The Comonotonic Sure-Thing Principle

CHEW SOO HONG
Department of Economics, University of California, Irvine, California

PETER WAKKER
Medical Decision Making Unit, University of Leiden, Leiden, The Netherlands

Abstract

This article identifies the common characterizing property, the comonotonic sure-thing principle, that underlies the rank-dependent direction in non-expected utility. This property restricts Savage's sure-thing principle to comonotonic acts, and is characterized in full generality by means of a new functional form—cumulative utility—that generalizes the Choquet integral. Thus, a common generalization of all existing rank-dependent forms is obtained, including rank-dependent expected utility, Choquet expected utility, and cumulative prospect theory.

Key words: Choquet integral, capacity, utility, sure-thing principle

1. Introduction

Expected utility has been the received theory of individual decision making under risk, where probabilities are known beforehand, and under uncertainty, where probabilities are not known beforehand. Its validity was however called into question due to the paradoxes of Allais and Ellsberg. These paradoxes have led to the development of various alternative theories. There are two main directions in the literature on axiomatic non-expected utility theories—the rank-dependent direction and the betweenness direction. Betweenness, the common characterizing property of the latter direction, requires that the preference for a probability mixture of two lotteries be between those two lotteries. The betweenness direction originated with weighted utility theory (Chew, 1983; Fishburn, 1983), followed by Dekel (1986), Hazen (1987), Chew (1989), Gul (1991), and others. The betweenness literature has concentrated on decision making under risk, and does not provide an account of the Ellsberg paradox.

We focus on the other main direction, the rank-dependent direction, that originated with Quiggin (1982) and Schmeidler (1989). The rank-dependent theories have a number of appealing features. While expected utility has often been criticized for modeling risk attitudes solely through the valuation of outcomes, the rank-dependent models incorporate the influence of the perception of risk and uncertainty. Further, they have tractable functional forms and explain both the Allais and the Ellsberg paradoxes.

Savage's sure-thing principle is directly violated by the Allais paradox in the context of risk, and by the Ellsberg paradox in the context of uncertainty. This article proposes a weakening of Savage's sure-thing principle, called the comonotonic sure-thing principle

or comonotonic independence axiom, which essentially restricts Savage's sure-thing principle to comonotonic acts, i.e., acts that induce the same ordering on the states of nature in terms of the associated outcomes. We will use both labels interchangeably noting that the comonotonic sure-thing principle label is clearer in linking its content with the sure-thing principle, while comonotonic independence is simpler and avoids the connotation associated with the term "sure-thing."

We identify the comonotonic independence axiom as the common characterizing property that underlies the rank-dependent direction. We characterize comonotonic independence in full generality by means of a new functional form called cumulative utility (CU). The axiomatizations currently provided in the literature suppose various conditions, none directly weakening Savage's sure-thing principle, and all stronger than comonotonic independence. The derivations of most of those results can be simplified through the results of this article; see table 1 in section 5. In particular, by dropping separability of belief and utility, CU generalizes Choquet expected utility, as well as cumulative prospect theory (rank- and sign-dependent utility theory). A more detailed discussion of the literature is provided in section 5.

The rank-dependent generalization of Savage's expected utility to Choquet expected utility became possible when a proper generalization of additive integrals was understood, in Quiggin (1982) for the special case of decision making under risk, and, independently, in Schmeidler (1989) for the general case. This generalization is the Choquet integral introduced by Choquet (1954). In section 4, the form introduced in this article, the CU-functional, is shown to be a genuine generalization of the Choquet integral, satisfying natural conditions such as monotonicity, linearity within comonotonic subsets, and sup-norm continuity. The latter is less restrictive than weak continuity, popularly used in decision making under risk and in statistics, and yields convenient necessary and sufficient conditions in the main theorems; see Observation 4.2. It only employs structure on the outcomes, and does not impose any restriction on the state space, which may be finite, infinite, contain atoms, or be atomless.

Section 2 presents some notation. Section 3 presents the representation theorem for simple acts. The main theorem, dealing with nonsimple acts, is given in section 4. Section 5 relates our result to the literature. Table 1 illustrates that the CU-characterization generalizes not only the existing rank-dependent characterizations for uncertainty, where no rank-linear forms had as yet been axiomatized, but also generalizes existing results for the context of risk. For risk, rank-linear forms had been proposed before. Besides generalizing the existing results, we provide a direct proof that avoids some mathematical problems that occurred in the literature; these are discussed or referenced in section 5. Appendix A gives proofs and appendix B generalizes the results to connected topological spaces.

2. Notation

Let Ω be the *state space*, endowed with an algebra $\mathcal{A}$. Elements of $\mathcal{A}$ are called *events*. As *consequence space* we take the real line $\mathbf{R}$. (The appendix will generalize to connected topological spaces, e.g., any convex subset of $\mathbf{R}^m$.) F is the set of all *acts*, i.e., measurable

functions (the inverse of every interval is an event) from Ω to $\mathbf{R}$. We denote by $f_{-A}x$ the act assigning x to all $\omega \in A$ and $f(\omega)$ to all $\omega \in A^c$, and by $f_{-A,B}x,y$ the act assigning x to all $\omega \in A$, y to all $\omega \in B$, and $f(\omega)$ to all $\omega \in A^c \cap B^c$. Here A,B are disjoint events. A function $v : \mathcal{A} \to [0,1]$ is a *capacity* if $v(\varnothing) = 0$, $v(\Omega) = 1$, and $v(A \cup B) \geq v(A)$ for all $A,B \in \mathcal{A}$; capacities are nonadditive generalizations of probability measures.

$\Pi = (A_1, \ldots, A_n)$ is an *ordered partition*, or *partition* for short (the usual unordered partitions will not occur in this article), if $A_1, \ldots, A_n$ are mutually exclusive and exhaustive subsets of Ω. It is always assumed that the A_j's are events. We write $\cup^i(\Pi)$ for $\cup_{j=1}^i A_j$; $\cup^0(\Pi) = \varnothing$; we often suppress Π. F^{Π} is the set of acts of the form $\sum_{j=1}^n x_j 1_{A_j}$ (1_A is the indicator function); whenever we use this notation, it is implicitly understood that the A_j's are disjoint. Note that the set F^{Π} does not depend on the ordering of events in Π. $F_{\downarrow}^{\Pi}$ is the subset of F^{Π} for which $x_1 \geq \cdots \geq x_n$; here obviously the ordering of events in Π does matter. The set F^s of *simple* acts is the union of F^{Π} over all partitions Π; F^b denotes the set of all bounded acts.

We are concerned with a preference relation $\succcurlyeq$ on a set of acts; the exact domain of $\succcurlyeq$ will be specified below. A function φ *represents* $\succcurlyeq$ if φ is a function from the domain of $\succcurlyeq$ to $\mathbf{R}$ such that $[f \succcurlyeq g \Leftrightarrow \varphi(f) \geq \varphi(g)]$. If a representing function exists, then $\succcurlyeq$ is a *weak order*, i.e., it is *complete* $[\forall f, g : f \succcurlyeq g$ or $g \succcurlyeq f]$ and transitive.

3. Cumulative utility for simple acts

This and the following section characterize CU theory in terms of properties of $\succcurlyeq$. CU employs an *outcome-dependent capacity*, i.e., a function $W : \mathbf{R} \times \mathcal{A} \to \mathbf{R}$ satisfying (i) $\forall A \in \mathcal{A} : W(0,A) = 0$, (ii) $\forall x \neq 0$, $W(x,\Omega) \neq 0$, and (iii) $\forall x \neq 0$, $W(x,\cdot)/W(x,\Omega)$ is a capacity. In this section, we restrict attention to the set F^s of simple acts. A representing function on F^s is a *cumulative utility* (*CU*) functional if there exists an outcome-dependent capacity W such that the function is of the form

$$f = \sum_{j=1}^n x_j 1_{A_j} \mapsto \sum_{j=1}^n [W(x_j, \cup^j) - W(x_j, \cup^{j-1})] . \tag{1}$$

The most well-known special case is subjective expected utility (SEU). It corresponds to $W(x,A) = P(A)u(x)$, where P is a probability measure and u a utility function. An intermediate special case is Choquet expected utility (CEU), where $W(x,A) = v(A)u(x)$, with v being a capacity instead of a probability measure. Note the central role of events of the form $\cup^j$, describing the occurrence of a fixed outcome x_j or anything better. That is, these events are "cumulative." This explains the term cumulative utility.

Throughout this article, for an outcome-dependent capacity W,

$$V(y, B, A) := W(y, B) - W(y, A).$$

The condition $B \supset A$ is implicit in the notation $V(y,B,A)$. With this notation, (1) can be rewritten as

$$\sum_{j=1}^{n} V(x_j, \cup^j, \cup^{j-1}).\tag{2}$$

One may interpret V as an event-dependent utility, i.e., the natural rank-dependent analogue of state-dependent utility. To generalize Choquet expected utility (the case where $W(x,A) = v(A)u(x)$ above) to incorporate state-dependence of utility, one can write $W(x,A) = v(A)u_A(x)$, where $u_A(\cdot)$ now depends on the event A. Then, however, with preferences over acts as the only empirical primitive, the factorization $v(A)u_A(\cdot)$ becomes meaningless. Only the product $W(x,A) = v(A)u_A(x)$ can be derived from preferences. Thus, the result of this article may be interpreted as a generalization of CEU that allows for state-dependent utility. Under CU, the outcome-dependent capacity and the functional in (1) on F^s are said to be *associated with each other*.

Acts f and g are *comonotonic* if there exist no states ω, ω' such that $f(\omega) > f(\omega')$ and $g(\omega) < g(\omega')$. That is, f and g order the states of nature the same way in terms of the appeal of the assigned outcomes. A set of acts is *comonotonic* if the acts are pairwise comonotonic. We say that $\geqslant$ satisfies *comonotonic independence* (*comonotonic sure-thing principle*) on F^s if

$$f_{-A}\, x \geqslant g_{-A}\, x \Leftrightarrow f_{-A}\, y \geqslant g_{-A}\, y\tag{3}$$

for all $f_{-A}\, x$, $g_{-A}\, x$, $f_{-A}\, y$, $g_{-A}\, y$ which are comonotonic. Observe that if the latter acts are not required to be comonotonic, then (3) is equivalent to Savage's sure-thing principle. Comonotonic independence weakens Savage's sure-thing principle by requiring the four involved acts to be comonotonic. It is apparent that comonotonic independence is compatible with both the Allais paradox and the Ellsberg paradox. Motivations have been provided in Yaari (1987), Schmeidler (1989), Wakker (1990a), and other articles. Schmeidler (1989) uses the term comonotonic independence for independence with respect to probabilistic mixing; in his model, consequences are probability distributions over a set of prizes. Our condition is the natural weakening of his condition to the setting in which a probabilistic mixture operation is not available. It can be seen (Wakker, 1989, Lemma VI.3.3) that the simple acts in (3) are comonotonic if and only if they are contained in one set $F_{\downarrow}^{\Pi}$.

The following lemma prepares for the main theorem. Its proof is given immediately because it illustrates the CU functional. Together with Corollary 1, it shows that comonotonic independence underlies all rank-dependent models.

Lemma 1. $\geqslant$ satisfies comonotonic independence on F^s whenever it can be represented by a CU functional.

Proof. Let $\sum_{i \neq j} f_i 1_{A_i} + x 1_{A_j} \geq \sum_{i \neq j} g_i 1_{A_i} + x 1_{A_j}$. Further, let $f_1 \geq \cdots \geq f_{j-1} \geq x \geq f_{j+1} \geq \cdots \geq f_n$, the same with g instead of f, and/or with y instead of x. Substitute

$$\sum_{i \neq j} V(f_i, \cup^i, \cup^{i-1}) + V(x, \cup^j, \cup^{j-1}) \geq \sum_{i \neq j} V(g_i, \cup^i, \cup^{i-1}) + V(x, \cup^j, \cup^{j-1}),$$

and replace the common term $V(x, \cup^j, \cup^{j-1})$ by $V(y, \cup^j, \cup^{j-1})$. Note that this reasoning could only be applied because the acts are comonotonic, and so can be related to one same set $F_\downarrow^\Pi$. $\qquad \square$

Another way to understand the result of the above lemma is to observe that CU gives an additive representation within each set $F_\downarrow^\Pi$; see (2). That implies the usual independence within each $F_\downarrow^\Pi$, i.e., comonotonic independence for simple acts.

We say that $\geq$ satisfies *pointwise monotonicity* on a general set of acts G, if for all acts $f, g \in G : [\forall \omega \in \Omega : f(\omega) \geq g(\omega)] \Rightarrow f \geq g$. For simple acts however, we need to develop a stronger definition of monotonicity. An event A is *inessential* if $f \sim g$ for all simple acts f and g that coincide on A^c. *Essential* is the opposite of inessential. We call $\geq$ *jointly monotonic* on F^s if, for all simple acts f and essential events A, $[y > x \Rightarrow f_{-A} y > f_{-A} x]$. In CU, joint monotonicity implies monotonicity with respect to outcomes as well as with respect to event inclusion; in decision making under risk, it is equivalent to stochastic dominance (see section 5). Pointwise monotonicity follows on F^s as an elementary consequence of joint monotonicity, transitivity, and completeness.

Continuity will be related to the supremum norm: $\geq$ *is continuous* on F^s if for each $f > g$ in F^s there exists $\delta > 0$ such that for all h in F^s : $([\forall \omega \in \Omega : |f(\omega) - h(\omega)| < \delta] \Rightarrow h > g)$, and $([\forall \omega \in \Omega : |h(\omega) - g(\omega)| < \delta] \Rightarrow f > h)$.

For an outcome-dependent capacity $W : \mathbf{R} \times \mathcal{A} \to \mathbf{R}$ and the CU functional on F^s associated with W, we use the following analogous conditions. W satisfies *joint monotonicity* if $y > x \Rightarrow V(y, B, A) > V(x, B, A)$ whenever $B \backslash A$ is essential. CU is *continuous* on F^s if for each $f \in F^s$ and each $\epsilon > 0$ there exists $\delta > 0$ such that for all $h \in F^s$: $[\forall \omega \in \Omega : |f(\omega) - h(\omega)| < \delta] \Rightarrow |CU(f) - CU(h)| < \epsilon$. Note that this implies continuity of $W(\cdot, A)$ for each A. Finally, a function is a ratio scale if it is unique up to a positive factor. The following theorem characterizes CU on the simple acts.

Theorem 1. Suppose there exist three disjoint essential events. The following two statements are equivalent for $\geq$ on F^s.

(i) $\geq$ is represented by

$$CU : \sum_{j=1}^n x_j 1_{A_j} \mapsto \sum_{j=1}^n \{ W(x_j, \cup_{i=1}^j A_i) - W(x_j, \cup_{i=1}^{j-1} A_i) \}$$

for $x_1 \geq \cdots \geq x_n$, where $W : \mathbf{R} \times \mathcal{A} \to \mathbf{R}$ is an outcome-dependent capacity that is continuous in the first argument and jointly monotonic.

(ii) $\geq$ satisfies transitivity, completeness, joint monotonicity, (supnorm-) continuity, and comonotonic independence (comonotonic sure-thing principle).

Further, given (i) or (ii), CU is (supnorm-)continuous, and the functions W and CU are ratio scales.

A detailed proof of the theorem is given in the appendix. We give here some comments on the proof of the most difficult implication, $(ii) \Rightarrow (i)$. (ii) implies that within each comonotonic subset $F^{[]}_\downarrow$, the regular sure-thing principle holds. This principle is identical to the independence condition proposed in Debreu (1960), shown to be necessary and sufficient for additive representation on full product sets. Here we need the extension of Debreu's result to comonotonic sets. This extension is nontrivial, and was provided by Wakker (1993a). Next the additive representations on comonotonic subsets must be "glued together" into one overall CU-functional that is defined on the set F^s of simple acts. For example, the additive representations must be made to coincide on the diagonal of constant acts; the required derivation turns out to be rather complicated. The "gluing together" of additive representations within different comonotonic subsets, the main problem in the proof, is trivially satisfied in the context of decision making under risk. It shows why our result for uncertainty is more complicated than existing results for risk.

4. Cumulative utility for nonsimple acts

This section extends the results of the previous section to F^b, the set of bounded acts. The definitions of comonotonicity, as well as continuity for $\geq$ and functionals CU, are extended to F^b in a straightforward manner. For simplicity, the definitions of (in)essentiality (and joint monotonicity as based on it) will not be changed, i.e., they are only related to simple acts. (In the presence of proper monotonicity, this can be seen to be equivalent to alternative definitions of (in)essentiality, related to larger sets of acts.) Joint monotonicity of $\geq$ will only be imposed on F^s; whether Statement (i) in Theorem 2 implies joint monotonicity of $\geq$ on the entire set F^b is an open question. *Comonotonic independence* holds on F^b if (3) holds for all comonotonic acts in F^b. For the extension to nonsimple acts, certainty equivalents are important; for an act, the *certainty equivalent* is the constant act/outcome equivalent to that act.

Lemma 2. Suppose $\geq$ satisfies joint monotonicity on F^s, and transitivity, completeness, and continuity on F^b. Then, for each $f \in F^b$, there exists a certainty equivalent.

If a CU functional as characterized in Theorem 1 is given on F^s, then, under the conditions of the above lemma, we can, for any $f \in F^b$, define $CU(f)$ as the CU-value of the certainty equivalent of f. Obviously this definition will only be useful if CU is a genuine generalization of integrals. That is, given boundedness, $CU(f)$ should be the limit of CU values of converging simple acts. This is guaranteed by continuity, together with the

well-known denseness (with respect to the supnorm) of the simple acts in F^b. As a consequence (derived below), CU-on-F^b inherits *pointwise monotonicity*, natural for (generalized) integrals, from CU-on-F^s : $[\forall \omega : f(\omega) \geq g(\omega)] \Rightarrow CU(f) \geq CU(g)$. Another property that CU-on-F^b inherits from CU-on-F^s is comonotonic independence for the represented preference relation. This is crucial as comonotonic independence is at the core of the rank-dependent stream, and of this article. The property follows directly from the following lemma.

Lemma 3. Suppose that $CU : F^b \to \mathbf{R}$ is continuous, and is associated with an outcome-dependent capacity on F^s. Then CU satisfies pointwise monotonicity, and $CU(f) - CU(g) = CU(f') - CU(g')$ for all comonotonic acts f, g, f', g' satisfying: $\forall \omega \in \Omega$,

$$[f(\omega) = g(\omega), f'(\omega) = g'(\omega)] \text{ or } [f(\omega) = f'(\omega), g(\omega) = g'(\omega)]. \tag{4}$$

Actually, a somewhat stronger condition than comonotonic independence follows for a preference relation represented by a CU-functional as in the above lemma. It says that, given comonotonicity, not only one or finitely many common outcomes can be changed without affecting preference, but rather any arbitrary number:

Corollary 1. Under the conditions of Lemma 3, the binary relation $\succcurlyeq$, represented by CU, satisfies the following condition: For all comonotonic f, g, f', g' satisfying (4), $[f \succcurlyeq g \Leftrightarrow f' \succcurlyeq g']$. Consequently, $\succcurlyeq$ satisfies comonotonic independence.

Sometimes in the literature stronger formulations for independence conditions, such as in the above corollary, have been used, where an infinite number of common outcomes may be replaced. Also Savage (1954) used this stronger formulation for his sure-thing principle. The above corollary shows, in combination with Observation 1 below, that the alternative definition and ours are equivalent for continuous CU, thus for most rank-dependent models in the literature; also in Savage (1954), restriction of the sure-thing principle to simple acts would have sufficed to imply it in general (see Wakker, 1993b).

Theorem 2. Suppose there exist three disjoint essential events. The following two statements are equivalent for $\succcurlyeq$ on F^b, the set of bounded acts:

(i) $\succcurlyeq$ is represented by a (supnorm-)continuous functional CU. On F^s, CU is associated with a jointly monotonic outcome-dependent capacity W, i.e.,

$$CU : \sum_{j=1}^{n} x_j 1_{A_j} \mapsto \sum_{j=1}^{n} \{W(x_j, \cup_{i=1}^{j} A_i) - W(x_j, \cup_{i=1}^{j-1} A_i)\}$$

for $x_1 \geq \cdots \geq x_n$.

(ii) $\geqslant$ satisfies joint monotonicity on F^s, transitivity, completeness, (supnorm-)continuity, and comonotonic independence (comonotonic sure-thing principle).

Next some observations and modifications are given.

Observation 1. The following observations apply to Theorem 2:

- The functions W and CU in (i) are ratio scales.
- In (i), $W : \mathbf{R} \times \mathcal{A} \to \mathbf{R}$ is continuous in outcomes.
- Both in statement (i) and statement (ii), pointwise monotonicity is implied by the other conditions. Continuity uniquely determines the extension of CU from F^s to F^b.
- Comonotonic independence can be weakened to hold only on F^s; alternatively, it can be strengthened to the condition in Corollary 1.
- Instead of F^b, any set of acts X with $F^s \subset X \subset F^b$ could have been taken as domain.

The next observation is another illustration that supnorm continuity is not very restrictive.

Observation 2. Under SEU and CEU, (supnorm-)continuity of the representing functional is equivalent to continuity of the utility function.

The representation of Theorem 2 can be extended to unbounded acts with finite CU-value, in a way similar to definitions of integrals. Supnorm-continuity then is to be restricted to F^b. For act f, the *truncation* f_υ^μ is υ if f is less/equal υ, it is f if f is between υ and μ, and it is μ if f is greater/equal μ. For any unbounded act f, the usual integrals are defined as limits of integrals of truncations of f. Analogously, if the limit of the certainty equivalents of the truncations f_υ^μ (υ going to $-\infty$, μ to ∞) exists and is real, then we can extend the definition of the CU functional to f by defining $CU(f)$ as the CU-value of the mentioned limit of certainty equivalents. Here the condition must be imposed on the preference relation that f is equivalent to the mentioned limit of certainty equivalents. Representations of unbounded acts in rank-dependent models have been discussed in Chew and Epstein (1989a, end of section 4), Wakker (1993b), and Nakamura (1992b).

5. Specializations of cumulative utility

This section presents the implications of Theorem 2 for the recent literature on rank-dependent preferences. We first show how, in general, decision making under risk can be considered a special case of decision making under uncertainty in which probabilities are known beforehand. In this case, a probability measure P on Ω is given, and each act f induces a probability distribution $F_f : x \mapsto P(f \leq x)$ on the outcome space $\mathbf{R}$. An act f *stochastically dominates* another act g if $F_f(x) \leq F_g(x)$ for all outcomes x. We require $\geqslant$ to satisfy *stochastic dominance* in the sense that $f \geqslant g$ whenever f stochastically dominates g, and $f > g$ if furthermore $F_f \neq F_g$. It follows that different acts inducing the same

probability distribution over outcomes are indifferent. It is then customary not to make explicit the underlying state space but rather to assume that it is rich enough to generate all (countably additive) probability distributions on **R**. An example of such a state space is [0,1] endowed with the uniform distribution on the Borel subsets. Then, as is well-known, for any probability distribution, there exists a random variable/act that generates it.

Table 1 presents the rank-dependent forms for simple acts $\sum_{j=1}^{n} x_j 1_{A_j}$ with $x_1 > \cdots > x_n$; the extension to general acts has been given in section 4. The forms to the right concern decision making under uncertainty in full generality; the forms to the left concern decision making under risk. In the latter case, the (objective) probabilities $P(A_j)$ are abbreviated as p_j.

CU was defined before; RLU (rank-linear utility) is the special case in which $W(x,A) = \psi(x,P(A))$ for a function ψ satisfying $\psi(x,0) = 0$, and for all $y > x$ and $q > p$,

$$\psi(y,q) - \psi(y,p) - \psi(x,q) + \psi(x,p) > 0. \tag{5}$$

Condition (5) is equivalent to stochastic dominance, and is the analogue of joint monotonicity in CU (see Lemma 4 below). CEU (Choquet expected utility) corresponds to the case where $W(x,A) = u(x)v(A)$, where v is a capacity; let us repeat that $\cup^j$ abbreviates $\cup_{i=1}^{j} A_i$. RDEU (rank-dependent expected utility) is the special case of CEU in which $v = \varphi \circ P$, for a known probability measure P and a strictly increasing function φ with $\varphi(0) = 0$, $\varphi(1) = 1$; stochastic dominance implies that u is strictly increasing. SEU (subjective expected utility) corresponds to CEU where the capacity v is additive, i.e., it is a (subjective) probability measure. EU (expected utility) is the special case of SEU in which the subjective probability measure coincides with a known objective probability measure P.

RDEU originated with Quiggin (1982). See also Yaari (1987), Chew (1990), Nakamura (1992a), and Wakker (1994). CEU was initiated by Schmeidler (1989; first version, 1982) with refinements by Gilboa (1987), Luce (1988), Nakamura (1990, 1992a), Wakker (1989, Theorem VI.5.1), Sarin and Wakker (1992), and Chew and Karni (1993). RLU was first proposed in Segal (1989, Theorem 1; first version, 1984) (see also Wakker, 1993c; Segal, 1993.) Green and Jullien (1988) gave sufficient conditions for this form. Chew and Epstein (1989a) gave necessary and sufficient conditions (see also Chew, Epstein, and Wakker, 1992). Our result is more general, because, first, it can be formulated for connected topological outcome spaces (see appendix B), and, second, the imposed continuity restrictions are weaker.

We obtain the following generalization of Wakker (1990b):

Table 1.

Risk	Uncertainty
EU: $\sum_{j=1}^{n} p_j u(x_j)$	SEU: $\sum_{j=1}^{n} v(A_j) u(x_j)$
RDEU: $\sum_{j=1}^{n} [\varphi(\sum_{i=1}^{j} p_i) - \varphi(\sum_{i=1}^{j-1} p_i)] u(x_j)$	CEU: $\sum_{j=1}^{n} [v(\cup^j) - v(\cup^{j-1})] u(x_j)$
RLU: $\sum_{j=1}^{n} [\psi(x_j, \sum_{i=1}^{j} p_i) - \psi(x_j, \sum_{i=1}^{j-1} p_i)]$	CU: $\sum_{j=1}^{n} [W(x_j, \cup^j) - W(x_j, \cup^{j-1})]$

Observation 3. All models from the left-hand side of Table 1 can be derived from those on the right-hand side: One relates probability distributions to random variables (acts) on the space [0,1] endowed with the uniform distribution, and invokes stochastic dominance.

The following lemma shows how RLU can be derived from CU. Thus, in conjunction with Theorem 2, the lemma provides a characterization of RLU. Here it should be noted that, for decision making under risk, stochastic dominance implies joint monotonicity of the preference relation, thus of the CU form. The requirement that P be atomless is necessary and sufficient for the possibility to generate all probability distributions on $\mathbf{R}$.

Lemma 4. Suppose P is an atomless countably additive probability measure on Ω. Under CU, the following two statements are equivalent:

(i) $\succcurlyeq$ satisfies stochastic dominance.
(ii) RLU holds.

Proof. (ii) immediately implies (i). So we suppose that (i) holds, and derive (ii). Under stochastic dominance, the equality $P(A) = P(B)$ implies, for all $x \geq y$, the indifference $x1_A + y1_{A^c} \sim x1_B + y1_{B^c}$. Substituting CU yields

$$W(x, A) - W(x, \varnothing) + W(y, \Omega) - W(y, A) = W(x, B) - W(x, \varnothing) + W(y, \Omega)$$

$$- W(y, B).$$

Substituting $y = 0$ and

$$W(x, \varnothing) = W(0, \Omega) = W(0,A) = W(x, \varnothing) = W(0,\Omega) = W(0,B) = 0$$

yields $W(x,A) = W(x,B)$ for all $x \geq 0$. Analogously, $W(y,A) = W(y,B)$ follows for $y < 0$ by substituting $x = 0$. So $W(\cdot,A)$ depends only on the probability of A, and $W(x,A) = \psi(x,P(A))$, for a function ψ. Mainly by stochastic dominance, ψ satisfies all required conditions. □

The ordinal independence condition of Green and Jullien (1988) is weaker than comonotonic independence. In Theorem 1, however, comonotonic independence can be weakened to the following condition, the equivalent of ordinal independence of Green and Jullien:

$$f_{-A} x \succcurlyeq g_{-A} x \Leftrightarrow f_{-A} y \succcurlyeq g_{-A} y \tag{6}$$

where either x and y are maximal outcomes, or minimal outcomes. This condition implies comonotonic independence in the presence of the other conditions of Theorem 2, by reasonings analogous to Gorman's (1968). (Further details will be given in Remark A1.[1])

Quiggin (1989, Proposition 4) and Ebert (1988a,b) also obtained such representations for simple probability distributions. Due to symmetry, in RLU, only nondecreasing random variables from [0,1] to **R** need to be considered, so that the additive representation only has to be derived on one maximal comonotonic set. The process in the proof of Theorem 1 of fitting together the additive representations from different comonotonic sets through certainty equivalents into one overall additive representation can thus be skipped for RLU. Hence its derivations are considerably simpler. (Wakker (1991, 1993a) pointed out mathematical problems for the derivations of RLU. Also, Chateauneuf (1990) characterized RLU.)

Cumulative prospect theory, introduced in Tversky and Kahneman (1992) is a special case of CU with $W(x,B) = v^+(B)u(x)$ for $x \geq 0$, and $W(x,B) = v^-(B)u(x)$ for $x < 0$. (See also Wakker and Tversky, 1993.) The preference functional used in cumulative prospect theory appeared also in Starmer and Sugden (1989) and Luce and Fishburn (1991). It modifies the earlier prospect theory of Kahneman and Tversky (1979) by adopting RDEU or CEU forms, and presupposes a status quo—the zero outcome—relative to which gains and losses are defined.

Within CU, we may extend the notion of a status quo to a normal region $[\upsilon,\mu]$ of outcomes within which behavior conforms to the prescriptions of expected utility. Only when extreme outcomes are involved, may the implications of expected utility be violated. This can be modeled by: For $x \in [\upsilon,\mu]$, $W(x,A) = u(x)P(A)$, i.e., coincides with expected utility; for $x < \upsilon$, $W(x,A) = u(x)v(A)$; for $x > \mu$, $W(x,A) = u(x)v'(A)$, where v and v' are capacities, $v(v')$ pertains to the perception of uncertainty when excessively low (high) outcomes are involved.

Next we suggest an application of the CU functional proposed in this article to robust statistics. There, the L-estimator corresponds to RDEU with $u(x) = x$ (Huber, 1981; Chew and Epstein, 1989b). A commonly used special case is the α-trimmed mean consisting of the arithmetic mean after trimming the α fractions of the highest and the lowest observations. One can be critical of trimmed means for indiscriminately discarding outliers, even when these turn out to be moderate. It seems prudent to determine a normal region $[\upsilon,\mu]$ and trim for outliers only if they are outside this region. Such an estimator corresponds to: For $x \in [\upsilon,\mu]$, $W(x,A) = xP(A)$, i.e., coincides with expected value, for $x < \upsilon$, $W(x,A) = x\varphi(P(A))$ where φ is continuous on [0,1], linear on $(\alpha,1]$, and vanishes on $[0, \alpha)$, for $x > \mu$, $W(x,A) = x\varphi'(P(A))$ where φ' is continuous on [0,1], linear on $[0,1 - \alpha]$, and equals 1 on $(1 - \alpha,1]$.

Appendix A: Proofs

Proof of Theorem 1.

SUPNORM-CONTINUITY OF CU: As a preparation we consider (supnorm-) continuity of CU.

Lemma A1. If W is jointly monotonic, then continuity of W in its first argument is equivalent to continuity on F^s of the functional CU associated with W.

Proof. If CU is continuous, then continuity of $W(x,A)$ in x follows from continuity of $CU(x1_A + y1_{A^c})$ in x, for $y < x$. Next suppose W is continuous in its first argument. To demonstrate the continuity of CU, let f^j be a sequence of simple acts, converging in supnorm to the simple act f. We shall only prove liminf $CU(f^j) \geq CU(f)$. We can write f as $\sum_{i=1}^{n} x_i 1_{A_i}$, with $x_1 > \cdots > x_n$. Let g^j be the simple act $\sum_{i=1}^{n} y_i^j 1_{A_i}$, where y_i^j is the minimum of x_i and $\inf\{f^k(\omega) : k \geq j, \omega \in A_i\}$. Then also these acts g^j converge in supnorm, monotonically from below, to f. By taking j large enough, $y_1^j > \ldots > y_n^j$ can be guaranteed, so that then all g^j are in the same $F_{\downarrow}^{\Pi}$ as f, with $\Pi = (A_1, \cdots, A_n)$. By continuity in x, for each i, of $W(x, \cup^i)$, consequently of each $V(x, \cup^i, \cup^{i-1})$, $CU(g^j)$ converges to $CU(f)$. Repeated application of joint monotonicity implies pointwise monotonocity of CU on F^s, so that $CU(f^j) \geq CU(g^j)$ for all j. Hence liminf $CU(f^j) \geq$ liminf $CU(g^j) = CU(f)$. □

The above lemma also shows that, on F^s, (supnorm-) continuity in the presence of joint monotonicity is equivalent to the usual (Euclidean) "finite-dimensional" continuities imposed within each F^{Π}, so is not very restrictive.

IMPLICATION $(i) \Rightarrow (ii)$: To derive $(i) \Rightarrow (ii)$ in Theorem 1, assume (i) holds. Comonotonic independence has been derived in Lemma 1. Transitivity and completeness are direct. Continuity of $\geqslant$ follows from the continuity of CU as established above. For joint monotonicity of $\geqslant$, consider the simple act $\sum_{i=1}^{n} x_i 1_{A_i} + z 1_A$, for $x_1 > \cdots > x_n$, A essential, and $(A_1, \ldots, A_n)$ an ordered partition of A^c. It suffices to show that the CU-value of this simple act is a strictly increasing function of z. On $(-\infty, x_n]$, it is because $V(z, A \cup (\cup''), \cup'')$ is, on $[x_{i+1}, x_i]$, because $V(z, A \cup (\cup^i), \cup^i)$ is, and on $[x_1, \infty)$, because $V(z, A, \varnothing)$ is.

IMPLICATION $(ii) \Rightarrow (i)$: Now we turn to the most difficult part of the proof. We suppose (ii) holds, and derive (i). Since $\geqslant$ satisfies the usual continuity and monotonicity conditions within each F^{Π}, we get:

Lemma A2. For each $f \in F^s$ there exists a certainty equivalent (defined in section 4).

We call one ordered partition $(A_1, \ldots, A_n)$ an *ordered refinement* of another ordered partition $(B_1, \ldots, B_m)$ if A_1 is $B_1 \cup B_2 \cup \cdots \cup B_{k1}$, A_2 is $B_{k1+1} \cup B_{k1+2} \cup \ldots \cup B_{k2}$, etc. Again, because this article only deals with ordered partitions, we shall suppress "ordered" everywhere. Note that thus (A_1, A_2, A_3) is not a refinement of $(A_2 \cup A_3, A_1)$. The following simple observations concerning essentiality of events can easily be verified:

If A is essential, then so is each superset.   (A1)

If A is essential, and $(A_1, \ldots, A_n)$ is a partition of A, then at least one A_i is essential.   (A2)

For each partition, there exists a refinement containing at least three essential events. (A3)

For (A3), intersect each element of the partition with each of the three disjoint essential events as required in Theorem 1, and with the complement of the union of those three events. Then apply (A2) with each of those three events in the role of A.

For every partition Π containing at least three essential events, $\geqslant$ restricted to $F_{\downarrow}^{\Pi}$ is a continuous monotonic weak order that satisfies usual independence (CI in Wakker, 1993a). Hence, by Corollary 3.6 of Wakker (1993a) (the increasingness condition there follows after suppression of inessential coordinates), there exists a cardinal additive representation $V^{\Pi}: \sum_{j=1}^{n} x_j 1_{A_j} \mapsto \sum_{j=1}^{n} V_j^{\Pi}(x_j)$ on $F_{\downarrow}^{\Pi}$. For each inessential A_j in Π, V_j^{Π} is constant, and can be taken to be 0; for each essential A_j, it is strictly increasing in outcome by joint monotonicity. By (A3) and standard uniqueness results (see Wakker, 1993a), the same holds true for partitions Π with only two essential events. Partitions containing at least two essential events are called *essential partitions*. By (A2), any refinement of an essential partition is again essential. For all essential Π, we set $V^{\Pi}(1_{\Omega}) = 1$ and, for all j, $V_j^{\Pi}(0) = 0$. Assume now that Π' is a refinement of an essential partition Π. Then the additive representation $V^{\Pi'}$ for Π', when restricted to $F_{\downarrow}^{\Pi}$, provides an additive representation that is alternative to V^{Π}. By cardinality, V^{Π} differs from the restriction of $V^{\Pi'}$ only by scale and location. By our choices of the 0 and 1 levels of the additive representations, V^{Π} must be identical to the restriction of $V^{\Pi'}$. In other words, we have:

For any essential partition Π, the additive representation V^{Π} coincides on common domain with the additive representation $V^{\Pi'}$ of any finer partition. (A4)

The major difficulty in the present proof is to show that the different additive representations for different partitions coincide on the "diagonal" of constant acts:

$$\sum_i V_i^{\Pi}(x) = \sum_j V_j^{\Pi'}(x) \text{ for all outcomes } x \text{ and essential partitions } \Pi \text{ and } \Pi'. \quad (A5)$$

We call two essential partitions Π and Π' *linked*, denoted by $\Pi \mathscr{L} \Pi'$, if they assign the same $V^{\Pi} = V^{\Pi'}$-value to constant acts. Of course, we must prove that all essential partitions are linked. For every essential partition there exists a coarsening that is essential and contains exactly two events. By (A4), the additive representations on the two partitions must be in agreement, so the two partitions are linked. From this it follows that it suffices to prove (A5) for essential partitions containing exactly two events.

Let us consider essential partitions (A,A^c) and (B,B^c), and prove $(A,A^c) \mathscr{L} (B,B^c)$. If $A \cap B$ is essential, then, since essentiality of A^c by (A1) implies essentiality of $(A \cap B)^c$, the partition $(A \cap B, (A \cap B)^c)$ is also essential; hence, by (A4), we can then link $(A,A^c) \mathscr{L} (A \cap B, A \backslash B, B \backslash A, A^c \cap B^c) \mathscr{L} (A \cap B, (A \cap B)^c) \mathscr{L} (B \cap A, B \backslash A, A \backslash B, B^c \cap A^c) \mathscr{L} (B,B^c)$. So, say $A \cap B$ is inessential. Then, by (A2), essentiality of A implies essentiality of $A \backslash B$, and essentiality of B implies essentiality of $B \backslash A$. So all partitions below are essential:

$(A,A^c)\mathcal{L}(A\backslash B,A\cap B,A^c\cap B^c,B\backslash A)\mathcal{L}(A\backslash B,(A\cap B)\cup(A^c\cap B^c),B\backslash A)\mathcal{L}$

$(A\backslash B,A^c\cap B^c,A\cap B,B\backslash A)\mathcal{L}((A\backslash B)\cup(A^c\cap B^c),(A\cap B)\cup(B\backslash A))=(B^c,B).$

It remains to be shown that $(B,B^c)\ \mathcal{L}\ (B^c,B)$, where both partitions are essential.

By (A3), there exists a refinement of (B,B^c) having at least three essential events. If B contains at least two of those essential events, then by (A1) the partition can be chosen of the form (B_1, B_2, B^c). In the other case, B^c must contain at least two of those essential events, and the partition can be chosen of the form (B,B_1, B_2). Say the former case holds. We have $(B,B^c)\ \mathcal{L}\ (B_1,B_2,B^c)\ \mathcal{L}\ (B_1,(B_2\cup B^c))\ \mathcal{L}\ (B_1,B^c,B_2)\ \mathcal{L}\ ((B_1\cup B^c), B_2)\ \mathcal{L}\ (B^c, B_1, B_2)\ \mathcal{L}\ (B^c,B)$. Formula (A5) has been established.

With (A5) available, it is easy to show that the additive representations on the sets $F_\downarrow^\Pi$ coincide on common domain, resulting in one overall function, and further to show that this one overall function is representing: Consider $f\in F_\downarrow^\Pi$, $g\in F_\downarrow^{\Pi'}$, Π and Π' essential. By Lemma A2, there exists a certainty equivalent x for g. We have $f\geqslant g\Leftrightarrow f\geqslant x\Leftrightarrow V^{\Pi}(f)\geqslant V^{\Pi}(x)$. Since by (A5) and the representation on $F_\downarrow^{\Pi'}$ we have $V^\Pi(x) = V^{\Pi'}(x) = V^{\Pi'}(g))$, we get

$$f\geqslant g \Leftrightarrow V^\Pi (f) \geq V^{\Pi'} (g). \tag{A6}$$

First, this shows, by setting $f = g$ and reflexivity, that different V^Π's coincide on common domain so can be considered the restriction of one function, denoted CU, on F^x. Note that the union of the F^Π's for essential Π is A^x; every simple act is contained, by (A3), in F^Π for an essential Π. Second, representation of $\geqslant$ by CU now is immediate from (A6).

Continuing the proof, we proceed to show that CU is of the form as described in table 1. For positive x, we define

$$x \geq 0 \text{ then } W(x,A) := CU(x1_A). \tag{A7}$$

For negative x, we define

$$x < 0 \text{ then } W(x,A) := CU(x1_\Omega) - CU(x1_{A^c}). \tag{A8}$$

Obviously, W is an outcome-dependent capacity. It follows that, for each partition $\Pi = (B_1, ..., B_n)$ and simple act $\sum_{j=1}^n x_j 1_{B_j}$ with $x_1 \geq \cdots \geq x_n$, we have

$$\sum_{j=1}^n V_j^\Pi(x_j) = \sum_{j=1}^n \{W(x_j, \cup_{i=1}^j B_i) - W(x_j, \cup_{i=1}^{j-1} B_i)\}. \tag{A9}$$

So W and CU are associated. Next we show joint monotonicity of W. Let $y > x$, $B \supset A$, $B\backslash A$ essential. Monotonicity of $\geqslant$ implies $y1_A + y1_{B\backslash A}+x1_{B^c} > y1_A + x1_{B\backslash A}+x1_{B^c}$. Sub-

stituting CU gives joint monotonicity of W. Continuity of $W(x, A)$ in x follows from continuity of CU on any F^{Π}, with Π an essential partition that refines (A,A^c).

UNIQUENESS: The only freedom in the construction of CU and W in the above proof was the choice $CU(1_\Omega) = (V^{\Lambda\Pi}(1_\Omega) =)1$. One easily verifies that this could have been any other positive real number, leading to an admissible change in scale. There is no further freedom. Given $W(0,A) = W(x,\Phi) = 0$, CU, as well as the additive representations $V^{\Lambda\Pi}$ within the sets $F^{\Pi}_\downarrow$ (i.e., the restrictions of CU to these sets) in the proof above, we must assign value 0 to the act constant 0. So indeed each of these is unique up to a positive scale. This scale might be thought to depend on $F^{\Pi}_\downarrow$, but since the functions must coincide on the constant acts, the scale is independent of $F^{\Pi}_\downarrow$. Indeed CU, and W as uniquely determined by CU, are ratio scales.

Remark A1. In Theorem 1, comonotonic independence can be replaced by ordinal independence (6).

Proof. In the proof of Theorem 1, comonotonic independence was used to obtain, through some other results in the literature (Corollary 3.6 in Wakker, 1993a, which in turn is based on Theorem 4a in Wakker, 1991) an additive representation V^{Π} on sets $F^{\Pi}_\downarrow$. This in turn was based on application of Debreu (1960, Theorem 3) on small Cartesian products (denoted $E^{\leq}$ in the mentioned references) on which comonotonic independence is identical to full-strength independence. Since **R** is an arc-connected separable topological space, this might as well have been obtained, without any further complication, by (5) and Gorman (1968) instead of by comonotonic independence and Debreu (1960).   □

Proof of Lemma 2.

Let $f \in F^b$. It is well-known that there are sequences of simple acts f^j, g^j, converging to f, monotonically from above and below respectively; so, $f^j(\omega) \geq f^{j+1}(\omega) \geq f(\omega) \geq g^{j+1}(\omega) \geq g^j(\omega)$ for all j, ω. By Lemma A2, which only needed the conditions of Lemma 2, these simple acts have certainty equivalents, abbreviated CE. By pointwise monotonicity, $CE(f^j) \geq CE(f^{j+1}) \geq CE(g^{j+1}) \geq CE(g^j)$ for all j. Let the limits be x_f and x_g, respectively; obviously $x_f \geq x_g$. $[\forall j : x_g \geq g^j]$ implies $x_g \geq f$. $[\forall j : f^j \geq x_f]$ implies $f \geq x_f$. So $x_f = x_g \sim f$.   □

Proof of Lemma 3.

If f pointwise dominates g, then every simple function pointwise dominated by g is pointwise dominated by f as well. Since $CU(f)$ and $CU(g)$ are supremum of CU-values of pointwise dominated simple acts, $CU(f) \geq CU(g)$ follows.

For the other condition, let f,g,f',g' be comonotonic, and satisfy (4). By Wakker (1989, Lemma VI.3.3.iii), there exists an ordering $\succcurlyeq'$ (transitive, complete, antisymmetric: $\omega \succcurlyeq'$ ω' and $\omega' \succcurlyeq' \omega \Rightarrow \omega = \omega'$) of Ω such that, for each act h in $\{f,f',g,g'\}$,

$$\omega \succ' \omega \Rightarrow h(\omega) \geq h(\omega'). \tag{A10}$$

Each other act h satisfying (A10) is comonotonic with f,f',g,g'. Let $A := \{\omega : f(\omega) = g(\omega), f'(\omega) = g'(\omega)\}$. Then for each $\omega \in A^c : f(\omega) = f'(\omega)$ and $g(\omega) = g'(\omega)$. Let f^j, f'^j, g^j, g'^j be four sequences of simple acts, converging to f, f', g, and g', monotonically from below; e.g., $f^j(\omega) \leq f^{j+1}(\omega) \leq f(\omega)$ for all j,ω. They can always be chosen such that they satisfy (A10) and (4). To wit, suppose that they do not satisfy these conditions. We transfer these properties from f, f', g, g' to the simple acts as follows. For (A10), we change each h^j from $\{f^j, f'^j, g^j, g'^j\}$ by assigning, to each ω, $\max\{h^j(\omega') : \omega' \preccurlyeq' \omega\}$. Henceforth, f^j, f'^j, g^j, g'^j will denote the changed acts. This change has not affected monotonic convergence, in particular, for all j, ω, and symbols $h = f, f', g, g'$, still $h^j(\omega) \leq h^{j+1}(\omega) \leq h(\omega)$. Observe that all simple acts in question satisfy (A10), and are comonotonic with f, f', g, g'. This change is called a *step towards* (A10). Next, for each $\omega \in A$, we replace both $f^j(\omega)$ and $g^j(\omega)$ by their maximum $\max\{f^j(\omega), g^j(\omega)\}$, and analogously replace $f'^j(\omega)$ and $g'^j(\omega)$ by their maximum. For each $\omega \in A^c$, we replace both $f^j(\omega)$ and $f'^j(\omega)$ by their maximum, and replace $g^j(\omega)$ and $g'^j(\omega)$ by their maximum. Again, the changed acts will be denoted as the original ones. This change is called a *step towards* (4), and, again, has not affected monotonic convergence from below. Now also $f^j = g^j$ and $f'^j = g'^j$ on A, and $f^j = f'^j$ and $g^j = g'^j$ on A^c. I.e.,(4) is satisfied. But (A10) may now be violated. So we make another step towards (A10), then one towards (4), etc., until these steps involve no more change. Each change gives rise to at least one more identity between the outcomes of the simple acts. Since the number of outcomes is finite, after a finite number of steps, the procedure stops, and (A10) and (4) are satisfied. Substituting the resulting simple acts in the CU form on F^s shows that $CU(f^j) - CU(g^j) = CU(f'^j) - CU(g'^j)$ for all j : by comonotonicity f^j, f'^j, g^j, g'^j belong to one same $F_\downarrow^\sqcap$. Π can be refined such that each element is either a subset of A, or of A^c. The terms $V(x^j, C, D)$ with $C\backslash D \subset A$ cancel both in $CU(f^j)-CU(g^j)$ and in $CU(f'^j)-CU(g'^j)$; the ones with $C\backslash D \subset A^c$ are identical in these two differences. The resulting equality $CU(f^j)-CU(g^j) = CU(f'^j) - CU(g'^j)$ is preserved after taking limits, i.e., $CU(f)-CU(g) = CU(f')-CU(g')$. □

Proof of Theorem 2.

IMPLICATION $(i) \Rightarrow (ii)$: Suppose (i) holds. Transitivity, completeness, and continuity, are direct. Joint monotonicity of $\succcurlyeq$ on F^s was already established in Theorem 1. Comonotonic independence was established Corollary 1.

IMPLICATION $(ii) \Rightarrow (i)$. Suppose (ii) holds. By Theorem 1, there exists an outcome-dependent capacity W, jointly monotonic and continuous in outcomes, such that on F^s the

associated functional represents $\geqslant$. By Lemma A1, CU is continuous on F^s. We define, for any $f \in F^b$, $CU(f) = CU(x)$, where x is the certainty equivalent outcome/constant act the existence of which is guaranteed by Lemma 2. Obviously, on F^s, this definition coincides with CU as already defined there, and obviously CU represents $\geqslant$. It remains to be shown that CU is continuous on F^b. Let the bounded f^j converge to the bounded f. By compactness, it suffices to show that, if $CU(f^j)$ converges, it converges to $CU(f)$. Suppose this were not the case, say the limit were larger. Then there is an outcome x strictly between that limit and $CU(f)$. For large j, then $f^j > x$. But $x > f$, contradicting continuity of $\geqslant$.

$\square$

Proof of Observation 1.

That CU and W are ratio scales and that W is continuous in outcomes follows from Theorem 1. Pointwise monotonicity follows from Lemma 3. That continuity uniquely determines the extension of CU from F^s to F^b follows from the denseness of the simple acts.

The only implication of comonotonic independence used in the proof of (ii) $\Rightarrow$ (i) was its restriction to F^s. For the extension of F^s to F^b, it was not used. Corollary 1 shows how it can be strengthened.

Only minor details in the proof need to be adapted when the set of acts is a set between F^s and F^b. Availability of simple acts was all that was needed in the proof.

$\square$

Proof of Observation 2.

Under CEU, thus also under SEU, we see that, by restriction to constant acts, continuity of CU implies continuity of the utility function. Conversely, continuity of the utility function implies continuity of the functional, mainly by uniform continuity of the utility function when restricted to an appropriate compact set.

$\square$

Appendix B: General outcome sets

For simplicity of presentation, we have formulated the results for the case where the outcome set is **R**. This appendix shows how to adapt the results to the case where the outcome set is any connected topological space. Let us first discuss a problem that may arise if there are maximal and/or minimal outcomes. Then it may happen that the representing functional CU is "driven to infinity," i.e., assigns value ∞ to maximal acts, and/or value $-\infty$ to minimal acts. Being a ratio scale, CU then necessarily is an extended representation. The following example illustrates:

Example B1. Let the set of outcomes be $[-1,1]$, $\Omega = \{\omega_1, \ldots, \omega_n\}$ $(n \leq 3)$, $\succeq$ is represented by $\sum_{i=1}^{n} x_i 1_{\{\omega_i\}} \mapsto (x_{\pi(1)}+1) \times \exp(x_{\pi(2)}) \times \cdots \times \exp(x_{\pi(n)})$, with $x_{\pi(1)} \geq \cdots \geq x_{\pi(n)}$ (see Example 3.8 in Wakker, 1993a). An extended CU-representation is obtained by taking logarithm: the act constant -1 gets assigned value $-\infty$; other acts get assigned value $\ln([x_{\pi(1)} + 1]) + x_{\pi(2)} + \cdots + x_{\pi(n)}$, for $x_{\pi(1)} \geq \cdots \geq x_{\pi(n)}$. Note that $W(-1,A) = -\infty$ for all essential, (i.e., nonempty) A. Since CU and W are ratio scales, there does not exist a real-valued CU-representation.

Let us now, similarly to appendix 2 in Wakker (1994), show how to adapt Theorem 2 to general outcome sets. We will continue to use the notation $\sum_{j=1}^{n} x_j 1_{A_j}$ for the simple act assigning outcome x_j to A_j for each j. In general, $\geq$ on outcomes is replaced by the preference relation $\succeq$ restricted to the constant acts. Equalities $x = y$ are sometimes replaced by equivalences $x \sim y$. We assume that replacement of outcomes by other equivalent outcomes does not affect preference. For an adaptation of Theorem 1 this would have to be made explicit; for the adaptation of Theorem 2 that we provide below, it will be implied by pointwise monotonicity, a condition that will be added. Thus, in the presence of joint monotonicity (now with $y > x$ instead of $y > x$), the ordering of outcomes is the same for all essential $B\backslash A$.

It was convenient (while not essential) for the derivations in this article, in particular for the uniqueness results, to let the representations have the value 0 at 0. Hence, for a general outcome set, we fix one outcome x^0, called *neutral*, at which the representation and the outcome-dependent capacity will be given the value 0. The neutral outcome should be neither maximal nor minimal to avoid the problems of being driven to infinity at this outcome; that would preclude normalization at 0.

An act f is *bounded* (in preference) if there exist outcomes υ,μ such that $[\forall \omega \in \Omega : \upsilon \preceq f(\omega) \preceq \mu]$. Certainty equivalents need no longer be uniquely determined. The choice between them is always immaterial. Acts are required to be *measurable:* the inverses of sets of outcomes of the form $\{x : x \succeq y\}$ and $\{x : x \preceq y\}$ should be events. In the presence of continuity of $\succeq$, measurability with respect to any algebra containing the topology on the outcome set will ensure this.

Supnorm-continuity cannot be defined in this general setting. One solution, which we inferred from Chateauneuf (1990), is to assume, somewhat less generally, that the outcome set is a separable connected metric space. Alternatively, one may make explicit the implications of (supnorm-)continuity needed in the proofs of Theorems 1 and 2. These implications are as follows. On F^s, one imposes *simple continuity*, which requires the restriction of $\succeq$ to any F^Π to be continuous with respect to the product topology there. This (and the possibility to replace outcomes by indifferent outcomes without affecting preference) is all that is needed to adapt Theorem 1. On F^s, the analogous simple continuity of CU is, by Lemma A1, equivalent to supnorm-continuity if the outcome set is $\mathbf{R}$. On F^b, more implications of supnorm-continuity are needed. First, we add pointwise monotonicity explicitly and assume the existence of a certainty equivalent for each $f \in A^b$. Then we impose a "simple-act denseness" condition to guarantee that the CU-value of an act is indeed the infimum of the CU-values of pointwise dominating simple acts, as well

as the supremum of pointwise dominated simple acts: We say *simple-act denseness* holds if, for all $f > h$, there exists a simple act g such that $f > g > h$, where g pointwise dominates h, and a simple act g' such that $f > g' > h$ where g' is pointwise dominated by f. An analogous condition is defined for CU. That the added restriction about pointwise dominance of g above is appropriate may appear from its necessity in the adaptation of Lemma 3, needed for the derivation of comonotonic independence. Next we define an Archimedean axiom that precludes the phenomenon of driven-to-infinity, illustrated in Example B1. Let $\Pi = \{A_1, \cdots, A_n\}$ be an ordered partition. A sequence of outcomes $\alpha^1, \alpha^2, \cdots$ is a *standard sequence* on event A_j in $F_{\downarrow}^{\Pi}$ if there exist $x, i \neq j$, and nonequivalent outcomes v_i, w_i such that (denoting only the outcomes for events A_i and A_j, assuming other outcomes identical to those of x) $(v_i, \alpha^k) \sim (w_i, \alpha^{k+1})$ for all $k = 1, 2, \cdots$, with furthermore all acts in question contained in $F_{\downarrow}^{\Pi}$. We call $\alpha^1, \alpha^2, \cdots$ a *second-order standard sequence* on A_j in $F_{\downarrow}^{\Pi}$ if for every k there exists a standard sequence $(\beta^1, \beta^2, \cdots)$ on A_j in $F_{\downarrow}^{\Pi}$, such that for some $m, l \in IN$, $\alpha^k \sim \beta^m$, $\alpha^{k+1} \sim \beta^{m+1}$, and $\alpha^{k+2} \sim \beta^{m+2l}$. The *second-order Archimedean axiom* requires that a second-order standard sequence be finite whenever it is *bounded* (i.e., there exist outcomes μ and v that are preferred and dispreferred, respectively, to each element of the sequence). Wakker (1993a) showed that the axiom is necessary and sufficient to preclude the driven-to-infinity phenomenon (and that a regular "first-order" version is not).

Theorem B1. Suppose the set of outcomes is a connected topological space with a fixed (neutral) outcome x^0 that is neither maximal nor minimal, and that there exist three disjoint essential events. The following two statements are equivalent for $\geqslant$ on F^b, the set of bounded measurable acts:

(i) $\geqslant$ is represented by a pointwise monotonic, simple-continuous, simple-act dense functional CU. On F^s, CU is of the form

$$\sum_{j-1}^{n} x_j 1_{A_j} \mapsto \sum_{j-1}^{n} \{W(x_j, \cup_{i=1}^{j} A_i) - W(x_j, \cup_{i=1}^{j-1} A_i)\}$$

for $x_1 \geqslant \cdots \geqslant x_n$. Here W is an outcome-dependent capacity ($W(x^0, A) = W(x, \varnothing) = 0$ for all A and x) that is jointly monotonic.

(ii) $\geqslant$ satisfies joint monotonicity on F^s, transitivity, completeness, pointwise monotonicity, simple-continuity, the existence-of-certainty-equivalent condition, simple-act denseness, the second-order Archimedean axiom, and comonotonic independence (comonotonic sure-thing principle).

Proof. For the implication (i) $\Rightarrow$ (ii), we make a few comments. For the certainty equivalent condition, note that by pointwise monotonicity, there exists a dominating and dominated constant act for every bounded act f. By continuity of CU on constant acts (as implied by simple continuity) and connectedness, f has a certainty equivalent. For the proof of comonotonic independence, we shortly comment on the proof of (the analogue

of) Lemma 3. For $h = f,f',g,g'$, one now takes a sequence of simple acts h^j such that h pointwise dominates h^{j+1} and h^{j+1} pointwise dominates h^j for all j, and $CU(h^j)$ converges to $CU(h)$. By simple-act denseness, such h^j exist. The other parts of the implication (i) $\Rightarrow$ (ii) are straightforward.

Next we assume that (ii) holds, and derive (i). First the CU-representation on F^s is derived. We adapt the proof of $(ii) \Rightarrow (i)$ of Theorem 1 given in appendix A. Lemma A2, ensuring the existence of certainty equivalents for simple acts, is now derived from Wakker (1989, Lemma VI.6.1) and connectedness. For every partition containing at least three disjoint essential events, now, by Theorem 3.3.C (instead of Corollary 3.6) of Wakker (1993a), a cardinal additive representation V^Π is obtained on $F^\Pi_\downarrow$. Again, such representations are subsequently obtained on all essential partitions. We take any non-maximal outcome (existing by continuity and connectedness) $x^1 > x^0$, and let all additive representations assign value 0 to the act constant x^0, and value 1 to the act constant x^1. With these modifications, the representation on F^s results. The extension to F^b requires no further comments.   □

The, somewhat complicated, second-order Archimedean axiom in statement (ii) above is only needed for maximal or minimal outcomes. If there exist maximal outcomes but no minimal outcomes and the second-order Archimedean axiom is dropped, then only a minor adjustment is required to permit an extended representation. It is then possible that for the maximal equivalence class of acts (containing the acts that assign maximal outcomes to each state of nature, and acts that do so except on an inessential event) the CU value must be ∞. For all nonmaximal acts the CU value then is finite. The representation in terms of the outcome-dependent capacity in Theorem B1.(i) can then be maintained if $W(x,A)$ is infinite whenever x is maximal and A^c is inessential, and if the convention is adopted that terms $\infty - \infty$ are zero (such terms occur only for the events $A_n, \cdots, A_j$ such that $A_n \cup \cdots \cup A_j$ is inessential).

If the second-order Archimedean axiom is dropped and there exist minimal outcomes, the case is more complicated. Then for the minimal equivalence class of acts (containing the acts that assign minimal outcomes to each state of nature, and acts that do so except on an inessential event) the CU value can be $-\infty$. The representation in terms of the outcome-dependent capacity in Theorem B1.(i) then cannot be maintained for acts for which the minimal outcome occurs (for the other acts it can be). The reason is that $W(x,A)$ in Formula (A8) must be $-\infty$ as soon as x is minimal and A is essential. Then in the summation in Theorem B1.(i) terms $\infty - \infty$ occur that cannot be defined as 0. The cumulative method of summation of the outcome-dependent capacity-differences turns out to be less convenient for minimal outcomes than for maximal outcomes. If one is willing to give up the convention that the representations are zero at the prespecified neutral outcome, then there is a way for maintaining the representation in terms of the outcome-dependent capacity in Theorem B1.(i); we inferred this method from Chateauneuf (1995, personal communication). If then a minimal outcome x exists, one sets $W(x,A)$ = 0 for all events A such that A^c is essential. Only if A^c is inessential, one sets $W(x^0, A)$ = 0 for a neutral outcome x^0 that is strictly perferred to the minimal outcome. Then the convention $\infty - \infty = 0$ always gives the right results in the summation.

Finally, let us note that Remark A1 can be extended to separable arc-connected outcome spaces by Gorman (1968), and to separable connected outcomes spaces by von Stengel's (1991) generalization.

Acknowledgement

We thank Jean-Yves Jaffray for helpful comments and we especially thank Alain Chateauneuf for pointing out an error in an earlier version.

Notes

1. The comonotonic sure-thing principle was tested empirically by Wakker, Erev, and Weber (1994) in the context of risk, and by Fennema and Wakker (1994) in the context of uncertainty.

References

Chateauneuf, A. (1990). "On the Use of Comonotonicity in the Axiomatization of EURDP Theory for Arbitrary Consequences," *CERMSEM*. University of Paris I. Extended abstract presented at Fifth International Conference on the Foundations and Applications of Utility, Risk and Decision Theory (FUR-90).

Chew, S.H. (1983). "A Generalization of the Quasilinear Mean with Applications to the Measurement of Income Inequality and Decision Theory Resolving the Allais Paradox," *Econometrica* 51, 1065–1092.

Chew, S.H. (1989). "Axiomatic Utility Theories with the Betweenness Property," *Annals of Operations Research* 19, 273–298.

Chew, S.H. (1990). "The Rank-Dependent Quasilinear Mean," Dept. of Economics, University of California, Irvine, CA.

Chew, S.H., and L.G. Epstein. (1989a). "A Unifying Approach to Axiomatic Non-Expected Utility Theories," *Journal of Economic Theory* 49, 207–240.

Chew, S.H., and L.G. Epstein. (1989b). "Axiomatic Rank-Dependent Means," *Annals of Operations Research* 19, 299–309.

Chew, S.H., L.G. Epstein, and U. Segal. (1991). "Mixture Symmetry and Quadratic Utility," *Econometrica* 59, 139–163.

Chew, S.H., L.G. Epstein, and P.P. Wakker. (1991). "A Unifying Approach to Axiomatic Non-Expected Utility Theories: Correction and Comment," *Journal of Economic Theory* 59, 183–188.

Chew, S.H. and E. Karni. (1994). "Choquet Expected Utility with a Finite State Space," *Journal of Economic Theory* 62, 469–479.

Choquet, G. (1954). "Theory of Capacities," *Annales de l'Institut Fourier* 5 (Grenoble), 131–295.

de Finetti, B. (1937). "La Prevision: ses lois logiques, ses sources subjectives," *Annales de l'Institut Henri Poincaré* 7, 1–68.

Debreu, G. (1960). "Topological Methods in Cardinal Utility Theory." In K.J. Arrow, S. Karlin, and P. Suppes (eds.), *Mathematical Methods in the Social Sciences*. Stanford, CA: Stanford University Press, pp. 16–26.

Dekel, E. (1986). "An Axiomatic Characterization of Preferences under Uncertainty: Weakening the Independence Axiom," *Journal of Economic Theory* 40, 304–318.

Ebert, U. (1988a). "Measurement of Inequality: An Attempt at Unification and Generalization," *Social Choice and Welfare* 5, 147–169.

Ebert, U. (1988b). "Rawls and Bentham Reconciled," *Theory and Decision* 24, 215–223.

Fennema, H. and P.P. Wakker. (1994). "A Test of Rank-Dependent Utility in the Context of Ambiguity," Journal of Risk and Uncertainty, forth-coming.

Fishburn, P.C. (1981). "Subjective Expected Utility: A Review of Normative Theories," Theory and Decision 13, 139–199.

Fishburn, P.C. (1983). "Transitive Measurable Utility," Journal of Economic Theory 31, 293–317.

Giboa, I. (1987). "Expected Utility with Purely Subjective Non-Additive Probabilities," Journal of Mathematical Economics 16, 65–88.

Gorman, W.M. (1968). "The Structure of Utility Functions," Review of Economic Studies 35, 367–390.

Green, J. and B. Jullien. (1988). "Ordinal Independence in Non-Linear Utility Theory," Journal of Risk and Uncertainty 1, 355–387. ("Erratum" (1989) 2, 119).

Gul, F. (1991). "A Theory of Disappointment Aversion," Econometrica 59, 667–686.

Hazen, G.B. (1987). "Subjectively Weighted Linear Utility," Theory and Decision 23, 261–282.

Huber, P.J. (1981). Robust Statistics. New York: Wiley.

Kahneman, D, and A. Tversky. (1979). "Prospect Theory: An Analysis of Decision under Risk," Econometrica 47, 263–291.

Luce, R.D. (1988). "Rank-Dependent, Subjective Expected-Utility Representations," Journal of Risk and Uncertainty 1, 305–332.

Luce, R.D. (1991). "Rank-and-Sign Dependent Linear Utility Models for Binary Gambles," Journal of Economic Theory 53, 75–100.

Luce, R.D., and P.C. Fishburn. (1991). "Rank-and-Sign Dependent Linear Utility Models for Finite First-Order Gambles," Journal of Risk and Uncertainty 4, 29–59.

Nakamura, Y. (1990). "Subjective Expected Utility with Non-Additive Probabilities on Finite State Space," Journal of Economic Theory 51, 346–366.

Nakamura, Y. (1992a). "Multi-Symmetric Structures and Non-Expected Utility," Journal of Mathematical Psychology 36, 375–395.

Nakamura, Y. (1992b). "A Generalization of Subjectrive Expected Utility without Transitivity and Additivity." paper presented at the Sixth FUR Conference, Cachan, France.

Quiggin, J. (1982). "A Theory of Anticipated Utility," Journal of Economic Behavior and Organization 3, 323–343.

Quiggin, J. (1989). "Sure Things—Dominance and Independence Rules for Choice under Uncertainty," Annals of Operations Research 19, 335–357.

Sarin, R. and P.P. Wakker. (1992). "Incorporating Attitudes towards Ambiguity in Savage's Set-up," Econometrica 60, 1255–1272.

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

Schmeidler, D. (1989). "Subjective Probability and Expected Utility without Additivity," Econometrica 57, 571–587.

Segal, U. (1989). "Anticipated Utility: A Measure Representation Approach," Annals of Operations Research 19, 359–373.

Segal, U. (1993). "The Measure Representation: A Correction," Journal of Risk and Uncertainty 6, 99–107.

Starmer, C., and R. Sugden. (1989). "Violations of the Independence Axiom in Common Ratio Problems: An Experimental Test of Some Competing Hypotheses," Annals of Operations Research 19, 79–101.

Tversky, A. and D. Kahneman. (1992). "Advances in Prospect Theory: Cumulative Representation of Uncertainty," Journal of Risk and Uncertainty 5, 297–323.

von Neumann, J., and O. Morgenstern. (1944, 1947, 1953). Theory of Games and Economic Behavior. Princeton, NJ: Princeton University Press.

von Stengel, B. (1993). "Closure Properties of Independence Concepts for Continuous Utilies," Mathematics of Operations Research 18, 346–389.

Wakker, P.P. (1989). Additive Representations of Preferences: A New Foundation of Decision Analysis. Dordrecht: Kluwer Academic Publishers.

Wakker, P.P. (1990a). "Characterizing Optimism and Pessimism Directly through Comonotonicity," Journal of Economic Theory 52, 453–463.

Wakker, P.P. (1990b). "Under Stochastic Dominance Choquet-Expected Utility and Anticipated Utility are Identical," *Theory and Decision* 29, 119–132.

Wakker, P.P. (1991). "Additive Representations on Rank-Ordered Sets. I. The Algebraic Approach," *Journal of Mathematical Psychology* 35, 501–531.

Wakker, P.P. (1993a). "Additive Representations on Rank-Ordered Sets II. The Topological Approach," *Journal of Mathematical Economics* 22, 1–26.

Wakker, P.P. (1993b). "Unbounded Utility for Savage's 'Foundations of Statistics', and other Models," *Mathematics of Operations Research* 18, 446–485.

Wakker, P.P. (1993c). "Counterexamples to Segal's Measure Representation Theorem," *Journal of Risk and Uncertainty* 6, 91–98.

Wakker, P.P. (1994). "Separating Marginal Utility and Probabilistic Risk Aversion," *Theory and Decision* 36, 1–44.

Wakker, P.P., and A. Tversky. (1993). "An Axiomatization of Cumulative Prospect Theory," *Journal of Risk and Uncertainty* 7, 147–176.

Wakker, P.P., I. Erev, and E. Weber. (1994). "Comonotonic Independence: The Critical Test between Classical and Rank-Dependent Utility Theories," *Journal of Risk and Uncertainty* 9, 195–230.

Yaari, M.E. (1987). "The Dual Theory of Choice under Risk," *Econometrica* 55, 95–115.