consistency (all comonotonicity and non-nullity conditions are satisfied) it must be that $aEb \sim \bar{b}$, contradicting the choice of E. Thus $aDh \succ bDh$ and Comonotonic Strong Monotonicity holds. ■

Having proved that $\succsim$ satisfies Comonotonic Strong Monotonicity, Lemma 24 of KW implies that it also satisfies KW Comonotonic Tradeoff Consistency [12].

REFERENCES

[11] By $f(E)$ for a set E it is meant $f(s)$ for some $s \in E$. Since f and g are measurable w.r.t. the partition $\{E_1, \ldots, E_n\}$ then for all i, $f(s) = f(t)$ and $g(s) = g(t)$ for all states $s, t \in E_i$.

[12] The lemma makes use of a finite state space and measures tradeoffs over single states, but, as explained in section 5.1 of KW, the exact same arguments work when the state space is infinite but an appropriate finite partition is considered.

Aczel, J. (1966), Lectures on functional equations and their applications. Academic Press, New York.

Alon, S. (2012), "Derivation of a cardinal utility through a weak tradeoff consistency requirement", manuscript.

Anscombe, F.J. and R.J. Aumann (1963), "A Definition of Subjective Probability", *The Annals of Mathematical Statistics*, 34, 199-205.

Blaschke, W. and G. Boll (1938), Geometric der Gewebe. Springer-Verlag, Berlin.

Casadesus-Masanell, R., P. Klibanoff and E. Ozdenoren (2000), "Maxmin expected utility over savage acts with a set of priors", *Journal of Economic Theory*, 92, 35-65.

Chateauneuf, A. (1991), "On the use of capacities in modeling uncertainty aversion and risk aversion ", *Journal of Mathematical Economics*, 20, 343-369.

Debreu, G. (1959), The Theory of Value: An axiomatic analysis of economic equilibrium. Yale University Press, CT.

Debreu, G. (1960), "Topological Methods in Cardinal Utility Theory." In Kenneth J. Arrow, Samuel Karlin, & Patrick Suppes (1960, Eds), Mathematical Methods in the Social Sciences, 16Uf02d26, Stanford University Press, Stanford, CA.

de Finetti, B. (1937), "La Provision: Ses Lois Logiques, ses Sources Subjectives," *Annales de l'Institut Henri Poincare* 7, 1-68. Translated into English by Henry E. Kyburg Jr., "Foresight: Its Logical Laws, its Subjective Sources," in Henry E. Kyburg Jr. & Howard E. Smokler (1964, Eds), Studies in Subjective Probability, 93-158, Wiley, New York; 2nd edition 1980, 53-118, Krieger, New York.

Fishburn, P. C. (1970), Utility Theory for Decision Making. Wiley, New York.

Ghirardato, P., P. Klibanoff and M. Marinacci (1998), "Additivity with Multiple Priors", *Journal of Mathematical Economics*, 30, 405-420.

Ghirardato, P., F. Maccheroni, M. Marinacci and M. Siniscalchi (2003), "A subjective spin on roulette wheels", *Econometrica*, 71, 1897-1908.

Ghirardato, P. and M. Marinacci (2001), "Risk, Ambiguity, and the Separation of Utility and Beliefs," Mathematics of Operations Research 26, 864-890.

Gilboa, I. (1987), "Expected Utility With Purely Subjective Non-Additive Probabilities", *Journal of Mathematical Economics*, 16, 65-88.

Gilboa, I. and D. Schmeidler (1989), "Maxmin expected utility with nonunique prior", *Journal of Mathematical Economics*, 18, 141-153.

Gilboa, I., F. Maccheroni, M. Marinacci and D. Schmeidler (2010), "Objective and Subjective Rationality in a Multiple Prior Model", *Econometrica*, Vol. 78, No. 2, 755–770.

Grant, S. and B. Polak (2011), "Mean-Dispersion Preferences and Constant Absolute

Uncertainty Aversion", Cowles Foundation Discussion Paper No. 1805.

Kobberling, V. and P. Wakker (2003), "Preference foundations for nonexpected utility: A generalized and simplified technique", *Mathematics of Operations Research*, 28, 395-423.

Kopylov, I. (2007), "Subjective probabilities on 'small' domains", Journal of Economic Theory, 133, 236-265.

Krantz, D.H., R.D. Luce, P. Suppes, and A. Tversky (1971), Foundations of Measurement. New York: Academic Press.

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

Maccheroni, F. M. Marinacci and A. Rustichini (2006), "Ambiguity aversion, robustness, and the variational representation of preferences", *Econometrica*, 74, 1447-1498.

von Neumann, J. and O. Morgenstern (1944, 2nd edition 1947, 1953), Theory of Games and Economic Behavior. Princeton University Press, Princeton NJ.

Ramsey, F. P. (1931), "Truth and Probability." In "The Foundations of Mathematics and other Logical Essays," 156-198, Routledge and Kegan Paul, London. Reprinted in Henry E. Kyburg Jr. & Howard E. Smokler (1964, Eds), Studies in Subjective Probability, 61-92, Wiley, New York. (2nd edition 1980, Krieger, New York.)

Savage, L.J. (1952, 1954), The Foundations of Statistics. Wiley, New York (2nd ed. 1972, Dover, New York).

Schmeidler, D. (1989), "Subjective probability and expected utility without additivity", *Econometrica*, 57, 571-587. (1984), CARESS WP #84-21, U. of PA.

Wakker, P. (1986), Representations of Choice Situations, Ph.D. Dissertation, University of Brabant, Department of Economics, Tilburg, the Netherlands.

Wakker, P. (1989), "Additive Representations of Preferences, A New Foundation of Decision Analysis", *Kluwer Academic Publishers, Dordrecht, The Netherlands.*

Wakker, P. (1989), "Continuous Subjective Expected Utility with Nonadditive Probabilities," Journal of Mathematical Economics 18, 1-27.

Data-Based Comparison of Likelihoods with Incompleteness*

Shiri Alon† and Gabrielle Gayer ‡

November 19, 2012

Abstract

This paper presents an axiomatic model of a process in which likelihoods of eventualities are compared based on a database. A central feature of this process is that the comparison is not required to be complete. The first step in the evaluation process is to establish the support provided by a record in the database to the likelihood of an eventuality. This support is determined by the relevance of a record to an eventuality. Then one eventuality is ranked above another when the total support it gains over the entire database is higher. However, the correct relevance of each record may be impossible to ascertain with any degree of surety. Therefore a set of relevance-weights is introduced, each element in this set corresponding to a possible interpretation of the data. A comparison can only be made when the rankings of eventualities established by the different relevance-weights, are in unanimous accord, otherwise the comparison is indecisive. The present model is a generalization of the Gilboa and Schmeidler inductive inference model (2003) which defines a complete ranking process over a database.

*The authors are grateful to Itzhak Gilboa, David Schmeidler and Peter Wakker for many helpful comments.

†School of Mathematical Sciences, Tel Aviv University, Tel Aviv 69978, Israel. e-mail: shiri.aloneron@gmail.com

‡Department of Economics, Bar-Ilan University, Ramat Gan, 52900, Israel. e-mail: gabi.gayer@gmail.com

Keywords: Incompleteness, Inductive Inference, Case-Based Decision Theory, Likelihood comparisons.

JEL classification: D80

1 Introduction

1.1 Motivation

There are many instances where people are faced with problems in which they have to compare the likelihood of several eventualities. These problems can range from predicting the weather to the chances of survival of the Eurozone or the achievement of peace in the Middle East. These predictions have obvious implications for market behavior as well. How do people make these evaluations? What is their process of performing comparisons? Are they always able to conclude that one eventuality is more likely than another? Understanding this evaluation mechanism may help us both to identify the circumstances under which such comparisons are possible as well as to predict their outcomes.

The present work, focusing on evaluation methods based on past observations, rests upon the widely accepted principle that similar circumstances lead to similar outcomes. One eventuality is considered to be more likely than another if the data support the occurrence of the former more than they do the latter. Take for example an individual trying to predict which of two teams is more likely to win a match. In the simple case where the teams played each other many times and Team A won a large majority of games, an evaluator will most surely conclude that Team A is more likely to win the current match. On the other hand, if the only available observations are the results of matches in which teams A and B played other teams, the relevance of these observations to the problem at hand is questionable. In this case it would be much harder to ascertain which team is more likely to win.

A more serious case of this type of quandary is that of a doctor who is asked to recommend whether a patient with a certain medical condition should undergo a treatment that, while it may cure the patient's illness, may also produce complications. The doctor will undoubtedly base the recommendation on the outcomes of relevant past cases. In particular, the doctor will tend to recommend treatment where success occurred much more frequently than failure, and oppose treatment if the opposite were true. The question is what will the doctor recommend if he or she is uncertain as

to the relevance of these past cases to the condition of the current patient. In such a situation the doctor will feel much less confident about recommending any specific option, and will only outline to the patient the advantages and disadvantages of each course of action.

In fact, it appears that the US Preventive Services Task Force (USPSTF) relies upon a similar approach when issuing guidelines for preventive health care services which are broken down by age, gender, medical condition and other criteria. However, the USPSTF often publishes 'I statements' where no recommendation is made, when it "concludes that the current evidence is insufficient to assess the balance of benefits and harms of the service". [1] In such cases the USPSTF advises that patients be informed about the uncertain balance of benefits and harms, stating that its recommendation may be altered after gathering more specific observations. The use of 'I statements' implies that uncertainty regarding the relevance of existing data is one of the primary causes for not making recommendations.[2]

The evaluation process described above contains two major components: the first that the comparison of eventualities depends on support of past observations that may change as a function of the available data, and the second that the evaluator may refrain from making a prediction due to the existence of several plausible methods to analyze or interpret the data. These different interpretations may indeed lead to diametrically opposite conclusions which would render impossible any comparison of the likelihoods of the eventualities under consideration.

In the context of decision-making, the inevitability of making a decision is the main critique of incompleteness. In fact, it is commonly held that not making a decision is in itself also a decision. However, from both positive and normative perspectives this criticism is invalid in connection to comparisons of likelihoods, for not only is it impossible to force an individual to make an honest prediction, it may be undesirable as well. Surely in the above medical example a patient would prefer to be told that the doctor could not assess whether success is more likely than failure, rather than be

[1] See the USPSTF website, http://www.uspreventiveservicestaskforce.org.

[2] We thank Irit Dror for bringing the example of USPSTF to our attention.

misled by an arbitrary, poorly-based prediction.

There are many situations in which there is an individual who makes a decision based on evaluations of likelihoods of a separate party, this medical example being only one instance. In such a case it seems more appropriate to allow for incomparability of likelihoods instead of insisting on completeness, for the decision maker may entertain various considerations and constraints, of which comparison of likelihoods is but one element. In particular, when the comparison of likelihoods is inconclusive the decision maker may want to put more weight on the other considerations. Thus, if the evaluator were to find it impossible to compare some likelihoods, then such difficulty is in itself informative. The decision maker would typically prefer that this predicament be revealed rather than have it papered over by an unfounded conclusion.

The purpose of this essay is to model evaluation processes in which data-based likelihoods are compared, but which also allow for non-determinability. The model aims to describe conditions in which a database is utilized to compare eventualities by evaluating the support that the database gives to their likelihoods of occurrence. When several plausible interpretations of the data lead to opposite conclusions, inability to make a valid comparison obtains.

1.2 The Model

According to the Gilboa and Schmeidler model of inductive inference (2003; henceforth abbreviated GS) a weight is assigned to each record in the database expressing its degree of support for the eventuality under consideration. The evaluation of an eventuality simply becomes its aggregate support. One eventuality is ranked above another if its aggregate support across the database is higher (the inductive inference model is closely related to Case-Based Decision theory, 1995 and 1997). For instance, when considering surgery for a patient, a past case in which surgery was successful lends more support to the eventuality 'success' than to the eventuality 'failure', this support being higher the more resemblance there is between the medical parameters of the patients in the past and current cases. 'Success' would be considered more likely than 'failure' if its total support across the database were higher.

Ranking eventualities according to the GS model requires the capability of assigning an *exact* weight, expressing degree of support, to each pair of eventuality and record in the database. This is a reasonable assumption when a good understanding exists of the factors responsible for relevance of past cases to the current situation. However, when the situation at hand is not completely understood or perhaps when there is a lot at stake, the evaluator may be reluctant to commit to specific relevance-weights. For instance, the doctor in the medical example above may not be able to determine whether the cases of two patients in the same age group but with different blood pressure levels resemble each other more than the cases of two patients from different age groups but with identical blood pressure levels.

The present work generalizes the GS model by allowing the evaluator to be unsure about the relevance of records in the database to the eventuality under consideration and therefore allows a *set* of relevance-weights that correspond to different interpretations of the data. In the model one eventuality is considered more likely than another, if and only if, it gains more support according to every interpretation of the data. If conflicting interpretations of the data obtain, that is when an eventuality is evaluated as more likely than another according to one relevance-weighting, but according to another relevance-weighting the opposite is true, then the evaluator will refrain from issuing a comparison of the likelihoods of the two eventualities. This model thus converts lack of confidence into a sign of caution.

Specifically, relevance-weighting function v assigns a weight to each pair of an eventuality and a record in a database. For two eventualities, x and y, and a database I, x is considered more likely than y, given database I, if and only if, for *every* relevance-weighting function v in some set V,

$$\sum_{c \in I} v(x, c) \geq \sum_{c \in I} v(y, c) \tag{1}$$

where c denotes a record in the database, and $v(x, c)$ is the support lent by record c to eventuality x, according to relevance-weighting function v. In other words, x is more likely than y, given database I, if and only if, there is unanimity over V that I lends more support to eventuality x than to eventuality y.

In this paper three axioms are imposed in order to derive the desired representation. In comparison to the GS model, completeness is suppressed, while GS's Combination axiom is replaced by a stronger axiom which we have termed Independence of Support. Both Combination and Independence of Support rely on the rationale that the support a certain record in the database lends to an eventuality is the same, regardless of the eventuality to which a comparison is made, and regardless of the other records in the database. Combination is a basic expression of this notion, whereas Independence of Support abstracts it even further. As the relations in this paper may be incomplete, the Independence of Support axiom not only requires that this notion of independence not be contradicted, but it also entails that comparisons motivated by this consideration be actively completed.

In a complementary result, our axioms are supplemented by Completeness in order to obtain a representation as in (1) with a single function v.[3] The strengthening of the Combination axiom allows this model to dispense with GS's Diversity axiom thus enabling the removal of the restrictions on v that appear in the GS representation. On the other hand, without the Diversity axiom, uniqueness is no longer guaranteed.

Independence of Support is closely connected to the finite cancellation condition which was first introduced by Kraft, Pratt and Seidenberg (1959) to derive a representation of a subjective probability. Later, Scott (1964), Kranz et al. (1971) and Narens (1974) stated other variants of this condition (for a more detailed discussion of the finite cancellation condition, see Alon and Lehrer, 2012). The primitives over which the finite cancellation condition is defined are events rather than databases as in our Independence of Support condition. Nevertheless, both conditions impose strong additivity requirements in order to derive the desired representation.

1.3 Related Literature

There is a vast body of literature within decision theory providing axiomatizations of relations that, due to ambiguity are allowed to be incomplete. Ambiguity in these

[3] To be precise, we require also a strict version of Independence of Support, see Subsection 3.2.

models is represented by a decision maker having a set of priors over the states of nature, which is analogous to our evaluator considering a set of relevance-weights of past records to eventualities. The decision maker prefers one alternative over another, if and only if, this alternative is preferred according to each prior in the set of priors (see, for instance, Giron and Rios (1980), Bewley (2002), Gilboa et al. (2010), and Galaabaatar and Karni (2011)). However, in these works the preferences do not explicitly depend on data. By contrast, the ability of the evaluator in the present model to rank the likelihoods of two eventualities is a function of the available data, which is a primitive of the model.

Ambiguity was introduced into the Case-Based framework by Eichberger and Guerdjikova (2010 and 2011), first in the context of belief formation and then in the context of decision-making. The decision-making model suggests an α-maxmin rule which is complete. Eichberger and Guerdjikova (2010 and 2011) refer to two types of ambiguity: the first, due to the small number of observations in the database, where absence of data gives rise to the largest amount of ambiguity that gradually disappears as the number of observations grows. As will be shown in Section 2.2, this aspect is not part of the present model, for when no data are available, it is possible to compare any pair of eventualities and all eventualities are considered to be equally likely. This impartiality property is consistent with the aim of the present work which is to model an evaluation process that is purely data-based. Thus, this model expresses no *a priori* bias towards any one of the eventualities, and further restricts incomparability of likelihoods only to contradictory interpretations of available data. The second source of ambiguity mentioned in Eichberger and Guerdjikova (2010 and 2011) is caused by uncertainty regarding the relevance of past observations to the current problem. This second type is in line with the reasoning considered here and does not disappear with the accumulation of data.

1.4 Outline of the Paper

The setup and assumptions are discussed in Section 2. Section 3 contains the main theorem of the paper, a unanimity representation of incomplete relations, followed by a representation when relations are assumed to be complete. Section 4 expands the applications to which this model may be relevant, going beyond the context of evaluation and comparison of likelihoods. Section 5 contains the proofs.

2 Setup and Assumptions

2.1 Setup and Notation

$\mathbb{X}$	a finite, non-empty set of eventualities, with typical elements $x, y, \ldots$.
$\mathbb{C}$	a finite, non-empty set of record types, with a typical element c. This is the set of all possible classes of past observations that can be found in a database.
$\mathbb{J} = Z_+^{\mathbb{C}}$	the set of all databases, which are functions from record types to nonnegative integers, with typical elements $I, J, \ldots$. For $I \in \mathbb{J}$ and $c \in \mathbb{C}$, $I(c)$ denotes the number of times record type c appears in database I.
$\{\succsim_I\}_{I \in \mathbb{J}}$	for a database I, $\succsim_I$ is a binary relation over $\mathbb{X}$. Relationship $x \succsim_I y$ means that eventuality x is at least as likely as eventuality y, given database I.

In the medical example above the eventualities are 'success' and 'failure', and a record type consists of various medical parameters and the outcome of treatment. For instance, for parameters age ('young' or 'old') and blood pressure ('high' or 'low'), the record type ('young', 'low','success') represents a case of a young patient with low blood pressure whose treatment was successful, while the record type ('old', 'high', 'failure') represents a case of an old patient with high blood pressure whose treatment was unsuccessful. A database is a summary of the actual number of times the doctor has

encountered each record type. Given a database the doctor tries to evaluate whether success is more likely than failure. The relationship 'success' $\succsim_I$ 'failure' implies that the doctor maintains that success is at least as likely as failure given database I and vice versa for 'failure' $\succsim_I$ 'success'. Lack of both 'success' $\succsim_I$ 'failure' and 'failure' $\succsim_I$ 'success' indicates that the doctor is unable to assess which of the two eventualities is more likely.

For two databases I and J, $I + J$ is the database obtained by pointwise addition, namely, $(I + J)(c) = I(c) + J(c)$ for every $c \in \mathbb{C}$.

2.2 Assumptions

First and foremost, the model does not impose completeness of relations $\{\succsim_I\}_{I \in \mathbb{J}}$. Without completeness weak relations allow for distinguishability between pairs of eventualities that are equally likely and those that, according to our model, cannot be compared. This distinction cannot be made when the relations are strict, thus weak relations are more appropriate for our framework. The model assumes reflexivity of relations in order to establish that they are indeed weak.

A1. Reflexivity: For every eventuality x and database $I \in \mathbb{J}$, $x \succsim_I x$.

2.2.1 *Combination and Related Axioms*

Next we introduce and discuss a few assumptions that eventually lead to the set of necessary and sufficient axioms employed in our model. A basic assumption presented first is Transitivity, which states that for every three eventualities x, y, z and any database I, if both $x \succsim_I y$ and $y \succsim_I z$ then $x \succsim_I z$. Another assumption, central to the GS model, is Combination, which states that if eventuality x is evaluated as more likely than eventuality y given two separate databases I and J, then it should also be evaluated as more likely than y when the two databases are combined. More formally, for every two eventualities x, y and databases I and J, if $x \succsim_I y$ and $x \succsim_J y$ then $x \succsim_{I+J} y$. Combination further states that if, in addition, either $\neg(y \succsim_I x)$ or

$\neg(y \succsim_J x)$, then $\neg(y \succsim_{I+J} x)$ (which translates to strict relationships if completeness is assumed). The assumption of Combination flows logically from the conception of a database as the sum of its components. Therefore, the relevance of a record to an eventuality cannot be altered by other records in the database. Combination requires that the support for an eventuality provided by any one record in the database be independent of all the other records in that database.

For a better understanding of this axiom consider an example, taken from GS, where it fails. Suppose an individual tries to compare the likelihoods of a coin being fair or biased using a database of past tosses. The support that a record of tails provides to the likelihood that the coin is biased is higher when the other records in the database are mainly tails than it is when the database contains a nearly equal number of 'heads' and 'tails' records. The support that a 'tails' record lends to the likelihood of 'biased coin' therefore depends upon the other records in the database (see GS for a detailed discussion of this point and more examples).

Under the assumption of completeness, Combination implies several other properties that do not follow when completeness is suppressed. As will be made clear, these properties match the type of ambiguity addressed in our model and are consistent with the idea of independence of support across records. Therefore it seems sensible that these properties continue to hold in our incomplete framework as well.

For the special case of $I = J$, Combination implies that if $x \succsim_I y$, then $x \succsim_{2I} y$. Combination and completeness together also imply the opposite, namely if $x \succsim_{2I} y$, then $x \succsim_I y$, or, equivalently, if $\neg(x \succsim_I y)$, then $\neg(x \succsim_{2I} y)$. Without completeness this implication does not necessarily follow. In our model the inability to make comparisons is caused by conflicting interpretations of the data. In particular, when $\neg(x \succsim_I y)$ there must be one interpretation of the records in I whereby x is evaluated as more likely than y, and another interpretation of the records in I whereby the opposite evaluation occurs. Independence of support across records means that the duplicated database, $2I$, is simply the sum of two identical copies of I. Hence, the same conflicting interpretations that apply for I apply for $2I$ as well, yielding $\neg(x \succsim_{2I} y)$. The next axiom is the applications of this argument to any number of replications of the same

database.

A2. Persistent Incomparability: For any two eventualities x and y, database $I \in \mathbb{J}$ and an integer $n \in N$, if $\neg(x \succsim_I y)$ then $\neg(x \succsim_{nI} y)$.

In the presence of completeness, Combination also entails that if $x \succsim_{I+J} y$ and $y \succsim_I x$ then $x \succsim_J y$. However, when completeness is not imposed, inability to compare x with y on J is still consistent with Combination. Nevertheless, the requirement that $x \succsim_J y$ once again conforms to independence of evaluations across records, as the support that records in I lend to the likelihood of y compared with that of x cannot be altered by records in J. Since y is perceived to be more likely than x on I, if this evaluation is reversed on $I + J$, then it must be the records in J that caused this reversal by supporting x more than y. The following axiom incorporates this idea.

Strong Combination: For any two eventualities x and y and databases I and J,

If $x \succsim_I y$ and $x \succsim_J y$, then $x \succsim_{I+J} y$.

If $x \succsim_{I+J} y$ and $y \succsim_I x$, then $x \succsim_J y$.

2.2.2 *Insufficiency of the Above Axioms*

Reflexivity, Transitivity, Persistent Incomparability, and Strong Combination are obviously necessary conditions for the representation (1), however, they are insufficient, as demonstrated by the following example.

Example 1. Let $\mathbb{X} = \{x, y, z, w\}$ and $\mathbb{C} = \{1, 2\}$. Suppose that the relations $\succsim_I$ are all reflexive, and that

$$\begin{array}{lllll} x \succsim_{(1,0)} y & , & x \succsim_{(1,1)} z & , & x \succsim_{(0,1)} w \quad , \\ z \succsim_{(1,0)} w & , & w \succsim_{(1,1)} y & , & z \succsim_{(0,1)} y \quad , \text{ and} \\ & & \neg(x \succsim_{(2,2)} y) & . & \end{array}$$

[4] For instance, $(1, 0)$ denotes the database which contains one record of type '1' and no records of type '2'.

It is easy to see that Transitivity, Persistent Incomparability, and Strong Combination are all satisfied,[5] yet these relationships cannot be represented by (1). Assuming there exists such a representation, the relationships above imply that for all $v \in V$,

$$\begin{aligned} v(x,1) &\geq v(y,1) \ , \\ v(z,1) &\geq v(w,1) \ , \\ v(x,1)+v(x,2) &\geq v(z,1)+v(z,2) \ , \\ v(w,1)+v(w,2) &\geq v(y,1)+v(y,2) \ , \\ v(x,2) &\geq v(w,2) \ , \text{ and} \\ v(z,2) &\geq v(y,2) \ . \end{aligned}$$

By summing over these inequalities $2v(x,1)+2v(x,2) \geq 2v(y,1)+2v(y,2)$ for all $v \in V$. Hence by the representation $x \succsim_{(2,2)} y$, contradicting $\neg(x \succsim_{(2,2)} y)$.

Example 1 not only demonstrates that the axioms discussed above are too weak to derive the desired representation, but also raises the question of whether it is reasonable to require that a stronger condition hold. In particular we need to examine whether, given the impetus of the model, the relationships stated in the example should in fact imply that $x \succsim_{(2,2)} y$. We argue the affirmative. On database $(1,1)$ it is possible to complete the relationship between z and w in two ways, either by $z \succsim_{(1,1)} w$ or by $w \succsim_{(1,1)} z$. In case $z \succsim_{(1,1)} w$, Transitivity implies $x \succsim_{(1,1)} y$, and by Combination it follows that $x \succsim_{(2,2)} y$. Otherwise, if $w \succsim_{(1,1)} z$, then by Strong Combination $w \succsim_{(0,1)} z$ leading to $x \succsim_{(0,1)} y$ by Transitivity. Then by applying Combination twice, both $x \succsim_{(1,1)} y$ and $x \succsim_{(2,2)} y$. The ranking of eventualities z and w on $(1,1)$ may not be known, yet if all possible rankings lead to the same conclusion that $x \succsim_{(2,2)} y$ then this relationship should be completed in that manner.

2.2.3 *Independence of Support : A Further Strengthening of the Above Axioms*

The previous discussion suggests that the strengthening of the GS axioms introduced above is in order. To obtain a representation in the GS model, the above conditions

[5] For this purpose assume that relationships on replications of the databases mentioned above are properly completed according to Combination. Moreover, any form of continuity may be added.

are supplemented with a richness assumption termed Diversity (along with a form of Continuity). Diversity states that every four eventualities can be ranked in any order given some database. Gilboa and Schmeidler show that with this set of axioms the relations are represented by (1) with a single weighting function v. The richness implied by Diversity constrains the representing v (a further discussion on Diversity appears in Subsection 3.2). Diversity is a strong assumption in our context, since it requires the evaluator to be able to determine many comparisons, and thus confines the type of incomplete relations that can be described. This work therefore strengthens Strong Combination instead of imposing Diversity.

In this model the strengthening of Strong Combination imposes that the support that a record provides to the likelihood of an eventuality is independent of the other records and the other eventualities. The impetus for independence across records is the same as that mentioned above in the discussion of Combination, where the support provided for the likelihood of an eventuality by a specific record is not altered by the other records in the database.

Independence across eventualities means that the support that a record lends to the likelihood of eventuality x is independent of the eventuality to which it is being compared. Hence eventualities can be evaluated separately from one another. Such separability would fail if in determining the support that a record lends to the likelihood of an eventuality x, different criteria were employed depending on the eventuality y to which x is being compared. For example, consider an investor who is comparing several investments based on past performance of these and other similar investments. Suppose that when comparing the likelihood of profit from a European equity index to that from a US equity index, the investor considers past performances of all indices in both continents, but when comparing the likelihood of profit from a European equity index to the likelihood of profit from a European bond, only previous European investments are considered. In that case, the support a past investment lends to the eventuality 'profit from a European equity index' depends on the eventuality to which it is being compared. Such a process of comparison is ruled out in the present model. The connection between these notions of independence and the next axiom will now be

clarified.

Notice that in Example 1, the relationships satisfy the following conditions: the aggregation of all records on which x is more likely than the other eventualities (database $K = (2, 2)$) equals the aggregation of all records on which y is less likely than the other eventualities. Moreover, for any other alternative, the aggregation of all records on which it is more likely is the same as the aggregation of records on which it is less likely. When a set of relationships satisfies such conditions we say that it satisfies *conditions of support* of eventuality x over eventuality y on database K.

For a general formulation, let $\mathcal{R} = \{b_1 \succsim_{J_1} w_1, \ldots, b_n \succsim_{J_n} w_n\}$ denote a set of relationships, and define, for eventuality z, $\mathcal{M}_{\mathcal{R}}(z) = \sum_{l:b_l=z} J_l$ and $\mathcal{L}_{\mathcal{R}}(z) = \sum_{l:w_l=z} J_l$. That is, $\mathcal{M}_{\mathcal{R}}(z)$ aggregates all records in $\mathcal{R}$ for which z is considered more likely, whereas $\mathcal{L}_{\mathcal{R}}(z)$ aggregates those records in $\mathcal{R}$ for which it is considered less likely. Both $\mathcal{M}_{\mathcal{R}}(z)$ and $\mathcal{L}_{\mathcal{R}}(z)$ are themselves databases. The next definition applies this notation in order to obtain a general formulation for relationships that satisfy conditions of support for one eventuality over another for a given database.

Definition 1. *A set of relationships $\mathcal{R}$ is said to satisfy conditions of support of eventuality x over eventuality y on database K, if:*

$$\mathcal{M}_{\mathcal{R}}(x) = \mathcal{L}_{\mathcal{R}}(x) + K, \tag{2}$$

$$\mathcal{L}_{\mathcal{R}}(y) = \mathcal{M}_{\mathcal{R}}(y) + K, \tag{3}$$

$$\mathcal{M}_{\mathcal{R}}(z) = \mathcal{L}_{\mathcal{R}}(z), \quad \textit{for every } z \neq x, y\ . \tag{4}$$

A necessary condition for representation (1) is that when a set of relationships satisfies conditions of support of eventuality x over eventuality y on database K, then $x \succsim_K y$. This may easily be verified by summing over all the relationships in $\mathcal{R}$, in the same manner as in Example 1. This condition is next stated as an axiom.

A3-E. Exact Independence of Support: Let x, y be eventualities and K a database. If there exists a set of relationships that satisfies conditions of support of x over y on K, then $x \succsim_K y$.

Exact Independence of Support implies Transitivity and Strong Combination.[6] The following discussion explains the concepts supporting this axiom and elucidates how they comply with the two notions of independence across both records and eventualities discussed above.

By equation (2), the aggregated database on which x is considered **more** likely than other eventualities in $\mathcal{R}$, $\mathcal{M}_{\mathcal{R}}(x)$, is the sum of K and the database on which x is considered **less** likely than other eventualities in $\mathcal{R}$, $\mathcal{L}_{\mathcal{R}}(x)$. Following the logic that underlies Strong Combination, the records in K must be responsible for the evaluation of x as more likely. It thus follows that K contains evidence to support the evaluation of x as more likely relative to the other eventualities in $\mathcal{R}$, while equation (3) upholds this by providing evidence in K to support the evaluation of y as less likely relative to the other eventualities in $\mathcal{R}$.

Nevertheless, it is undesirable to conclude that x should be considered more likely than y on K based on equations (2) and (3) alone. This is clearly demonstrated by the situation in which for some database I there is a most likely eventuality, z^*, and a least likely eventuality, z_* (so that $z^* \succsim_I x \succsim_I z_*$ for all eventualities x). If only conditions (2) and (3) were to be imposed, then by $x \succsim_I z_*$ and $z^* \succsim_I y$, every two eventualities x, y (apart from z^* and z_*) would have to be evaluated as equally likely on database I. The problem is that x and y are compared to eventualities that are not neutral. If, on the other hand, the likelihood assessments of x and y would rely on comparisons to eventualities that are neutral in a proper sense, then these could serve as a common reference point between the likelihood evaluation of x on K and the likelihood evaluation of y on K. The conclusion that x is more likely than y on K would then be natural (akin to Transitivity).

Conditions (4) of the definition imply that all eventualities in $\mathcal{R}$ other than x and y gain overall neutral support, as each has an equal amount of support both for and against it from each record type in the databases involved. The neutrality of these eventualities suggests that they can be employed as a common reference point. In this case, since K contains evidence that x is more likely compared to these neutral

[6] An explicit statement and proof appear below in Remark 1.

eventualities, while y is less likely compared to them, it is sensible to consider x to be more likely than y, given K, as Exact Independence of Support requires.

The axiom makes use of the two notions of independence, across both records and eventualities, in the interpretation of $\mathcal{M}_{\mathcal{R}}(\cdot)$ and $\mathcal{L}_{\mathcal{R}}(\cdot)$. Evidence that an eventuality is more likely, as well as evidence that it is less likely, is aggregated across databases, in order to obtain both the 'more likely' and the 'less likely' databases for the eventuality, $\mathcal{M}_{\mathcal{R}}(\cdot)$ and $\mathcal{L}_{\mathcal{R}}(\cdot)$. The resulting databases are independent of the order and manner in which they were aggregated, thus expressing independence of support across records. In addition, $\mathcal{M}_{\mathcal{R}}(\cdot)$ and $\mathcal{L}_{\mathcal{R}}(\cdot)$ are formed on the basis of comparisons of a particular eventuality to various other eventualities, reflecting the idea that the support a record lends to the evaluation of an eventuality is independent of the eventualities to which it is being compared.

Finally Exact Independence of Support is generalized to express the requirement that independence of support across records and eventualities be a continuous property. Namely if there are relationships that adequately approximate conditions of support of x over y on K, then again it must be that $x \succsim_K y$. The following definition states those conditions of ε-approximate support. For the definition recall that for a set of relationships $\mathcal{R}$ and eventuality z, $\mathcal{M}_{\mathcal{R}}(z)(c)$ and $\mathcal{L}_{\mathcal{R}}(z)(c)$ denote the number of times record type c appears in the databases $\mathcal{M}_{\mathcal{R}}(z)$ and $\mathcal{L}_{\mathcal{R}}(z)$, respectively.

Definition 2. *For $\varepsilon > 0$, a set of relationships $\mathcal{R}$ is said to ε-satisfy conditions of support of eventuality x over eventuality y on database K, if there exists $m \in N$ such that for every $c \in \mathbb{C}$,*

$$\begin{aligned} |\mathcal{M}_{\mathcal{R}}(x)(c) - \mathcal{L}_{\mathcal{R}}(x)(c) - mK(c)| &< m\varepsilon \\ |\mathcal{L}_{\mathcal{R}}(y)(c) - \mathcal{M}_{\mathcal{R}}(y)(c) - mK(c)| &< m\varepsilon \\ |\mathcal{M}_{\mathcal{R}}(z)(c) - \mathcal{L}_{\mathcal{R}}(z)(c)| &< m\varepsilon, \quad \textit{for every } z \neq x, y. \end{aligned}$$

These conditions are meant to capture ε-closeness of $\mathcal{M}_{\mathcal{R}}(x) - \mathcal{L}_{\mathcal{R}}(x)$ and $\mathcal{L}_{\mathcal{R}}(y) - \mathcal{M}_{\mathcal{R}}(y)$ to K, and of $\mathcal{M}_{\mathcal{R}}(z) - \mathcal{L}_{\mathcal{R}}(z)$ to the all-zeros database. The approximated conditions of support are formulated such that the size of the neighborhood around the approximated databases is proportional to the number of replications of this database

(i.e., neighborhood $m\varepsilon$ is proportional to the number of replications of K and the zeros database). This is a natural restriction since a given difference in the number of records is much more substantial when considering databases with a small number of records as compared to databases with a large number of records.

The continuous version of Exact Independence of Support may now be stated.

A3. Independence of Support: Let x, y be eventualities and K a database. If for every $\varepsilon > 0$ there exists a set of relationships that ε-satisfies conditions of support of x over y on K, then $x \succsim_K y$.

The axioms that are finally imposed in the model are Reflexivity, Persistent Incomparability, and Independence of Support. Independence of Support and Reflexivity entail that on an empty database (the all-zeros database) all eventualities are comparable and deemed equally likely (this attribute is termed *Impartiality*; See Remark 1 below). Together with Persistent Incomparability, this confines inability to compare eventualities only to situations created by conflicting interpretations of available data. This and other attributes that follow from these three axioms are summarized in the next remark.

Remark 1. For every three eventualities x, y, z and databases I and J, Reflexivity, Persistent Incomparability and Independence of Support imply:

(a) Transitivity: If $x \succsim_I y$ and $y \succsim_I z$, then $x \succsim_I z$.

(b) Combination: If $x \succsim_I y$ and $x \succsim_J y$, then $x \succsim_{I+J} y$.

(c) Homogeneity: $x \succsim_I y$ if and only if $x \succsim_{nI} y$.

(d) Impartiality: $x \succsim_0 y$, where 0 denotes the all-zeros database.

3 Results

3.1 Main Result

The conclusion of this inquiry is that assumptions A1, A2 and A3 are equivalent to the representation of the GS model with multiple relevance-weighting functions. As the range of the relevance-weighting functions is $\mathbb{X} \times \mathbb{C}$, they are referred to as matrices throughout the remainder of the paper.

Theorem 1. *The following statements are equivalent:*

(i) The relations $\{\succsim_I\}_{I \in \mathbb{J}}$ satisfy assumptions A1, A2 and A3.

(ii) There exists a nonempty set V of matrices $v : \mathbb{X} \times \mathbb{C} \longrightarrow R$, such that for any two eventualities x, y and database I,

$$x \succsim_I y \quad \Leftrightarrow \quad \sum_{c \in \mathbb{C}} v(x,c)I(c) \geq \sum_{c \in \mathbb{C}} v(y,c)I(c), \quad \textit{for every } v \in V \ . \tag{5}$$

With this representation each matrix $v \in V$ can be viewed as a conceivable interpretation of the data. The likelihood of eventuality x being greater than y is accepted only when every possible interpretation of the data leads to that conclusion. If, however, x is evaluated as more likely than y according to one interpretation, and less likely according to another interpretation, then no ranking of the likelihoods of the two eventualities is possible.

In order to illustrate the evaluation process described by our theorem let us return to the medical example presented in the introduction. Assume that the doctor has the following set of data on a particular treatment:

Age	Blood Pressure	Outcome	Num. of Records
young	low	success	150
young	high	success	4
old	low	success	60
old	high	success	5
young	low	failure	40
young	high	failure	0
old	low	failure	30
old	high	failure	40

Now suppose a young person with high blood pressure consults the doctor about undergoing this treatment. The doctor is generally in favor of treatment as the overall success rate in the entire population is approximately 2/3. However, the rate varies dramatically within different subgroups of the population, therefore the degree of relevance of these past cases to this specific patient is crucial.

In this situation there are only four cases identical to the current case. As this number is so small, the doctor will have to take into account other cases in order to evaluate whether success is more likely than failure. If the doctor were to consider only past cases within the patient's age group ('young'), he or she will conclude that success is more likely, as it occurs much more frequently than failure in that group. This corresponds to assigning, for $x \in \{success, failure\}$, $v(x, c) = 1$ when $c = (young, \cdot, x)$ and $v(x, c) = 0$ otherwise. If, instead, the doctor were to consider only past cases with high blood pressure, the doctor would come to the opposite conclusion, as failure within this group is much more frequent than success. This corresponds to assigning, for $x \in \{success, failure\}$, $v(x, c) = 1$ when $c = (\cdot, high, x)$ and $v(x, c) = 0$ otherwise.

It is precisely this contradiction in the conclusions that makes the doctor unsure whether treatment is more likely to succeed than fail in the current case. It should be noted that such uncertainty does not exist when considering treatment for a young patient with low blood pressure, since both the age and the blood pressure level indicate that success is more likely than failure in this case.

3.1.1 *Uniqueness of the Representing set V*

The set of matrices V in the theorem need not be unique. First, if V represents the relations $\{\succsim_I\}_{I\in\mathbb{J}}$, then so does the closed convex cone generated by V, and second it is possible to shift the matrices in V by adding any matrix with identical rows (i.e. a matrix with constant columns) and still obtain a representation of the same relations. These issues can be resolved by normalizing some row of each matrix in V, for example, by fixing the last row in each matrix to equal an all-zeros row. Yet, even if all matrices in V are normalized, and only closed convex cones are considered, uniqueness is still not guaranteed. The following example demonstrates this point.

Example 2. Let $\mathbb{X} = \mathbb{C} = \{1, 2, 3\}$, and define V to be the closed convex cone generated by the following three matrices:

$$v_1 = \begin{pmatrix} 3 & 0 & -1 \\ 0 & 1 & -1 \\ 0 & 0 & 0 \end{pmatrix} \quad , \quad v_2 = \begin{pmatrix} 1 & 0 & -1 \\ 0 & 3 & -1 \\ 0 & 0 & 0 \end{pmatrix} \quad , \quad v_3 = \begin{pmatrix} 1 & 0 & -3 \\ 0 & 1 & -3 \\ 0 & 0 & 0 \end{pmatrix} \quad .$$

The relations obtained are $x \succsim_I y \;\Leftrightarrow\; I(x) \geq 3I(y)$ for any two eventualities $x, y \in \{1, 2, 3\}$ and database I. Now consider the matrix

$$v_4 = \begin{pmatrix} 5/12 & 0 & -1 \\ 0 & 1 & -1 \\ 0 & 0 & 0 \end{pmatrix} .$$

If the matrices above are supplemented with v_4 the relations do not change in any way, in spite of v_4 not being contained in V. In other words, if W is the closed convex cone generated by $v_1, \ldots, v_4$, then W and V are two distinct sets, each representing the same relations $\{\succsim_I\}_{I\in\mathbb{J}}$.

Without uniqueness, a 'standard' representing set of matrices is the maximal representing set w.r.t. set inclusion. This is the union of all representing sets, which is a closed convex cone and unique by definition. The maximal set has the advantage of containing all possible values of relevance-weights. If, after employing a consensus rule over

various interpretations of data as a first stage of comparison, a further examination of completion of relations is undertaken by employing a different rule (e.g. a minimum-support rule), then no relevance-weights are eliminated *a priori*. The second rule then relies on the full range of interpretations that are compatible with the initial incomplete relations.

3.2 Complete Relations

If completeness is assumed for relations $\{\succsim_I\}_{I\in\mathbb{J}}$, then a conclusion similar to GS can be obtained. For this to occur, Independence of Support needs to be augmented as follows in order to deal with the case of strict rankings.

A1'. Completeness: for any two eventualities x, y and database I, either $x \succsim_I y$ or $y \succsim_I x$.

For a complete relation $\succsim_I$, $\succ_I$ denotes its asymmetric part, while $\sim_I$ denotes its symmetric part.

A3-S. Strict Independence of Support: Let x, y be eventualities and K be a database. If there exists a set of relationships that satisfies conditions of support of x over y on K, where at least one of the relationships is strict, then $x \succ_K y$.

Proposition 2. *The following statements are equivalent:*

(i) The relations $\{\succsim_I\}_{I\in\mathbb{J}}$ satisfy assumptions A1', A3 and A3-S.

(ii) There exists a matrix $v : \mathbb{X} \times \mathbb{C} \longrightarrow R$, such that for any two eventualities x, y and database I,

$$x \succsim_I y \quad \Leftrightarrow \quad \sum_{c\in\mathbb{C}} v(x,c)I(c) \geq \sum_{c\in\mathbb{C}} v(y,c)I(c) \quad . \tag{6}$$

There are two significant differences between the representation obtained here and that of the GS model. First, the assumptions used here imply a broader range of

relations, as the relations are not required to be diverse, and correspondingly the representing matrix (in the terminology of GS) is not necessarily diversified.[7]

On the other hand, while in the GS model the representing matrix is essentially unique (up to multiplication by a positive constant and shift by a matrix with identical rows), Proposition 2 does not provide such uniqueness. The following is an example of a family of complete relations which satisfy (i) of Proposition 2, yet admit two different representations by two distinct matrices.

Example 3. Let $\mathbb{X} = \{x, y, z\}$, $\mathbb{C} = \{1, 2, 3\}$, and

$$v = \begin{pmatrix} 1 & 0 & 3 \\ 0 & 1 & 2 \\ 0.5 & 0.5 & 2.5 \end{pmatrix}.$$

The relationships defined are: if $I(1) + I(3) \geq I(2)$, then $x \succsim_I z \succsim_I y$, otherwise $y \succsim_I z \succsim_I x$. This representation, however, is not unique. The same relationships are induced, for instance, by the matrix:

$$w = \begin{pmatrix} 1 & 0 & 3 \\ 0 & 1 & 2 \\ 0.25 & 0.75 & 2.25 \end{pmatrix}.$$

The difference between the representation of Proposition 2 and that of GS stems from the fact that the axioms of Combination and Diversity used in the GS model (See Subsection 2.2 for their formulation) are replaced by Independence of Support in both its strict and non-strict versions. As discussed in 2.2, Independence of Support generalizes and strengthens Combination. In so doing, it makes Diversity unnecessary. The Independence of Support axioms together with Completeness are sufficient to obtain representation (6) (In GS both strict and weak versions of Combination are required as well). However, Independence of Support, as mentioned before, does not

[7] A matrix $v : \mathbb{X} \times \mathbb{C} \longrightarrow R$, where $|\mathbb{X}| \geq 4$, is diversified if there exist no four distinct elements $x, y, z, w \in \mathbb{X}$ and $\lambda, \mu, \theta \in R$ with $\lambda + \mu + \theta = 1$, such that $v(x, \cdot) \leq \lambda v(y, \cdot) + \mu v(z, \cdot) + \theta v(w, \cdot)$. If $|\mathbb{X}| < 4$, v is diversified if no row in v is dominated by an affine combination of the others.

yield a unique representing matrix. Note that in the example above Diversity is not satisfied, as eventuality z is always ranked between x and y. This example therefore cannot be accommodated in the GS model.

4 Alternative Applications

In addition to the evaluation process discussed in this paper, there are a variety of problems in areas of interest as diverse as statistics, welfare analysis, and economics where our model may be a useful tool. The basic commonality that these problems share with the current model is that the analysis of data can be interpreted in different ways yielding conflicting rankings.

4.1 Classification Problems

In a classification problem one must predict the class of a new data point based on past observations. The prediction will be positioned in the most common class among the most relevant observations in the database. The relevance of an observation in the database is determined by a kernel function that depends on the distance between the attributes of the new data point and the attributes of the observation in question. Gilboa and Schmeidler (2003) establish that a classification problem can be viewed as a special instance of their model by defining $v(x, c)$ to equal the kernel function if observation c belongs to class x, otherwise $v(x, c) = 0$. In fact, the medical example discussed in section 3 is also a classification problem in which the doctor must predict whether the class of the current patient will be success or failure. However, instead of employing a single kernel to predict the outcome, the doctor considers several kernels. As demonstrated in the example, classification of the new data point may greatly depend upon which kernel is employed. Indeed when there is no good reason to choose one kernel over another, it would be hard to classify the new data point, and care must be taken to require that classification be made only upon consideration of all kernels.

4.2 Welfare Comparisons

Standard of living is used to compare welfare among societies. Generally speaking, standard of living depends on a variety of factors such as average income, inequality of income, poverty, education, life expectancy, etc. However, there are many ways to calculate this measure, each calculation producing a different ranking of countries. For example the United States, that has a high income per capita that is, however, distributed very unevenly, will rank higher in measures that put more weight on average income, and will rank lower in measures that put more weight on income inequality relative to other developed countries. In 2011, the United States was ranked above Germany according to the Human Development Index (HDI), but below Germany according to the Inequality-adjusted HDI (IHDI).[8]

The model suggested in this paper may incorporate ranking of countries based on different welfare indices by applying a unanimity rule. A corresponding database may be composed of observations of income, education, state of health, etc. of residents of each country, with each index corresponding to a different matrix v. Under the unanimity rule it would be impossible to compare the welfare of the United States with that of Germany, as different indices produced opposite rankings. On the other hand, the welfare of the United States and Norway can be compared, since the latter ranked higher on both indices.

4.3 Voting Games

Many voting systems can be represented as cooperative games. If N is the set of voters, the value $w(T)$ of coalition $T \subset N$ is set to 1 if T can accumulate enough votes to become a winning coalition, and to 0 if it cannot. Power indices provide a measure of the voting power of each voter and play an important role in examining the fairness of voting systems. Two central power indices are those of Shapley-Shubik (1954) and Banzhaf (1965). The Shapley-Shubik power index of voter i in a voting system w is

[8] This data can be found on the website of the United Nations Development Program (UNDP).

given by:

$$S_i(w) = \sum_{T \subset N \setminus i} \frac{|T|!(|N| - |T| - 1)!}{|N|!} \left(w(T \cup i) - w(T)\right) \ ,$$

while the Banzhaf power index of voter i in a voting system w is given by:[9]

$$B_i(w) = \sum_{T \subset N \setminus i} \frac{1}{2^{|N|-1}} \left(w(T \cup i) - w(T)\right) \ .$$

It is well known that in some games the Shapley-Shubik and the Banzhaf power indices yield similar results, while for others they vary considerably. In fact, they sometimes rank the power of two voters in opposite ways. It has been argued that the Shapley-Shubik index is more appropriate when the order of the players casting a vote matters, while the Banzhaf index is more appropriate when order does not matter. However, it is not always obvious when the order in fact matters, making it unclear which index to use.

Ranking the voting power of players may be described by the model presented here by allowing the eventualities to be the n players, and a record to be the value of the coalitions, $w(T)$, for all $T \subset N$. Two matrices, the Shapley-Shubik matrix, v^S, and the Banzhaf matrix, v^B, may be defined in the following manner, for all players i and coalitions $T \subset N$:

$$v^S(i,T) = \begin{cases} \frac{|T-1|!(n-|T|-2)!}{n!} & i \in T \\ -\frac{|T|!(n-|T|-1)!}{n!} & i \notin T \end{cases} \quad , \quad v^B(i,T) = \begin{cases} \frac{1}{2^{n-1}} & i \in T \\ -\frac{1}{2^{n-1}} & i \notin T \end{cases} \quad .$$

For each player i, $S_i(w) = \sum_{T \subset N} v^S(i,T)w(T)$, and $B_i(w) = \sum_{T \subset N} v^B(i,T)w(T)$. Using a unanimity rule it is possible to rank the voting power of two voters only when their rankings in both indices match.

4.4 Group Rankings

Another scenario for employing multiple relevance-weights is in describing a ranking process performed by a group of individuals. The process of ranking eventualities only

[9] This is the non-normalized version of the Banzhaf power index since the weights of all voters do not necessarily add to 1.

when all relevance-weighting functions are in agreement corresponds to the application of a unanimity rule. That is, one eventuality will be ranked above another only if all members agree on the ranking. Here, each member may hold a solid view of the appropriate degree of support for each eventuality; however, as members' views may differ, a unanimity decision rule will generally lead to incompleteness.

5 Proofs

5.1 Proof of Remark 1

For Transitivity, apply A3 to the relationships $x \succsim_I y,\ y \succsim_I z$, which satisfy conditions of support of eventuality x over eventuality z on database I. For Combination, apply A3 to the relationships $x \succsim_I y,\ x \succsim_J y$, which satisfy conditions of support of eventuality x over eventuality y on database $I+J$. In order to see that one direction of Homogeneity holds, apply A3 to n copies of the relationship $x \succsim_I y$, which then satisfy conditions of support of eventuality x over eventuality y on database nI. In the other direction apply axiom A2. Last, Impartiality is seen to be satisfies by applying A3 to the relationships $x \succsim_I x,\ y \succsim_I y$, which hold true by Reflexivity (for any database I). These relations satisfy conditions of support of eventuality x over eventuality y, as well as of eventuality y over eventuality x, on the all-zeros database.

5.2 Proof of Theorem 1

5.2.1 *Proof of the Direction (i)⇒(ii)*

The proof of this direction, as well as the proof of Proposition 2, use definitions and arguments from Ashkenazi and Lehrer (2001).

Extend $\mathbb{J}$ to nonnegative rational numbers, so that $\mathbb{J} = Q_+^{\mathbb{C}}$. The relations $\{\succsim_I\}_{I\in\mathbb{J}}$ are accordingly extended to rational-valued databases in the following manner: For a database $I \in \mathbb{J}$ (with rational entries), let $k \in N$ be such that $kI \in Z_+^{\mathbb{C}}$, and define $\succsim_I = \succsim_{kI}$. By Homogeneity (see Remark 1) the extension is well defined, and satisfies $\succsim_I = \succsim_{qI}$ for every $q \in Q$, $q > 0$. Assumptions A1 and A2, as well as A3-E, immediately

carry through to the extended $\mathbb{J}$, and will thus be employed without further mention. Axiom A3, when used in course of the proof, will explicitly be shown to work with rational-valued databases.

Let S denote the number of elements in $\mathbb{X}$ and T the number of elements in $\mathbb{C}$. Consider the Euclidean space R^{ST} as consisting of vectors with S blocks, T entries in each block. Each block corresponds to an eventuality, so that for $\varphi \in R^{ST}$, φ_x denotes the block matching eventuality x. For any two eventualities $x, y \in \mathbb{X}$ and database I, the vector $\varphi(I, x, y)$ denotes the vector in R^{ST} for which the block corresponding to eventuality x is I, the block corresponding to eventuality y is $-I$, and those corresponding to other eventualities are all zero.

Consider $\mathcal{B} = conv\{\varphi(I, x, y) \mid x, y \in \mathbb{X},\ I \in \mathbb{J}\} \subset R^{ST}$, with its relative topology. For $v \in R^{ST}$, eventuality x and record type c, $v(x, c)$ denotes the entry of v which belongs to block x, and inside block x, to record type c.

Define $E = cl(conv\{\varphi(I, x, y) \mid x \succsim_I y,\ x, y \in X,\ I \in \mathbb{J}\})$. That is, E is the closed convex hull generated by all vectors in $\mathcal{B}$ that indicate ranking. As Homogeneity and Impartiality are satisfied (see Remark 1), E is a closed convex cone with vertex at zero. By its definition, E contains all vectors that indicate ranking. That is, if $x \succsim_I y$ then $\varphi(I, x, y) \in E$. The following claims show that the opposite is also true, namely that if a vector of the form $\varphi(I, x, y)$ is contained in E, then $x \succsim_I y$.

Recall that if there are relationships $b_1 \succsim_{J_1} w_1, \ldots, b_n \succsim_{J_n} w_n$ that satisfy conditions of support of x over y on database K, then these relationships also ε-satisfy these conditions of support for every $\varepsilon > 0$, thus A3 implies that $x \succsim_K y$.

Claim 3. *Let $x, y, b_1, w_1, \ldots, b_n, w_n$ be eventualities and $K, J_1, \ldots, J_n$ databases. Suppose that $b_l \succsim_{J_l} w_l$ for $l = 1, \ldots, n$, and that there are strictly positive coefficients $q_1, \ldots, q_n \in Q$, such that*

$$\varphi(K, x, y) = \sum_{l=1}^{n} q_l \varphi(J_l, b_l, w_l) \ .$$

Then $x \succsim_K y$.

Proof. The vector $\varphi(K, x, y)$ has zeros in all blocks but those that correspond to eventualities x and y. Hence, for any eventuality $z \neq x, y$, by setting $I_l = q_l J_l$, it is

implied that

$$\sum_{l:b_l=z} I_l - \sum_{l:w_l=z} I_l = 0 \ .$$

By Homogeneity, as $b_l \succsim_{J_l} w_l$, then also $b_l \succsim_{I_l} w_l$ (since the q_l's are rational). In Addition, by the structure of the vector $\varphi(K, x, y)$,

$$\begin{aligned} \sum_{l:b_l=x} I_l - \sum_{l:w_l=x} I_l &= K, \quad \text{and} \\ \sum_{l:b_l=y} I_l - \sum_{l:w_l=y} I_l &= -K \ . \end{aligned}$$

Thus, Independence of Support implies that $x \succsim_K y$. ■

Claim 4. *Let $x, y, b_1, w_1, \ldots, b_n, w_n$ be eventualities and $K, J_1, \ldots, J_n$ databases. Suppose that $b_l \succsim_{J_l} w_l$ for $l = 1, \ldots, n$, and that there are strictly positive coefficients $\alpha_1, \ldots, \alpha_n \in R$, such that*

$$\varphi(K, x, y) = \sum_{l=1}^{n} \alpha_l \varphi(J_l, b_l, w_l) \ .$$

Then $x \succsim_K y$.

Proof. Consider the following finite system of equations, with $\alpha_1, \ldots, \alpha_n$ as variables:

$$\begin{aligned} &\sum_{l:b_l=z} \alpha_l J_l - \sum_{l:w_l=z} \alpha_l J_l = 0 \ , \quad \text{for all } z \neq x, y \\ &\sum_{l:b_l=x} \alpha_l J_l - \sum_{l:w_l=x} \alpha_l J_l = K \\ &\sum_{l:b_l=y} \alpha_l J_l - \sum_{l:w_l=y} \alpha_l J_l = -K \end{aligned}$$

All the coefficients (entries of the databases) are rational numbers. Therefore, if the system has a positive solution $\alpha_1, \ldots, \alpha_n$, then, by denseness of the rational numbers in the reals, it must also have a positive rational solution $q_1, \ldots, q_n$. In other words, employing again the structure of the vector $\varphi(K, x, y)$, there are positive rational coefficients such that $\varphi(K, x, y) = \sum_{l=1}^{n} q_l \varphi(J_l, b_l, w_l)$. By the previous claim it follows that $x \succsim_K y$. ■

It is now established that if a vector of the form $\varphi(K, x, y)$ is obtained as a convex combination of vectors which indicate ranking, then this vector itself indicates ranking. Therefore, taking the convex hull of vectors which satisfy $x \succsim_I y$ does not add any 'unwanted' vectors to E, that is, does not add any $\varphi(I, x, y)$ for which $\neg(x \succsim_I y)$.

It is next proved that applying a closure operation does not add any vectors $\varphi(I, x, y)$ for which $\neg(x \succsim_I y)$ as well.

Claim 5. *If $\varphi(I, x, y) \in E$ then $x \succsim_I y$.*

Proof. In light of the previous claim it remains to show that if $\varphi(I, x, y)$ is on the boundary of E then $x \succsim_I y$. As E is a convex cone with vertex at zero, if $\varphi(I, x, y)$ is on the boundary of E then so is $\varphi(iI, x, y)$ for iI an integer-valued database. Therefore there exists a sequence of points in E, $\varphi^n = \sum_{t=1}^{T_n} \lambda_t^n \varphi(I_t^n, x_t^n, y_t^n)$, which converges to $\varphi(iI, x, y)$, where $\lambda_t^n > 0$ and $x_t^n \succsim_{I_t^n} y_t^n$. By denseness of the rational numbers in the reals there must also be a sequence $q^n = \sum_{t=1}^{T_n} r_t^n \varphi(I_t^n, x_t^n, y_t^n) = \sum_{t=1}^{T_n} \varphi(r_t^n I_t^n, x_t^n, y_t^n)$, with $r_t^n > 0$ rational numbers, that also converges to $\varphi(iI, x, y)$.

By the structure of $\varphi(iI, x, y)$, if q^n converges to $\varphi(iI, x, y)$, then for any $\varepsilon > 0$ there exists $n \in N$ such that for every $c \in \mathbb{C}$,

$$
\begin{aligned}
&|\sum_{t:x_t^n=x} r_t^n I_t^n(c) - \sum_{t:y_t^n=x} r_t^n I_t^n(c) - iI(c)| < \varepsilon \\
&|\sum_{t:y_t^n=y} r_t^n I_t^n(c) - \sum_{t:x_t^n=y} r_t^n I_t^n(c) - iI(c)| < \varepsilon \\
&|\sum_{t:x_t^n=z} r_t^n I_t^n(c) - \sum_{t:y_t^n=z} r_t^n I_t^n(c)| < \varepsilon, \quad \text{for every } z \neq x, y\ .
\end{aligned}
$$

Let m be an integer such that $L_t := m r_t^n I_t^n$ is an integer-valued database for all t, then still $x_t^n \succsim_{L_t} y_t^n$, and for every $c \in \mathbb{C}$,

$$
\begin{aligned}
&|\sum_{t:x_t^n=x} L_t(c) - \sum_{t:y_t^n=x} L_t(c) - m(iI(c))| < m\varepsilon \\
&|\sum_{t:y_t^n=y} L_t(c) - \sum_{t:x_t^n=y} L_t(c) - m(iI(c))| < m\varepsilon \\
&|\sum_{t:x_t^n=z} L_t(c) - \sum_{t:y_t^n=z} L_t(c)| < m\varepsilon, \quad \text{for every } z \neq x, y\ .
\end{aligned}
$$

That is, the relationships $x_t^n \succsim_{L_t} y_t^n, t = 1, \ldots, T_n$, ε-satisfy conditions of support of eventuality x over eventuality y on database iI. Since this holds true for any $\varepsilon > 0$, A3 implies that $x \succsim_{iI} y$, thus also $x \succsim_I y$. ■

The previous claim yields that E contains exactly those vectors $\varphi(I, x, y)$ for which $\neg(x \succsim_I y)$, as summarized in the next conclusion.

Conclusion 6. *$\varphi(I, x, y) \in E$ if and only if $x \succsim_I y$.*

Define $V = \{v \in R^{ST} \mid v \cdot \varphi \geq 0 \text{ for all } \varphi \in E\}$ (where $v \cdot \varphi$ denotes the inner product of v and φ). The set V is not empty, since it contains zero. If v_1 and v_2 are in V then so is $\alpha v_1 + \beta v_2$ for $\alpha, \beta \geq 0$, and if $v_n \to_{n\to\infty} v$ for $v_n \in V$, then $v \in V$. Hence V is a closed convex cone, with vertex at zero.

By the definitions of V and E, if $x \succsim_I y$ then $\sum_{c\in\mathbb{C}} v(x, c)I(c) \geq \sum_{c\in\mathbb{C}} v(y, c)I(c)$ for every $v \in V$. In the other direction, suppose that $\neg(x \succsim_I y)$, that is, $\varphi(I, x, y) \notin E$. Since E is closed and convex, then by a separation theorem there exists a vector $w \in R^{ST}$ which separates E and $\varphi(I, x, y)$. Following Remark 1, the zero vector belongs to E, and if $\varphi(I, x, y) \notin E$ then $\varphi(qI, x, y) = q\varphi(I, x, y) \notin E$ for any $q \in Q$. Hence the separating scalar is zero, that is, $w \cdot \psi \geq 0 > w \cdot \varphi(I, x, y)$, for every $\psi \in E$. In other words, there exists $w \in V$ such that $\sum_{c\in\mathbb{C}} w(x, c)I(c) < \sum_{c\in\mathbb{C}} w(y, c)I(c)$. Thus it is established that there exists a set $V \subseteq R^{ST}$, such that for every pair of eventualities x, y and database I,

$$x \succsim_I y \quad \text{if and only if} \quad \sum_{c\in\mathbb{C}} v(x, c)I(c) \geq \sum_{c\in\mathbb{C}} v(y, c)I(c) \text{ for all } v \in V \ .$$

5.2.2 *Proof of the Direction (ii)⇒(i)*

Suppose that the relations $\{\succsim_I\}_{I\in\mathbb{J}}$ obtain a representation as in (ii) of Theorem 1. It is immediate to see that Reflexivity (A1) is satisfied. Assumption A2 is proved by multiplying both sides of the inequality by the same constant.

For Independence of Support (A3), let x, y be eventualities and K a database, such that for every $\varepsilon > 0$ there are relationships (ε-dependent) $b_1 \succsim_{J_1} w_1, \ldots, b_n \succsim_{J_n} w_n$

and an integer m_ε that satisfy, for every $c \in \mathbb{C}$,

$$
\begin{aligned}
|\sum_{l:b_l=x} J_l(c) - \sum_{l:w_l=x} J_l(c) - m_\varepsilon K(c)| &< m_\varepsilon \varepsilon \ , \\
|\sum_{l:w_l=y} J_l(c) - \sum_{l:b_l=y} J_l(c) - m_\varepsilon K(c)| &< m_\varepsilon \varepsilon \ , \\
|\sum_{l:b_l=z} J_l(c) - \sum_{l:w_l=z} J_l(c)| &< m_\varepsilon \varepsilon, \quad \text{for every } z \neq x, y \ .
\end{aligned}
$$

First, for every $\varepsilon > 0$, by summing over all the inequalities implied by the relationships,

$$
\sum_{l=1}^{n} \left[\sum_{c \in \mathbb{C}} v(b_l, c) J_l(c) - \sum_{c \in \mathbb{C}} v(w_l, c) J_l(c) \right] =: R_v^\varepsilon \geq 0 \ , \ \text{for every } v \in V \ .
$$

After rearranging the addends in the above expression according to the eventualities involved, it is obtained that:

$$
\begin{aligned}
0 \leq R_v^\varepsilon &= \sum_{l:b_l=x} \sum_{c \in \mathbb{C}} v(x,c) J_l(c) - \sum_{l:w_l=x} \sum_{c \in \mathbb{C}} v(x,c) J_l(c) + \\
&\quad \sum_{l:b_l=y} \sum_{c \in \mathbb{C}} v(y,c) J_l(c) - \sum_{l:w_l=y} \sum_{c \in \mathbb{C}} v(y,c) J_l(c) + \\
&\quad \sum_{z \neq x,y} \left[\sum_{l:b_l=z} \sum_{c \in \mathbb{C}} v(z,c) J_l(c) - \sum_{l:w_l=z} \sum_{c \in \mathbb{C}} v(z,c) J_l(c) \right] \\
&= \sum_{c \in \mathbb{C}} v(x,c) \left(\sum_{l:b_l=x} J_l(c) - \sum_{l:w_l=x} J_l(c) \right) + \qquad (7) \\
&\quad \sum_{c \in \mathbb{C}} v(y,c) \left(\sum_{l:b_l=y} J_l(c) - \sum_{l:w_l=y} J_l(c) \right) + \qquad (8) \\
&\quad \sum_{z \neq x,y} \left[\sum_{c \in \mathbb{C}} v(z,c) \left(\sum_{l:b_l=z} J_l(c) - \sum_{l:w_l=z} J_l(c) \right) \right] \qquad (9)
\end{aligned}
$$

Given the inequalities assumed on the databases involved, for every $v \in V$ (which is possibly all zeros),

$$
\begin{aligned}
(7) &\leq m_\varepsilon \sum_{c \in \mathbb{C}} v(x,c) K(c) + m_\varepsilon \varepsilon \sum_{c \in \mathbb{C}} |v(x,c)| \ , \\
(8) &\leq -m_\varepsilon \sum_{c \in \mathbb{C}} v(y,c) K(c) + m_\varepsilon \varepsilon \sum_{c \in \mathbb{C}} |v(y,c)| \ , \\
(9) &\leq m_\varepsilon \varepsilon \sum_{z \neq x,y} \sum_{c \in \mathbb{C}} |v(z,c)| \ .
\end{aligned}
$$

Therefore, for every $v \in V$,

$$\sum_{c\in\mathbb{C}} K(c)[v(x,c) - v(y,c)] \geq -\varepsilon \sum_{x\in\mathbb{X}} \sum_{c\in\mathbb{C}} |v(x,c)| \ .$$

Since there is a finite number of eventualities and records, it follows that for every $v \in V$, $\sum_{c\in\mathbb{C}} K(c)[v(x,c) - v(y,c)] \geq 0$, implying $x \succsim_K y$.

5.3 Proof of Proposition 2

Let the set E be as in the proof of Theorem 1. In the degenerate case where $x \succsim_I y$ for all x, y, I, E is the entire space, and v may be taken to be zero. Otherwise, applying Completeness, assume that for some x, y, I, $x \succ_I y$.

From the proof of that theorem, the set E is closed and therefore E^c is open. By definition, the sets E and E^c are disjoint, where
$x \succsim_I y \Leftrightarrow \varphi(I,x,y) \in E$, and with Completeness $x \succ_I y \Leftrightarrow \varphi(I,y,x) = -\varphi(I,x,y) \in E^c$.

Let $D = conv\{-\varphi(I,x,y) \mid x,y \in X,\ I \in \mathbf{J},\ x \succ_I y\}$. By A3-S and Completeness, if $-\varphi(I,x,y) \in D$ then $x \succ_I y$ (the proof is analogous to claims 3 and 4), and D is a convex cone with vertex at zero, not containing zero. It follows that $D \subseteq E^c$, and since E^c is open then so is D, being composed of convex combinations of vectors from E^c. By assumption there is $-\varphi(I,x,y) = \varphi(I,y,x) \in D$, thus the interior of D is nonempty.

The above yields two disjoint convex cones, one open (D) and one closed (E), with vertex at zero, one of them with a nonempty interior (D). By a basic separation theorem, there exists a nonzero separating vector v for which $v \cdot \psi < 0 \leq v \cdot \varphi$ for every $\psi \in D$ and every $\varphi \in E$. It follows that $x \succsim_I y \Leftrightarrow \varphi(I,x,y) \in E \Rightarrow I \cdot (v_x - v_y) \geq 0$, and $x \succ_I y \Leftrightarrow \varphi(I,y,x) \in D \Rightarrow I \cdot (v_x - v_y) > 0$. In other words, for any database I and eventualities x and y,

$$x \succsim_I y \quad \Leftrightarrow \quad \sum_{c\in\mathbb{C}} v(x,c)I(c) \geq \sum_{c\in\mathbb{C}} v(y,c)I(c) \quad .$$

The proof of the direction (ii)$\Rightarrow$(i) is the same as for Theorem 1, with Completeness

and the strengthening of Relevance Separability, A3-S, implied by the fact that there is a single representing matrix.

REFERENCES

Alon, S., and E. Lehrer (2012), "Subjective multi-prior probability: A representation of a partial likelihood relation", mimeo.

Ashkenazi, G., and E. Lehrer (2001), "Relative Utility", mimeo.

Banzhaf, J. F. (1965), "Weighted Voting Doesn't Work: A Mathematical Analysis", *Rutgers Law Review*, 19, 317-343.

Bewley, T. (2002), "Knightian Decision Theory: Part I", *Decisions in Economics and Finance*, 25, 79–110.

Eichberger, J., and Guerdjikove, A. (2010), "Case-Based Belief Formation under Ambiguity", *Mathematical Social Sciences*, 60, 161-177.

Eichberger, J., AND Guerdjikove, A. (2011) "Ambiguity, Data and Preferences for Information— A Case-Based Approach", mimeo.

Gajdos, T., T. Hayashi, J. M. Tallon, J. C. Vergnaud. (2008), "Attitude toward Imprecise Information", *Journal of Economic Theory,* 140, 27-65.

Galaabaatar, T. and E. Karni (2011) "Objective and Subjective Expected Utility with Incomplete Preferences", mimeo.

Gilboa, I., F. Maccheroni, M. Marinacci, and D. Schmeidler (2010), "Objective and Subjective Rationality in a Multiple Prior Model", *Econometrica*, 78, 755-770.

Gilboa, I., and D. Schmeidler (1995), "Case-Based Decision Theory", *Quarterly Journal of Economics*, 110, 605-639.

Gilboa, I., and D. Schmeidler (1997), "Act similarity in case-based decision theory", *Economic Theory*, 9, 47-61.

Gilboa, I., and D. Schmeidler (2003), "Inductive Inference: An Axiomatic Approach", *Econometrica*, 71, 1-26.

Giron, F. J. and S. Rios (1980), Quasi-Bayesian behaviour: A More Realistic Approach to Decision Making? In J. M. Bernardo, J. H. DeGroot, D. V. Lindley,

and A. F. M. Smith, editors, Bayesian Statistics,17-38. University Press, Valencia, Spain.

Kraft, C.H., J.W. Pratt, and A. Seidenberg (1959), "Intuitive Probability on Finite Sets", *The Annals of Mathematical Statistics*, Vol. 30, No. 2, 408-419.

Krantz, D.H., R.D. Luce, P. Suppes, and A. Tversky (1971), *Foundations of Measurement*, New York: Academic Press.

Narens, L. (1974), "Minimal conditions for additive conjoint measurement and qualitative probability", *Journal of Mathematical Psychology*, Vol. 11, Issue 4, 404-430.

Scott, D. (1964), "Measurement Structures and Linear Inequalities", *Journal of Mathematical Psychology*, Vol. 1, 233-247.

Shapley, L. S. and M. Shubik (1954), "A Method for Evaluating the Distribution of Power in a Committee System," *The American Political Science Review*, 48, 787-792.