FOUNDATIONS FOR THE THEORY OF RATIONAL CHOICE WITH VAGUE PRIORS

Klaus Nehring*

Abstract Vagueness of belief is modeled by allowing the decision maker to entertain (convex) sets of probabilities over the n possible states. Such convex belief sets can be obtained by dropping the completeness assumption in otherwise standard expected-utility axiomatizations, as Bewley and others have shown. The task of this paper is to demonstrate that decision problems with arbitrary belief sets can be reduced to "complete ignorance" problems, i.e., to problems whose belief sets are n-simplexes Σ^n and thus maximally inclusive.

This is formally captured in the condition of "complete ignorance reduction" (CIR). CIR asserts that any decision problem with vague priors (X, P) has an equivalent complete ignorance problem (Y, Σ^m), whose states correspond to the m extreme points of P, and whose payoffs in some state are given by the expected utility of the considered act under the extremal belief associated with that state.

The axioms employed in the derivation of CIR capture the idea that only payoff-relevant differences in beliefs should matter. They also provide a foundation for the axiom of "Replication Invariance" (sometimes also called "Invariance with respect to the merger of states") which is the hallmark of the existing literature on complete ignorance problems. Our paper thus integrates two approaches to vagueness which so far have lead completely separate lives.

1. INTRODUCTION

This paper attempts to provide the foundations for a general decision theory for vague beliefs. "Vagueness of belief" is assumed to obtain when the decision-maker is (or should rationally be) unwilling to commit himself to a unique subjective probability (vector) $p \in \Sigma^S$, where Σ^S denotes the unit simplex of $\mathbf{R}^s$ and S the set of "possible" states. Vagueness will be modeled by allowing the decision-maker to entertain convex sets of probabilities $P \subset \Sigma^S$.

This kind of approach has been forcefully advocated by Levi (1980) among others; it has received axiomatic support for instance in the work of Bewley (1986) who, building upon Smith (1961), has characterized convex "belief sets" in terms of incomplete preferences.

A more extensive discussion of the concept of vagueness and its modeling in terms of belief sets can be found in Nehring (1991a). Nehring (1991b) provides qualified support for the convexity assumption.

The literature has focused almost exclusively on the limiting cases of vagueness, of a determinate subjective probability on the one hand ("Bayesian Decision Theory") and of maximal vagueness on the other ("Complete Ignorance").

---

*Harvard University, July 1990.

"Complete Ignorance" has been characterized in the literature by two axioms of "Symmetry" and of "Replication Invariance" (also: "Merger of States"); the latter axiom rules out expected-utility maximization. We show that if Complete Ignorance is represented by maximally inclusive belief sets $P = \Sigma^S$, the classical axioms can be derived from more basic and general ones.

The major contribution of this paper is to provide an axiomatically justified method whereby choice-functions defined on the class of "complete ignorance problems" and satisfying these axioms ("CI-solutions") can be extended to general belief sets. The only restriction is that the belief sets have to have a finite number of extreme points, i.e., that they must be polyhedra; it is imposed solely for reasons of tractability.

The method we develop, called "CI-reduction," is a very simple one. It is shown that any decision problem with belief sets is equivalent to a CI-problem in which each state corresponds to an extreme point of the original belief set, and in which acts are identified with the vectors of expected utilities under these extremal probabilities.

It is important that the method of CI-reduction allows the extension of arbitrary CI-solutions. Nehring (1991a) axiomatizes a CI-solution of "Simultaneous Utility Maximization" ("SUM") which is superior to existing solutions and can lay claims to rationality by thoroughly satisfying the sure-thing principle applied to belief sets. It can be understood as a radicalized version of the classical "Savage rule."[1] The applicability of CI-reduction to arbitrary CI-solutions is crucial to ensure the extendability of SUM.

The only other work that seeks to develop a reasonably general decision theory for belief sets is Jaffray's (1988, 1989). His idea is also to extend CI-solutions, but his method is rather different. It has two serious drawbacks which our method avoids: its domain is still severely limited (to belief sets that can be characterized in terms of "belief functions"). Moreover, it has to be formulated in terms of preferences rather than choice-functions, and is thereby incompatible with SUM or the Savage rule.

2. FRAMEWORK AND NOTATION

Let S^∞ denote an (infinite) "universe of states," and let $\mathbb{S}$ be the class of its finite subsets. A belief set is a closed convex set P contained in some Σ^S, the unit simplex of $\mathbf{R}^S$. "$P \subset \Sigma^S$" simply means that the decision maker attributes probability zero (unambiguously) to all state of $S^\infty \setminus S$. Let $\mathbb{P}^S$ be the class of belief sets with a finite number of extreme points.[2]

An *act* x is an element of $\mathbf{R}^S$. Its consequences are to be interpreted as "payoffs" in terms of cardinal utility; the existence of a cardinal utility could be derived without difficulty by the standard "horse-lottery" technique due to Anscombe and Aumann (1963).

---

[1]Treated at length in the second half of Savage's "Foundations of Statistics" (1972^2).

[2]Recall that p is called an extreme point of P if and only if it is not a convex combination of two different elements of P.

Choice sets X are subsets of $\mathbb{R}^S$. $\mathbb{X}^S$ is the class of such sets; it may or may not be restricted by conditions such as compactness or convexity.

$\mathcal{E}(P)$	, the set of extreme points of P.
$\mathcal{CH}\, X$	, the convex hull of X
$x \cdot p$	, the scalar product of x and p: $x \cdot p = \sum_{s \in S} x_s p_s$.
$x_{-s} \in \mathbb{R}^{S \setminus s}$	, the vector of non-s components of x.

At the heart of this paper lies the notion that decision problems can often be described in different but equivalent ways by eliminating payoff-irrelevant aspects of the original description. As plain as it seems, it turns out to have remarkably strong consequences in the context of belief sets.

The possibility of redescription can be captured formally by introducing partitions of the set of possible states, as follows:

Let $\mathbb{F}$ be the set of some partitions of some $S \in \mathbb{S}$.

For any partition $F = \{F_i\}_{i \in I}$ of S into "events" F_i, let $\mathbb{R}^S_F$ be the #F-dimensional linear subspace of acts which are constant on the elements of partition, i.e.

$\mathbb{R}^S_F = \left\{ x \in \mathbb{R}^s \mid x_s = x_t \text{ if } s, t \in F_i, \forall\, s, t \in S,\, i \in I \right\}$, and define
$\mathbb{X}^S_F = \{ x \in \mathbb{X}^S \mid X \subset \mathbb{R}^S_F \}$.

If $X \in \mathbb{X}^S_F$, the payoff-structure may be specified without loss of information in terms of F; X may thus be redescribed as a subset of $\mathbb{R}^F$. The redescribed choice-set X^F is defined as $v^F(X)$, where v^F is the canonical bijection from $\mathbb{R}^S_F$ to $\mathbb{R}^F$, with $v^F(x)_{F_i} = x_s$, for any $s \in F_i$ and $i \in I$. We will typically write $v^F(x) = x^F$.

For any $P \in \mathbb{P}^S$, the marginal belief set $P^F \subset \Sigma^F$ induced by P on F is specified in the natural way, elementwise:

$$p^F = \{ p^F \mid p \in P \}, \text{ with } p^F(F_i) = \sum_{s \in F_i} p_s \quad \text{for } F_i \in F.$$

Note that since the mapping $p \mapsto p^F$ is linear, the marginals of polyhedra are polyhedra.

Abbreviate $\mathbb{D}^F = \mathbb{X}^F \times \mathbb{P}^F$, and let $\mathbb{D}^{\tilde{}} = \{\mathbb{D}^F \mid F \text{ is a partition of some } S \in \mathbb{S}\}$, the class of all problems "under general description". Similarly, $\mathbb{D} = \{\mathbb{D}^S \mid S \in \mathbb{S}\} \subset \mathbb{D}^{\tilde{}}$. Also, let $\mathbb{D}^{\tilde{}CI} = \{(X, \Sigma^F) \in \mathbb{D}^{\tilde{}}$ and $\mathbb{D}^{CI} = \{(X, \Sigma^S) \in \mathbb{D}\}$.

A solution is a mapping $C : \mathbb{D}^{\tilde{}} \rightarrow \bigcup_{F \in \mathbb{F}} \mathbb{R}^F$ such that $C(X,P) \subset X \; \forall (X,P) \in D^{\tilde{}}$.

3. DERIVATION OF REPLICATION INVARIANCE

Choices should not be affected by redescription, if the redescription preserves all the relevant information. This leads to the following condition RDI; in it, the qualifying condition is captured by the requirement that the belief set in the redescribed problem is the marginal of the original belief set plus an extreme point clause. For the purpose of reading this paper, take this clause simply as a matter of technical expediency, to allow to prove the "if-part" of proposition 2 below. In Nehring (1991b, section 1) it is explained as the result of a misspecification of the true logical structure of vague beliefs as sets rather than as equivalence classes of such sets.

DEF Redescription Invariance (RDI):

For any $(X,P) \in \mathbb{X}^S_F \times \mathbb{P}^S$ such that $\mathscr{E}(P)^F = \mathscr{E}(P^F)$: $C(X,P)^F = C(X^F, P^F)$.

Remark: The crucial issue concerning the validity of the axiom is that of "information preservation" mentioned above; the occurrence of an extreme-point clause shows its non-triviality. In Nehring (1991b, section 2 and 3) it is argued that the DM's belief attitudes "about F" are indeeed fully summarized by the marginal belief set P^F, given the satisfaction of the extreme-point condition.

RDI becomes powerful in conjunction with a condition of "General Isomorphism" (ISO). ISO asserts that belief set and choice set determine the optimal choice completely; the identity of the event $\{F_i\}_{i \in I}$ is irrelevant. In particular, it does not matter whether the events are "atomic" ($F_i = \{s\}$) or "composite" ($\#F_i > 1$).

With any bijection $\phi : F \rightarrow F'$ associate a bijection $\Phi : \mathbb{R}^F \rightarrow \mathbb{R}^{F'}$ by $\Phi(x)_{\phi(G)} = x_G \;\; \forall G \in F$.

DEF General Isomorphism (ISO):

For any $(X,P) \in \mathbb{X}^F \times \mathbb{P}^F$, $F, F' \in \mathbb{F}$ and any bijection $\phi : F \rightarrow F'$:

$$C(\Phi(X), \Phi(P)) = \Phi(C(X,P)).$$

Remark: The axiom of "General Isomorphism" is stronger than that of purely "Formal" Isomorphism ISO′, which is ISO restricted to problems in state-specification. Only ISO′ is explicitly invoked in the literature, often under the name "invariance with respect to the labeling of states"[3]. ISO needs, and is capable of, stronger justification: it reflects the principle of "consequentialist" rationality that the choice-worthiness of an act is fully determined by the valuation of its consequences and the estimate of their occurrence.

RDI and ISO together have a strong implication, the "Replication Invariance" axiom which has been the hallmark of the literature on complete ignorance problems. It may be stated thus:

DEF Replication Invariance (RI):

For any $X \in \mathbb{X}_F^S$ with $F = \{\{t\} \mid t \in S, t \neq s, s'\} \cup \{s, s'\}\}$:
$C(X_{-s}, \Sigma^{S \setminus s}) = C(X, \Sigma^S)_{-s}$.

In words: If a CI-problem has two states $s, s' \in S$ with identical payoffs for any act, elimination of one of the replicated states should not affect the optimal choice.

Denoting the restriction of RDI to $\mathbb{D}^{\sim}$ by RDI^{CI}, it is easy to show that

Proposition 1: $ISO \Rightarrow (RDI^{CI} \Leftrightarrow RI)$.

Proof: Take any $X \in \mathbb{X}_F^S$ with $F = \{\{t\} \mid \{t \in S, t \neq s, s'\} \cup \{s, s'\}$.

Define $\phi : F \to S \setminus s$ by $\phi(\{s,t\}) = t$ and $\phi(i) = i$ for $i \neq \{s,t\}$.
ISO implies

$$C(X_{-s}, \Sigma^{S \setminus s}) = \Phi(C(X^F, \Sigma^F)). \qquad 1)$$

Suppose C satisfies RDI^{CI} as well.
Then also $C(X^F, \Sigma^F) = C(X, \Sigma^S)^F$,
and thus $C(X_{-s}, \Sigma^{S \setminus s}) = \Phi(C(X, \Sigma^S)^F) = C(X, \Sigma^S)_{-s}$.
This proves $ISO \Rightarrow (RDI^{CI} \Rightarrow RI)$.

Conversely, suppose that RI holds besides ISO.
Then $C(X_{-s}, \Sigma^{S \setminus s}) = C(X, \Sigma^S)_{-s}$, implying with 1)

$$C(X^F, \Sigma^F) = \Phi^{-1}(C(X, \Sigma^S)_{-s}) = C(X, \Sigma^S)^F.$$

Repeating this argument for general F if necessary shows
$ISO \Rightarrow (RI \Rightarrow RDI^{CI})$. □

Remark: We have emphasized the importance of both assumptions for proposition 1. By contrast, the focus of the literature is on RDI^{CI}, used as a verbal argument for RI. ISO does not appear explicitly in the standard treatment; it is implicit in the notation, as if the equivalence

[3] Note that "invariance with respect to labeling" has to refer to the labeling of states, since events are not independently "labeled" but rather named as sets of states.

of states $s \in S \in \mathbb{S}$ and events $F_i \in F \in \mathbb{F}$ was a purely formal matter. But strong conclusions require strong premises; the camouflage of ISO makes RI appear as a rabbit jumping out of the magician's hat.

The second advance in the understanding of RI beyond the literature consists in the formulation of RDI in terms of (arbitrary) belief sets, and thereby making the issue of its validity explicitly one of "informational equivalence".

4. REDUCTION TO COMPLETE IGNORANCE PROBLEMS

We now want to tackle the second part of our task, to show that any decision problem can be reduced to a CI-problem "in expected utilities".

Let T be any set of states with $\#T = \#\mathscr{E}(P)$ and $\beta : T \to \mathscr{E}(P)$ a one-to-one map. Define $(\Psi(X), \Sigma^T)$, the "CI-reduction" of $(X,P) \in \mathbb{D}^F$ under β and ψ, by

$$\Psi : \mathbf{R}^F \to \mathbf{R}^T,\ \Psi(x) = \left(\sum_{G \in F} x_G \beta_G(t) \right)_{t \in T}, \text{ and } \Psi(X) = \{ \Psi(x) \mid x \in X \}.$$

DEF: CI-Reduction (CIR)

For any CI-reduction $(\Psi(X), \Sigma^T)$ of $(X,P) \in \mathbb{D}^F$,
$C(X,P) = \{x \in X \mid \Psi(x) \in C(\Psi(X), \Sigma^T)\}$.

In words: An act x is optimal in (X,P) if its extremal expected utility vector $\Psi(x)$ is optimal in the associated CI-problem in extremal expected utilities $(\Psi(X), \Sigma^T)$.

Example

Let C on $\mathbb{D}^{CI}$ be the maximin rule MM, defined by

$$MM(X) = \{ x \in X \mid \min_s x_s \geq \min_s y_s \ \forall y \in X \}.$$

CIR extends MM uniquely to $\mathbb{D}^{\sim}$, yielding the "maximum of expected utilities" (MMEU) rule, with

$$MMEU(X,P) = \{ \{ x \in X \mid \min \{ x \cdot p \mid p \in P \} \geq \min \{ y \cdot p \mid p \in P \} \ \forall y \in X \}.$$

- Note that in the case of MMEU, extreme points play no special role, since $\min \{ x \cdot p \mid p \in P \} = \min \{ x \cdot p \mid p \in \mathscr{E}(P) \}$. MMEU has been proposed by Gaerdenfors/Sahlin (1980) and - from a very different perspective than the present one - axiomatized by Gilboa/Schmeidler (1989).

An example of a choice-rule proposed in the literature that satisfies ISO and RDI but violates

CIR is Levi's "two-tier rule"[4]:

$$\mathrm{LEVI}\,(X,P) = \mathrm{MM}\,(\mathcal{A}(X,P)),$$
$$\text{where } \mathcal{A}\,(X,P) = \{\, x \in X \mid \exists\, p \in P \quad \forall y \in X: \quad x \cdot p \;\geq\; y \cdot p \,\}\,.$$

LEVI differs from MMEU in the following problem

$(X^0, P^0) \in \mathbb{X}^3 \times \mathbb{P}^3$, for instance, with

$X^0 = \mathcal{CH}\,\{\,(2,0,0),(1,1,-1)\,\}$,

$P^0 = \mathcal{CH}\,\{\,(1-\varepsilon,0,\varepsilon),(0,1-\varepsilon,\varepsilon)\,\}$, with $0 < \varepsilon < ½$.

$\mathcal{A}\,(X^0,P^0) = X^0$, hence $\mathrm{LEVI}\,(X^0,P^0) = (2,0,0)$.

On the other hand, for all $x \in X$:

$(0,1-\varepsilon,\varepsilon) \in \operatorname{argmin}\{\, p \cdot x \mid p \in P \,\}$; hence $\mathrm{MMEU}\,(X^0,P^0) = (1,1,-1)$.

To derive CIR, a new axiom is needed:

DEF Expected-Utility Equivalence (EUE):

For any $(X,P), (X',P) \in \mathbb{D}^S$ such that, for some mapping θ from X to X',

$\forall x \in X,\ \forall p \in P : p \cdot x = p \cdot \theta(x)$:

$x \in C(X,P) \Leftrightarrow \theta(x) \in C\,(X',P)$.

In words: Replacement of acts x by other acts $\theta(x)$ that are equivalent in terms of expected utility, under any acceptable probability $p \in P$, should affect the optimal choice only by the corresponding replacement θ.

Proposition 2: ISO & RDI & EUE $\Leftrightarrow$ CIR & RI .

The proof comes in four parts:

Part 1: ISO & RDI $\Rightarrow$ RI .

This is implied by proposition 1. □

Part 2: ISO & RDI & EUE $\Rightarrow$ CIR.

Fix some $(X,P) \in \mathbb{D}^S$, $S \in \mathbb{F}$. By ISO, w.l.o.g. $S \in \mathbb{S}$.

Again by ISO it suffices to show for some $J \in S$ and $\beta : J \to \mathscr{E}(P)$ (with associated Ψ) that $C(X,P) = \{x \in X \mid \Psi(x) \in C(\Psi(X), \Sigma^J)\}$.

We proceed by constructing three auxiliary problems and exploit appropriate invariance-conditions to determine their solution given $C(X,P)$; the last problem in the chain will be a CI-reduction of $(X,P), (\Psi(X), \Sigma^J)$.

[4]As in Levi (1980), ch.7.

Step 1: Take any $S' \in \mathbb{S}$ with $\# S' = \# S \cdot \# \mathscr{E}(P)$.
S' will be written as $S \times I$.
I will be used as a set of payoff-irrelevant background states. Its function is to allow the specification of a belief-set Q on $S \times I$ of a particular structure whose marginal Q^S on S is P. The belief set Q to be defined can be charaterized by two conditions:

i) For all $i \in I$, there is agreement among the $q \in Q$ that the conditional probability on $\{ S \times i \}$ is $\beta(i)$.

ii) There is complete ignorance about which i obtains.

Q can thus be interpreted as "complete-ignorance mixture of the extremal probabilities of P".

Formally, define $\gamma : \mathbf{R}^S \rightarrow \mathbf{R}^{S'}$ by $\gamma(x)_{s,i} = x_s$, and let $X' = \gamma(X)$. Also, let $J = \{S \times \{i\}\}_{i \in I} \in \mathbb{F}$, and fix a bijection $\beta : J \rightarrow \mathscr{E}(P)$; β will also be referred to as a function of $i \in I$.
Define $q^i \in \Sigma^{S \times I}$ by $q^i_{s,i} = \beta(i)_s$, and $q^i_{s,j} = 0$ for $j \neq i$.
Let $Q_0 = \{q^i\}_{i \in I}$ and $Q = \mathcal{CH}\, Q_0$.
Thanks to ISO, the partition $\{ s \times I \}_{s \in S}$ can be identified with S.

Assertion 1: $x \in C(X, P) \Leftrightarrow \gamma(x) \in C(X', Q)$.
Verification:
During the proof of part 1, we shall refer a number of times to the following elementary mathematical fact[5]:

Fact 1: If P is a compact convex subset of $\mathbf{R}^S$, $P = \mathcal{CH}\, \mathscr{E}(P)$.

Clearly $(q^i)^S = \beta(i)$, and thus $Q_0^S = \mathscr{E}(P)$. 1)
Since $Q_0 = \mathscr{E}(Q)$ according to 9) below, it follows that also $\mathscr{E}(Q)^S = \mathscr{E}(P)$. 2)
Moreover, due to fact 1, $Q^S = (\mathcal{CH}\ Q_0)^S = \mathcal{CH}\, Q_0^S = P$. 3)
Thanks to 2) and 3), RDI can be applied (in combination with ISO) to yield
$x \in C(X,P) \Leftrightarrow \gamma(x) \in C(X',Q)$. 4)

Step 2: We now exploit the agreement about conditional probabilities to take conditional expectations using EUE.

Define $\theta : \mathbf{R}^{S \times I} \rightarrow \mathbf{R}^{S \times I}$ by $\theta(x)_{i,s} = \sum_{t \in S} \beta(i)_t\, x_{i,t}$

and set $X'' = \theta(X')$.

Assertion 2: $x' \in C(X', Q) \Leftrightarrow \theta(x') \in C(X'', Q)$.

---

[5]See Rockafellar (1970, p.167), Corollary 18.5.1.

Verification: For $q \in \mathscr{E}(Q)$, i.e. $q = q^i$:

$$q^i \cdot \theta(x) = \sum_{s \in S} \beta(i)_s\, \theta(x)_{i,s} = (\sum_{s \in S} \beta(i)_s)(\sum_{t \in S} \beta(i)_t\, x_{i,t}) = q^i \cdot x. \qquad 5)$$

By the linearity of expectation and fact 1, $\forall q \in Q,\ x \in \mathbf{R}^{S \times I}: q \cdot \theta(x) = q \cdot x.$ 6)

EUE yields $x' \in C(X', Q) \Leftrightarrow \theta(x') \in C(X'', Q).$ 7)

Step 3: The taking of conditional expectations in step 2 has made $s \in S$ payoff-irrelevant, i.e. $X'' \in \mathbb{X}_J^{S \times I}$. (X'', Q) can thus be redescribed in terms of J as CI-problem in expected utilities. Setting $Y = \nu^J(X'')$, we get from RDI

Assertion 3: $x \in C(X'', Q) \Leftrightarrow \nu^J(x) \in C(Y, \Sigma^J).$

Verification: Clearly, $Q_o^J = \mathscr{E}(\Sigma^J).$ 8)

By the linearity of the mapping $\nu^J : q \mapsto q^J$, $\# \mathscr{E}(Q) \geq \# \mathscr{E}(Q^J) = \# J.$

From $\mathscr{E}(Q) \subset Q_o$, it follows that in fact $\mathscr{E}(Q) = Q_o$, 9)

hence that $\mathscr{E}(Q)^J = \mathscr{E}(\Sigma^J)$ from 8), and 10)

therefore also $Q^J = \Sigma^J$ by fact 1. 11)

10) and 11) allow to apply RDI again, yielding

$$x \in C(X', Q) \leftrightarrow \nu^J(x) \in C(Y, \Sigma^J). \qquad 12)$$

Step 4: Now $\nu^J \circ \theta \circ \gamma = \Psi : \mathbf{R}^S \to \mathbf{R}^J$, with $\Psi(x) = (\sum_{s \in S} \beta(j)_s\, x_s)_{j \in J}$;

it follows that $Y = \Psi(X)$.

Combining the assertions 1,2, and 3, we obtain

$$x \in C(X, P) \Leftrightarrow \Psi(x) \in C(Y, \Sigma^J). \quad \square \qquad 11)$$

Part 3: i) CIR $\Rightarrow$ EUE .

ii) CIR $\Rightarrow$ ISO .

i) is true simply because any CI-reduction of (X, P) is also a CI-reduction of (X', P) if X and X' satisfy the presupposition of EUE.

The analogous argument works for ISO as well. $\square$

Part 4: CIR & RI $\to$ RDI .

Take any $(X, P) \in \mathbb{D}^S$ such that $X \in \mathbb{X}_T^S$ and $\mathscr{E}(P)^T = \mathscr{E}(P^T)$. It has to be shown that $C(X, P)^T = C(X^T, P^T)$.

Let (Y, Σ^I) be a CI-reduction of (X, P) under β and Ψ, with $Y = \Psi(X)$.

Let J be the partition of I into equivalence classes [i] defined by the following equivalence relation $\sim$:

$$i \sim j \text{ iff } \beta(i)^T = \beta(j)^T. \qquad 1)$$

Lemma 2: i) $Y \subset \mathbf{R}^I_J$.

ii) (Y^J, Σ^J) is the CI-reduction of (X^T, P^T) under (β', Ψ') defined by $\beta' : [i] \mapsto \beta(i)^T$ and $\Psi' : x \mapsto (\beta(i)^T \cdot x)_{[i] \in J}$.

iii) For all $x \in \mathbf{R}^S_T$: $\Psi'(x^T) = \Psi(x)^J$.

Proof:

i) For $y = \Psi(x) \in Y$ and $i \sim j$: $y_i = y_j$, because $y_i = \beta(i)^T \cdot x = \beta(j)^T \cdot x = y_j$: the outer equalities are true by definition, the inner follows from 1) and the fact that any $x \in X$ is constant by assumption within elements of T. Hence $Y \subset \mathbf{R}^I_J$.

iii) For any $X \in \mathbf{R}^S_T$: $\Psi'(x^T) = (\beta'([i]) \cdot x^T)_{[i] \in J} = (\beta(i) \cdot x)^J_{i \in I} = \Psi(x)^J$.

ii) By iii), $\Psi'(X^T) = \Psi(X)^T = Y^J$.
By definition β' is one-to-one and $\beta'(J) = \mathscr{E}(P)^T$; since by assumption $\mathscr{E}(P)^T = \mathscr{E}(P^T), \beta'(J) = \mathscr{E}(P^T)$. □

Proposition 1 and Part 3, ii) imply CIR & RI $\Rightarrow$ RDI ^{CI}; we can therefore conclude that

$x \in C(X,P)$	iff $\Psi(x) \in C(Y, \Sigma^I)$	by CIR
	iff $\Psi(x)^J \in C(Y, \Sigma^I)^J$	by lemma 2,i)
	iff $\Psi(x)^J \in C(Y^J, \Sigma^J)$	by RDI^{CI} and i)
	iff $\Psi'(x^T) \in C(Y^J, \Sigma^J)$	by iii)
	iff $x^T \in C(X^T, P^T)$	by CIR and ii). ■

The role of CIR is summarized by

Lemma 3: A choice-function C on $\mathbb{D}^{CI}$ can be extended to $\mathbb{D}^{\sim}$ satisfying CIR if and only if C satisfies ISO on $\mathbb{D}^{CI}$. The extension is unique.

Proof: Uniqueness is trivial.

The "only-if" part holds because "isomorphic" CI-problems are CI-reductions of each other. The "if" part follows from the fact that CI-reductions of the same problem must be isomorphic to each other. □

Proposition 2 and lemma 3 imply directly

Proposition 3: A choice-function C on $\mathbb{D}^{CI}$ can be extended to $\mathbb{D}^{\sim}$ satisfying ISO, RDI and EUE if and only if C satisfies ISO and RI on $\mathbb{D}^{CI}$.
The extension is unique.

The extension of solutions for CI-problems to "CI-mixtures of probabilities" is nothing new. It has been proposed by Hurwicz (1951) and Milnor (1954) and has been axiomatized by Cohen and Jaffray (1985). Whereas Cohen and Jaffray extend CI-solutions in one step with the help of a "conditional-preference axiom", we decompose this extension into two steps (steps two and three of part 1), avoiding reference in the axioms to either preferences or conditional probabilities and without relying on any separability argument.

But the key innovation is step 1 in which general decision problems with belief sets are interpreted as problems with complete ignorance about the extremal probabilities.

The extension to CI-mixtures has a certain obviousness; in our treatment (steps 2 and 3) this is reflected in the fact that only a weak version of RDI is being used, in which the mapping $p \mapsto p^F$ is required to be invertible on P ("RDI^{1-1}"). On the other hand, step 1 utilizes Redescription Invariance in its strong general form RDI, and has therefore much more meat. Whereas RDI^{1-1} fails to restrict CI-solutions, even combined with ISO, RDI has a very strong implication for CI-problems, Replication Invariance.

A different line of attack is Jaffray's (1988, 1989) "mixture-approach", to our knowledge the only work that attempts to develop a reasonably general axiomatic decision theory for belief sets. It can be seen as "dual" to our approach in step 1, by interpreting decision-problems as mixtures of CI-problems rather than as CI-mixtures of probabilistic problems as we do.

Jaffray's approach is limited in two major ways: it can deal only with a special class of belief sets, those characterizable on terms of "belief-functions"; this class does not include CI-mixtures of probabilities, among others! Moreover, the theory has to be formulated in terms of preferences rather than choice-functions; this prejudices the search for a rational CI-solution decisively, as we argue in Nehring (1991a) which develops a theory of "Simultaneous Expected Utility Maximization" that violates standard choice-consistency conditions; it is shown there that any candidate for a rational solution has to do so.

REFERENCES

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

Bewley, T.F. (1986). "Knightian Decision Theory, Part I"; *Cowles Foundation Discussion Paper* **807**.

Cohen, M. and J.Y. Jaffray (1985). "Decision Making in a Case of Mixed Uncertainty: A Normative Model"; *Journal of Mathematical Economics*, **18**, (141-153).

Gaerdenfors, P. and Sahlin, N.E. (1982). "Unreliable Probabilities, Risk Taking, and Decision Making"; *Synthese*, **53**, (361-386).

Gilboa, I. and D. Schmeidler (1989). "Maxmin Expected Utility with a Non-Unique Prior"; *Journal of Mathematical Economics*, **18**, (141-153).

Hurwicz, L. (1951). "Some Specification Problems and Applications to Econometric Models"; *Econometrica*, **19**, (343-344.)

Jaffray, J.Y. (1988). "Linear Utility Theory for Belief Functions: A Discussion"; mimeo.

Jaffray, J.Y. (1989). "Linear Utility Theory for Belief Functions"; *Operations Research Letters*, **9**, (107-112).

Levi, I. (1980). *The Enterprise of Knowledge*. [MIT Press: Cambridge, MA].

Luce, R.D. and H. Raffia (1957). *Games and Decisions*. [John Wiley: New York].

Milnor, J. (1954). "Games Against Nature", in Thrall, Coombs and Davis (eds.), *Decision Processes*. [John Wiley: New York] (49-59).

Nehring, K. (1991a). "A Theory of Simultaneous Expected Utility Maximization with Vague Beliefs"; Chapter 1 of Ph.D. dissertation, "A Theory of Rational Decision with Vague Beliefs"; Harvard University.

Nehring, K. (1991b). "On the Logical Structure of Belief Sets"; Chapter 3 of Ph.D. dissertation, Harvard University.

Rockafellar, R.T. (1970). *Convex Analysis*, Princeton University Press.

Savage, L.J. (1972). *The Foundations of Statistics*, second revised edition. [Dover: New York].

Smith, C.A.B. (1961). "Consistency in Statistical Inference and Decision"; *Journal of the Royal Statistical Society, Series B*, **22**, (1-25).