Institute for Operations Research and the Management Sciences

Unbounded Utility for Savage's "Foundations of Statistics," and Other Models
Author(s): Peter Wakker
Source: *Mathematics of Operations Research*, Vol. 18, No. 2 (May, 1993), pp. 446-485
Published by: INFORMS
Stable URL: http://www.jstor.org/stable/3690290
Accessed: 03/05/2011 09:25

UNBOUNDED UTILITY FOR SAVAGE'S "FOUNDATIONS OF STATISTICS," AND OTHER MODELS

PETER WAKKER

A general procedure for extending finite-dimensional "additive-like" representations for binary relations to infinite-dimensional "integral-like" representations is developed by means of a condition called truncation-continuity. The restriction of boundedness of utility, met throughout the literature, can now be dispensed with, and for instance normal distributions, or any other distribution with finite first moment, can be incorporated. Classical representation results of expected utility, such as Savage (1954), von Neumann and Morgenstern (1944), Anscombe and Aumann (1963), de Finetti (1937), and many others, can now be extended. The results are generalized to Schmeidler's (1989) approach with nonadditive measures and Choquet integrals, and Quiggin's (1982) rank-dependent utility. The different approaches have been brought together in this paper to bring to the fore the unity in the extension process.

Introduction. In decision science it is usually assumed that a decision maker, when choosing between alternatives, maximizes a quantitative functional, satisfying certain desirable properties. Often the functional is assumed to be an integral. In decision theory the integral of a utility function may be taken in an expected utility criterion, in statistics the integral of the negative of a loss function leads to the minimization of expected loss. In dynamic contexts the integral of income over the time axis is taken, with a density function representing discounting. The assumption of maximization of a quantitative functional is made operational by *behavioral foundations*, i.e., (*decision-theoretic*) *representation theorems*. These take preferences (binary choices) between alternatives as observable primitives, and give necessary and sufficient conditions ("characterizations") for the preferences to maximize a functional of the desired kind. Obviously, this requires a "translation" of the desired conditions of the functionals to be maximized into conditions for the preferences, and vice-versa. Hence conditions for integrals must be found that are suited for such a translation.

For the continuation of this introduction some terminology must be introduced, and the organization of the paper must be sketched; see Figure 1. Results are given for five approaches. In each of these the term alternative is used. The first approach, the "functional approach," appears in §1. Here functionals on a set of (alternatives =) functions to $\mathbb{R}$ are taken as primitives, and characterized as integrals. The obtained results will be used in the other approaches. In these other approaches binary preferences are taken as primitives, and those representable by integrals are characterized. Section 2 presents three approaches, each dealing with decision making
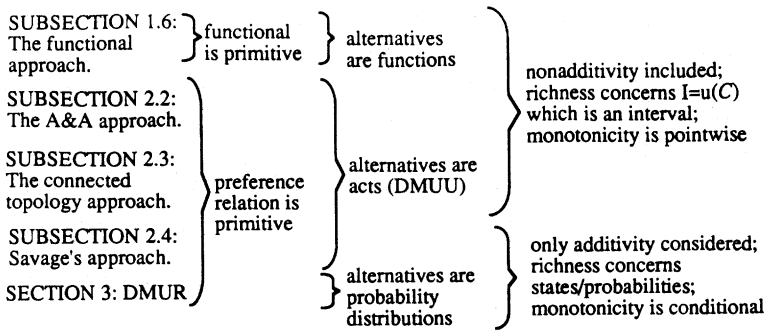
FIGURE 1. Organization of this paper. The (sub)sections of this paper are self-contained, and can be studied independently of other sections. Only Subsection 1.6 is needed for the proofs in Subsections 2.2 and 2.3.

under uncertainty (DMUU). Here (alternatives =) "acts" map a state space into a consequence space. The A&A approach of Subsection 2.2 assumes that the consequence space is a set of probability distributions; the connected topology approach of Subsection 2.3 assumes that the consequence space is a connected (separable) topological space; so both approaches impose richness restrictions upon the consequence space. While the three approaches mentioned so far include nonadditive integrals, the fourth and fifth approaches only deal with additive integrals. The fourth approach, i.e., the third and final approach to DMUU, is Savage's approach, considered in Subsection 2.4. It imposes richness restrictions upon the state space. Finally, §3 considers decision making under risk (DMUR), with (alternatives =) probability distributions. The richness restriction of this approach, requiring all step probability distributions to be available, is similar to the richness restriction of Savage's approach. To conclude the sketch of the organization of the paper, let us mention that in all approaches integrals designate expected utility in some sense.

The usual procedure in decision-theoretic integral representations reflects the definition of integrals. In a first stage the representation is obtained for "step" alternatives (having finite range). The integral representation for step alternatives is of an additive nature, and the equivalent conditions for preference relations through their finite (-dimensional) nature are proper for empirical testing. Next, in a second stage, some monotonicity (or continuity) condition is used to extend the integral representation from the step alternatives to bounded alternatives. The literature usually stops here. This paper adds a third stage. It also considers unbounded alternatives, as long as these have finite integrals. The desirability of such an extension was pointed out for instance in DeGroot (1970, end of §7.9), and for dynamic contexts in Lippman (1975). The main problem in the few approaches in the literature that address unbounded utility is that no proper condition was available to distinguish between alternatives with bounded integral and with unbounded, or undefined, integral. See for instance the beginning of §7.10 in DeGroot (1970).

"Truncation-continuity" (see Figure 3) turns out to be such a proper condition. By proper I mean that the condition can be translated into a condition for preferences that is observable. An example of a condition, proper in the above sense, is pointwise monotonicity. It requires that a function that pointwise dominates another should have at least as large an integral; see (1.5) below. In the usual setup of utility theory this translates into the condition for preferences that one alternative should be preferred to another if it dominates the other on all dimensions; see (2.1) below. Related conditions are used in §4.5 of Bhaskara Rao and Bhaskara Rao to define integrals. The conditions for defining integrals as used in §I.III.2 of Dunford and

Schwartz (1958), and in §4.4 in Bhaskara Rao and Bhaskara Rao (1983), can, to the best of my knowledge, not be translated directly into conditions for preferences. So these conditions are not proper in the above sense. The truncation-continuity condition for integrals is given in Lemma 1.8; formulas (2.2a, b) give the translation into a condition for preferences.

I agree (as in Wakker, 1989b, §III.1) with Aumann (1962, p. 446) and Schmeidler (1989, p. 576) that completeness of preferences (over "many" alternatives) is the most questionable condition in decision theory. Hence this paper strives for maximal flexibility concerning the alternatives required to be available. Apart from requiring availability of the "step" alternatives, and with every alternative some of its "truncations", the approach of this paper applies to *any* set of alternatives which have a finite integral; see Example 4.7. Step alternatives are very useful, first because they are empirically most meaningful, and second because they are theoretically useful in avoiding measurability problems and problems about infinite-dimensionality. Hence this paper generalizes results from the literature by restricting conditions as much as possible to step alternatives. As a price to pay, mainly, the formulation of "conditional monotonicity" for Savage's approach and DMUR is more complicated; it is however less restrictive than versions that have appeared in the literature.

The most important application of our results may be to the famous result in the first five chapters of Savage (1954). Savage was the first to give a sound behavioral foundation of subjective expected utility in which neither probabilities nor utilities are supposed to be given a priori. (The earlier Ramsey (1931) is not complete.) This changed the meaning of subjective probabilities and utilities from obscure-ad hoc into scientifically well founded. A foundation of Bayesian statistics, decision analysis, and the measuring of information had been laid. Still Savage's axioms have not been taken as point of departure in other analyses. It was discovered by Fishburn and Savage that Savage's axioms implied, in a way not foreseen by Savage, that the utility function had to be bounded. See Fishburn (1970, §14.1) and the second edition (1972, footnote on p. 80) of Savage (1954). Arrow (1971, pp. 64–65) found the same for a similar axiom system. Ever since, the extension of Savage's theorem to unbounded utility has been an open question, and with that the question "what is wrong with Savage's axioms?". This paper gives the extension of Savage's result to unbounded utility functions in Theorem 2.17. It incorporates (acts generating) the usual probability distributions over utilities, such as the normal distribution. While the implication of boundedness in Savage's theorem has been ascribed mainly to his postulates P6 and P7, I think that "what is wrong with Savage's axioms", is primarily his requirement of completeness of the preference relation on the set of all (alternatives =) acts, indeed, under preservation of the other conditions. This completeness is the only postulate of Savage that is relaxed in Theorem 2.17; his other postulates are all satisfied there. The underlying problem was already observed by Menger (1934). As soon as utility is unbounded, there exist acts with unbounded expected utility; to the best of my knowledge there is no "sensible" complete ordering of these. The well-known St. Petersburg Paradox builds on this; see Samuelson (1977) and Shapley (1977) and the references therein for many discussions. Also Toulet (1986) weakens completeness in Savage's axioms.

The results of this paper are useful for every field of science where integral representations are important, such as welfare theory (where alternatives are for instance allocations of income), dynamic decision making (where alternatives are for instance streams of income), and the theory of continuous systems (where the performance of a system may be an integral of the performances of the components). Wakker (1989b, §II.1) gives more examples. Given the largeness of the literature, I restrict attention to my own research specialization, DMUU/R, in this paper.

Results in the literature which can be extended by the techniques of this paper, but will not be discussed in the sequel, include de Finetti (1937, 1972, 1974), Koopmans (1972; our approach allows for discrete time as well as continuous time or mixed cases), Mak (1987, for infinitely many components), and Weibull (1985; truncation-continuity characterizes the set L^1 without assuming, as Weibull does, utility to be known).

Similarly to the extension of Savage's (1954) result, the results of Anscombe and Aumann (1963) and Wakker (1989b) are extended below. In the latter two no restriction is imposed on the state space. It may be finite, infinite, a continuum, and may or may not contain atoms. Also the result of von Neumann and Morgenstern (1944) for DMUR will be extended. The models of Savage and Anscombe and Aumann have not yet been extended to unbounded utility functions in the literature. For the result of von Neumann and Morgenstern several contributions have been made; nevertheless the generality provided by truncation-continuity in Theorem 3.6 has not been attained before.

Theorems 2.9 and 2.13, and Corollary 4.5 deal with generalizations of expected utility. While the axioms of Savage make subjective expected utility operational, they (primarily the sure-thing principle) are nowadays mostly taken to show that subjective expected utility has too many drawbacks. This was discussed in the early Allais (1953). The most influential result in the recent literature to deviate from expected utility has been Machina (1982). There expected utility was deprived of its most celebrated results, the results on risk aversion of Pratt and Arrow and the results on stochastic dominance. Machina showed that these results, extended in a natural way, hold in the general (differentiable) case, without needing the assumption of expected utility maximization. See also the survey in Machina (1987) and Fishburn (1988), for recent deviations from expected utility. Another impulse for the deviation from expected utility stems from the recent developments in artificial intelligence. The prevailing view is that subjective probabilities are too restricted and intractable to describe reasoning with uncertainty, or for the updating of knowledge. See for instance the discussion in *Statistical Science* (1987, 2, no. 1), or Wakker (1990c). Hence, besides expected utility, this paper considers the generalization of subjective expected utility called Choquet expected utility, as initiated by Schmeidler (1989; first version 1982). Here probabilities are allowed to be nonadditive, which may express vagueness about the values of the probabilities, and optimism/pessimism. Integration is done by means of the Choquet integral. As a corollary of these results, Subsection 4.2 extends Quiggin's (1982) rank-dependent utility to unbounded probability distributions.

In §4 countable additivity, uniqueness results, and necessity of conditions are discussed. The motivation to bring together representation results for several different approaches into one paper, rather than present them separately, is to bring to the fore their unity. This is sketched in Subsection 4.6. For that reason the similar intermediate steps will be made explicit in the analyses.

Subsections 2.2, 2.3, 2.4, §3, and Subsection 4.2, dealing with five different approaches to representations of preferences, can be read independently of each other. If a (sub)section uses definitions given before, mention will be made in that (sub)section. Also all main results are formulated so as to be immediately accessible without consultation of the text, other than for definitions. These main results are Theorems 1.12, 1.13, 1.15, 2.5, 2.9, 2.13, 2.17, 3.6, as well as Corollaries 2.10, 2.14, and 4.5.

1. Properties of integrals. This section deals with the functional approach, and studies properties of functionals which are (Choquet) integrals. Alternatives, the arguments of the functionals, are functions.

1.1. *Introduction*. The characterizations of integral representations in this paper reflect the definitions of integrals with respect to bounded measures. The start is with a measure, or, equivalently, with the integral of indicator functions, subsequently extended to all linear combinations of indicator functions (i.e., step functions). This paper concentrates on subsequent technical stages, where integral representations are extended to functions with infinitely many values. For such functions the integral can be defined as the limit of integrals of step functions that approximate the function in an appropriate manner.

This leads, in the second stage, to a well-defined real value whenever the function is bounded. Bounded functions can be enclosed from above and below by step functions, resulting in upper and lower sums which in the limit uniquely enclose the value of the integral. The exact way of constructing the enclosing step functions varies from context to context. In most contexts the range is "rich", i.e., an interval; then the enclosing step functions can be dominating/dominated "pointwise", i.e., for every argument. This is the usual procedure for defining integrals. In some contexts in this paper however, not the range, but the domain is "rich". Then the range of the measure is convex and the enclosing step functions are taken, somewhat more complicated, "conditionally dominating/dominated".

For unbounded functions complications occur. Integrals can be plus or minus infinity, or, which is the more difficult problem, they can be undefined if both tails are too dense. The third and final stage of the extension includes the functions that are unbounded but still have a well-defined and finite integral. The characteristic feature of such functions is that for each positive ϵ a "truncation from above" of the function can be made such that the integral above the truncation is smaller than ϵ (see Figure 3), with similar observations for truncation from below. This condition can be translated into a condition for preference relations, truncation-continuity, introduced in Wakker (1989b) for the context of subjective expected utility. The condition has not appeared elsewhere in the literature. Still I think it is the natural translation of the above characteristic feature. The aim of this paper is to show that this condition gives a simple and general tool to characterize integral representations for unbounded functions in many contexts.

The extension of preferences to functions with infinite and unbounded integrals is a topic for future research. In DMUU the only attempt along these lines known to us is Toulet (1986). Many open problems remain in her paper, however, such as how to define formally "may be compared" in Axiom 1(iv) while having correctness of Propositions 4 and 10, how to deal with the case where utility is both unbounded from above and below, and how to define preferences such as in Counterexample 5.1 while preserving conditions such as transitivity. The take-over criterion from economic growth theory, considering the integral of differences rather than the difference of integrals (see for instance von Weizsäcker 1965) seems better suited.

The results in this section are given for general nonadditive measures (capacities). A reader interested only in (integrals with respect to) additive measures may simply keep in mind throughout the sequel the special case that the capacities v are additive, i.e., are probability measures, and that the Choquet integrals are regular integrals. The reader need only observe then that (1.1) below is an alternative way to define additive integrals.

In the functional-analysis-literature it is usually assumed that the domain of a functional is a linear space of functions. In representations of preferences, however, the domain of the functional is usually not closed under addition and/or multiplication. Hence integrals must be characterized on more general domains. Given an algebra $\mathscr{A}$ on a set Ω, the set of *step functions*, denoted $\mathscr{F}^s$, is the set of functions from Ω to an interval I that are linear combinations of the indicator-functions of

elements of $\mathcal{A}$. The only restriction imposed on $\mathcal{F}$ below is that it contain all step functions. Step functions meet no measurability problems. They are measurable whatever algebra is imposed on the range, and so are their sums and mixtures. Hence, many conditions will be restricted to $\mathcal{F}^s$ below. Obviously, $\mathcal{F}^s$ is a subset of the set of measurable *simple functions* (i.e., functions with finite range) from Ω to I. In the functional approach of this section the algebra $\mathcal{D}$ on I is assumed to contain all one-point subsets; then $\mathcal{F}^s$ is identical to the set of measurable simple functions. So we assume throughout this section:

ASSUMPTION 1.1. (Functional Assumption). Ω is a nonempty set and $\mathcal{A}$ is an algebra of subsets of Ω. $I \subset \mathbb{R}$ is a nondegenerate interval and $\mathcal{D}$ is an algebra of subsets of I containing all intervals. $\mathcal{F}$ is a set of measurable functions from Ω to I, containing $\mathcal{F}^s$. V is a function from $\mathcal{F}$ to $\mathbb{R}$. □

The above assumption describes the *functional approach*. Note that the measurability assumptions are the usual ones if $\mathcal{D}$ is the Borel-σ-algebra and $\mathcal{A}$ is a σ-algebra. The more general approach of this paper, including (non-σ-) algebras, calls for some caution. For example, the sum of two measurable functions may be nonmeasurable. Because this has sometimes been overlooked in the literature, an example is given. Another example for a somewhat different context has been given in Fishburn (1982, §10.2). The sum/mixture of step functions is always measurable. This is one reason for restricting conditions to step functions.

EXAMPLE 1.2 (measurability problems for algebras). Let $\Omega = [0,1] \times [0,1]$, with $\mathcal{A}$ the set of all rectangles and their finite unions. $I = \mathbb{R}$ and $\mathcal{D}$ is the set of all intervals and their finite unions. Now f, the projection on the first coordinate, and g, the projection on the second coordinate, are measurable, but their sum is not; $(f + g)^{-1}([0,1])$ is not a finite union of rectangles. Similarly, mixtures (e.g., $\frac{1}{2}f + \frac{1}{2}g$) of f and g are not measurable. □

Real numbers are sometimes identified with constant functions. A function f *pointwise dominates* a function g, *notation* $f \geqslant_p g$, if $[\forall \omega \in \Omega: f(\omega) \geqslant g(\omega)]$. The notations $>_p$, $\leqslant_p$, $<_p$ are similar.

1.2. *Choquet integration.*

This subsection can be skipped by readers only interested in additive measures and integration. A function $v: \mathcal{A} \to [0,1]$ is a *capacity* if $v(\varnothing) = 0$, $v(\Omega) = 1$, and v is *monotonic with respect to set-inclusion*, i.e., $A \supset B \Rightarrow v(A) \geqslant v(B)$. In the literature the normalization $v(\Omega) = 1$ is sometimes omitted, and often the (set-) continuity condition of Subsection 4.1 is added; v is a (*finitely additive*) *probability measure* if furthermore v is *additive*, i.e., $v(A \cup B) = v(A) + v(B)$ for all disjoint A, B.

For a capacity v, and a measurable function $f: \Omega \to \mathbb{R}$, the *Choquet integral* of f (with respect to v), denoted $\int_\Omega f \, dv$, or $\int f \, dv$, or $\int f$, and introduced in Choquet (1953–1954), is

$$(1.1) \quad \int_{\mathbb{R}_+} v(\{s \in \Omega: f(s) \geqslant \tau\}) \, d\tau + \int_{\mathbb{R}_-} [v(\{s \in \Omega: f(s) \geqslant \tau\}) - 1] \, d\tau.$$

The integrals have been denoted as Lebesgue integrals; the integrands being nondecreasing, extended Riemann integrals could have been taken as well. As usual, the Choquet integral is undefined if in (1.1) $\infty - \infty$ would result. Obviously, a Choquet integral and the capacity in question are uniquely related through the equality $v(A) = \int 1_A \, dv$ (1_A denotes the *indicator function* of A). For countably additive (capacities $=$) probability measures, the Choquet integral is identical to the Lebesgue integral. For the more general finitely additive (capacities $=$) probability measures,

the Choquet integral coincides with the S-integrals in Bhaskara Rao and Bhaskara Rao (1983, §4.5), which uses upper and lower sums. An alternative definition of integration is given in Dunford and Schwartz (1958, Definition III.2.17), dealing with ranges more general than $\mathbb{R}$; see also §4.4 in Bhaskara Rao and Bhaskara Rao (1983). With $\mathbb{R}$ as range, the Dunford and Schwartz integral is defined, and then is identical to the Choquet integral, whenever the latter is finite. Further comments on integration with respect to finitely additive measures are given in de Finetti (1974, §3.12 and Chapter 6).

In Wakker (1989b, Chapter VI), elucidations are given for the Choquet integral. If a functional V is a Choquet integral, then the following straightforward results hold (with all V values assumed well defined):

$$(1.2) \qquad\qquad \forall \mu \in \mathbb{R}: V(\mu) = \mu,$$

$$(1.3) \qquad \forall \mu \in \mathbb{R}, \quad \forall f: V(f + \mu) = V(f) + \mu \quad (\textit{translation invariance}),$$

$$(1.4) \qquad \forall \lambda \in \mathbb{R}_+, \quad \forall f: V(\lambda f) = \lambda V(f) \quad (\textit{positive homogeneity}),$$

$$(1.5) \qquad \forall f, g: f \geqslant_p g \Rightarrow V(f) \geqslant V(g) \quad (\textit{pointwise monotonicity}).$$

Dellacherie (1970) was the first to observe the importance of the following condition for the Choquet integral. A set C of functions is *comonotonic* if

$$\nexists f, g \in C, \omega, \omega' \in \Omega \quad \text{such that} \quad f(\omega) > f(\omega'), g(\omega) < g(\omega').$$

The intuition for this condition can be inferred from Yaari (1987, p. 104), or Wakker (1990c, Example 1 up to Definition 4), or Wakker (1990b, §4). The following theorem has been proved in Dellacherie (1970) for positive functions under the assumption that v satisfies the first implication in Definition 4.1 below and that $\mathscr{A} = 2^\Omega$. Schmeidler (1986, Remark 4) was the first to prove the result in full generality.

THEOREM 1.3 (*comonotonic additivity for Choquet integrals*). *Suppose the Functional Assumption 1.1 holds. Let $\sigma, \tau \geqslant 0$, and suppose $f, g, \sigma f + \tau g \in \mathscr{F}$, and f, g are comonotonic. Then $\int (\sigma f + \tau g) \, dv = \sigma \int f \, dv + \tau \int g \, dv$, whenever all integrals in question are defined.* □

1.3. *Enclosing bounded functions by step functions when I is rich (an interval).* This subsection deals with the case that I is an interval. Then bounded functions $f \in \mathscr{F}$ (and their integrals) can be enclosed within step functions (and their integrals) mainly by pointwise monotonicity, as follows (see Figure 2(a)).

A sequence of functions $(f_j)_{j=1}^\infty$ *uniformly approximates* a function f *from below* if

$$(1.6) \qquad \forall \epsilon > 0 \; \exists j \in \mathbb{N} \; \forall i \geqslant j, \quad \forall \omega: f(\omega) - \epsilon \leqslant f_i(\omega) \leqslant f(\omega).$$

Similarly *uniform approximation from above* is defined. These approximations are of course special cases of supnorm-approximation. In agreement with customs in the field, which deviate somewhat from traditional terminology, we say that f is *bounded (from) above* if there exists an upper bound $\mu \in I$ for f; f is *bounded (from) below* if there exists a lower bound $\nu \in I$ for f, and f is *bounded* if it is both bounded from above and from below. $\mathscr{F}^b$ denotes the set of all bounded functions in $\mathscr{F}$. Thus a bounded function that is not contained in $\mathscr{F}$, is neither contained in $\mathscr{F}^b$. If I is bounded from some side but does not contain its infimum/supremum, then a function may not be bounded in the sense just introduced, but still be bounded in the
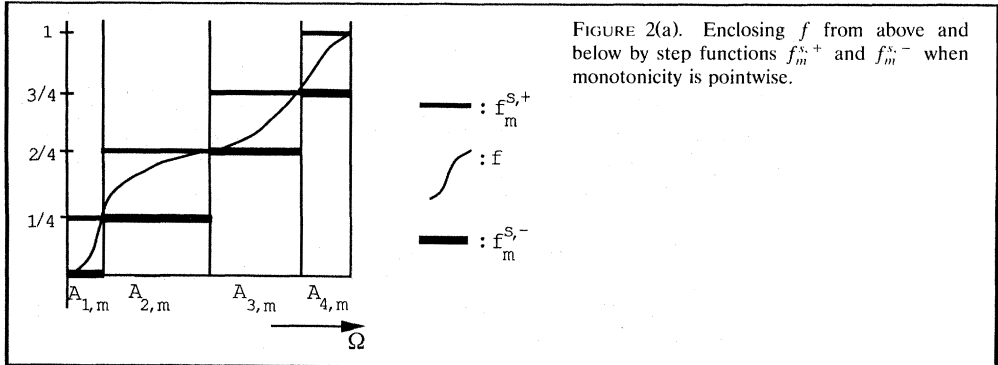
FIGURE 2(a). Enclosing f from above and below by step functions $f_m^{s,+}$ and $f_m^{s,-}$ when monotonicity is pointwise.

———— : $f_m^{s,+}$

$\int$: f

———— : $f_m^{s,-}$
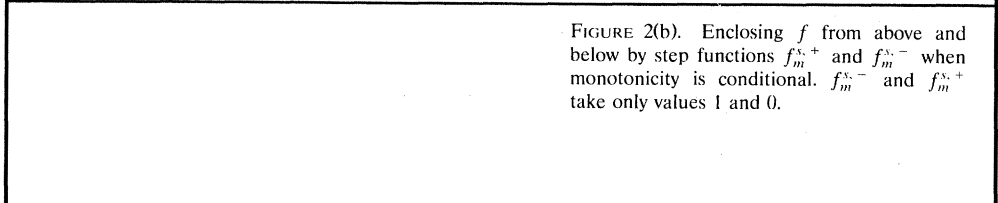
FIGURE 2(b). Enclosing f from above and below by step functions $f_m^{s,+}$ and $f_m^{s,-}$ when monotonicity is conditional. $f_m^{s,-}$ and $f_m^{s,+}$ take only values 1 and 0.
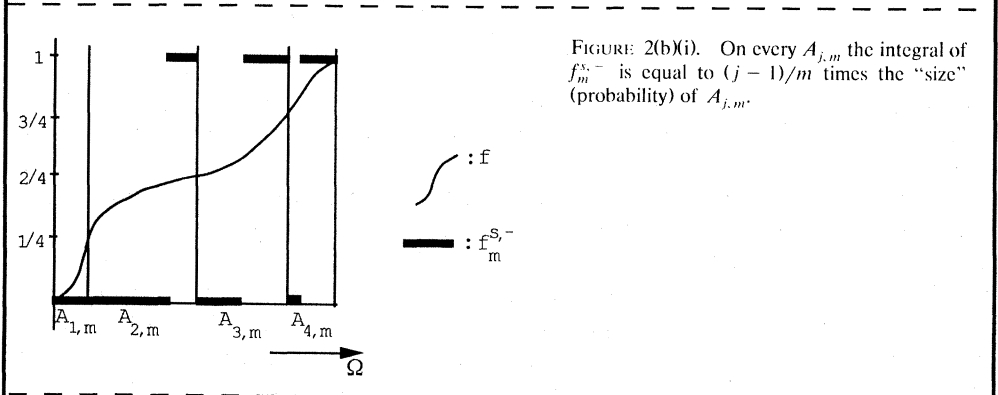
FIGURE 2(b)(i). On every $A_{j,m}$ the integral of $f_m^{s,-}$ is equal to $(j-1)/m$ times the "size" (probability) of $A_{j,m}$.

$\int$: f

———— : $f_m^{s,-}$
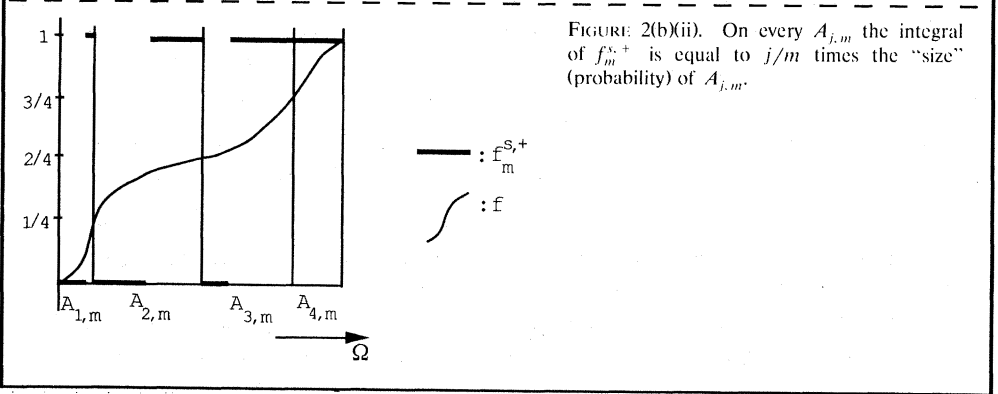
FIGURE 2(b)(ii). On every $A_{j,m}$ the integral of $f_m^{s,+}$ is equal to j/m times the "size" (probability) of $A_{j,m}$.

———— : $f_m^{s,+}$

$\int$: f

FIGURE 2. Enclosing a bounded function by step functions. A function f with bounds, say 0 and 1, is enclosed by step functions $f_m^{s,-}$ (from below) and $f_m^{s,+}$ (from above) for $m = 4$.

traditional sense (e.g., f' in Example 1.7). Hence Wakker (1989b) used the term strongly bounded for the above definition of bounded.

LEMMA 1.4. *Suppose the Functional Assumption 1.1 holds. If $f \in \mathcal{F}^b$, then a sequence of step functions approximating f uniformly from below, and a sequence of step functions approximating f uniformly from above, can be obtained. For all such sequences of step functions the Choquet integral tends to the Choquet integral of f.*

PROOF. The standard proof for additive measures can directly be applied to Choquet integrals. For the sake of reference, and comparison with reasonings below, details are given. Let $f \in \mathcal{F}^b$, with bounds say $\nu < \mu \in I$ (the case $\nu = \mu$ is trivial). Let $\nu = 0$, $\mu = 1$ for simplicity of notation. By translation invariance and positive homogeneity these choices are immaterial. Define for $m \in \mathbb{N}$

$$(1.7) \quad A_{1,m} := f^{-1}\left(\left[0, \frac{1}{m}\right]\right); \quad \forall 1 < j \leqslant m: A_{j,m} := f^{-1}\left(\left]\frac{j-1}{m}, \frac{j}{m}\right]\right), \quad \text{and}$$

$$f_m^{s,-} := \sum_{j=1}^{m} \frac{j-1}{m} 1_{A_{j,m}}, \qquad f_m^{s,+} := f_m^{s,-} + \frac{1}{m}.$$

Then $(f_m^{s,-})_{m=1}^{\infty}$ is a sequence of step functions, approximating f uniformly from below, and $(f_m^{s,+})_{m=1}^{\infty}$ is a sequence of step functions approximating f uniformly from above.

Next let $(f_j^{s,-})_{j=1}^{\infty}$ be an arbitrary sequence of step functions approximating f uniformly from below (from above is similar, and will not be elaborated). We may assume that, for all $j, [\forall \omega: f(\omega) - f_j^{s,-}(\omega) \leqslant 1/j]$. With $f_j^{s,+} := (f_j^{s,-} + 1/j)$, then $(f_j^{s,+})_{j=1}^{\infty}$ is a sequence of step functions approximating f uniformly from above. By pointwise monotonicity, the Choquet integrals of the $f_j^{s,-}$s are all smaller than or equal to the Choquet integral of f, and the Choquet integrals of the $f_j^{s,+}$s are all greater or equal. Because the Choquet integrals of the $f_j^{s,+}$s and $f_j^{s,-}$s differ by $1/j$ (as follows from translation invariance), the desired convergence follows. □

Note the interplay in the above proof of pointwise monotonicity and "below-and-above-enclosing-denseness" of step functions, which did not use any continuity.

LEMMA 1.5. *Suppose the Functional Assumption 1.1 holds. V is a Choquet integral on $\mathcal{F}^b$ whenever V is a Choquet integral on $\mathcal{F}^s$ and is pointwise monotonic.*

PROOF. Let $f \in \mathcal{F}^b$. By Lemma 1.4, f can be uniformly approximated from above by a first sequence of step functions, and from below by a second sequence of step functions. By pointwise monotonicity of $V, V(f)$ is the unique real number between the two sequences of Choquet integrals of the step functions. This, by pointwise monotonicity of the Choquet integral, is equal to the Choquet integral of f. □

1.4. *Enclosing bounded functions by step functions when Ω is rich.* This section only sketches the ideas leading to the results in Subsection 2.4 and §3. See Figure 2(b). Instead of the richness assumption concerning the range I as in the previous subsection, in the literature often a richness assumption is made concerning the domain Ω. It will entail the availability of all probabilities. The step functions obtained in Lemma 1.4 need not be available anymore and an alternative, somewhat more complicated way must be used to enclose bounded functions between step functions. For nonadditive contexts this would be very complicated, and will not be pursued here. *So we assume in this subsection that V is an integral on $\mathcal{F}^s$ with respect*

to an additive capacity v; v is also denoted as P. The richness assumption that we use is as follows:

$$\forall A \in \mathscr{A}, 0 < \sigma < 1 \; \exists B \in \mathscr{A} \; \text{ such that } \; B \subset A, \, P(B) = \sigma P(A).$$

It can be inferred from Example 4.10 that indeed the approach of the previous subsection will not work under this richness assumption. Suppose $f \in \mathscr{F}^b$ has bounds $\nu < \mu$. Suppose, for simplicity of notation, that $\nu = 0$, $\mu = 1$. Let, for any $m \in \mathbb{N}$ and $j = 1, \ldots, m$, $A_{j,m}$ be as in (1.7). As in the proof of Lemma 1.4, step functions $f_m^{s,-}$ and $f_m^{s,+}$ are defined for any m such that

$$(1.8) \qquad \int_\Omega f_m^{s,-} \, dP \leqslant \int_\Omega f \, dP \leqslant \int_\Omega f_m^{s,+} \, dP \quad \text{and}$$

$$(1.9) \qquad \int_\Omega f_m^{s,+} \, dP - \int_\Omega f_m^{s,-} \, dP = \frac{1}{m}.$$

The step functions will now only assign values 0 and 1. They are not dominating/dominated pointwise, but "conditionally", as follows. To define $f_m^{s,-}$, we split up any $A_{j,m}$ into two events $A_{j,m}^{1,-}$ and $A_{j,m}^{0,-}$. The superscript minus refers to the minus of $f_m^{s,-}$; versions with $+$ will appear below. $A_{j,m}^{1,-}$ has probability $((j-1)/m)P(A_{j,m})$; here $f_m^{s,-}$ is 1. $A_{j,m}^{0,-}$ has probability $((m-j+1)/m)P(A_{j,m})$; here $f_m^{s,-}$ is 0. To define $f_m^{s,+}$, we split up any $A_{j,m}$ into two events $A_{j,m}^{1,+}$ and $A_{j,m}^{0,+}$. $A_{j,m}^{1,+}$ has probability $(j/m)P(A_{j,m})$; here $f_m^{s,+}$ is 1. $A_{j,m}^{0,+}$ has probability $((m-j)/m)P(A_{j,m})$; here $f_m^{s,+}$ is 0. Note that

$$\int_{A_{j,m}} f_m^{s,-} \, dP = \frac{j-1}{m} P(A_{j,m}), \qquad \int_{A_{j,m}} f_m^{s,+} \, dP = \frac{j}{m} P(A_{j,m}) \quad \text{for all } j.$$

Hence (1.9) follows, and for any $A_{j,m}$

$$(1.10) \qquad \forall \omega \in A_{j,m} \colon \int_{A_{j,m}} f_m^{s,-} \, dP \leqslant f(\omega) P(A_{j,m}) \leqslant \int_{A_{j,m}} f_m^{s,+} \, dP.$$

Conditionally dominating/dominated refers to the appropriate inequality in (1.10). Because of this condition, for any $A_{j,m}$,

$$\int_{A_{j,m}} f_m^{s,-} \, dP \leqslant \int_{A_{j,m}} f \, dP \leqslant \int_{A_{j,m}} f_m^{s,+} \, dP,$$

and consequently (1.8) follows. Indeed, f has been enclosed between step functions "conditionally given each $A_{j,m}$".

At this stage an analogue of Lemma 1.5 could be formulated, with pointwise monotonicity replaced by a condition for functionals that might be called conditional monotonicity. I think however that for its own sake the condition is not very natural for functionals; hence I abstain from making it explicit. It will be implicit in Savage's approach to DMUU, and in the DMUR approach.

1.5. *Unbounded functions with finite (Choquet) integrals.* In this subsection, functionals are characterized as Choquet integrals on domains where the Choquet integral is well defined and finite. These domains may include unbounded functions. We assume that it has already been established that the functionals are Choquet

integrals on the set $\mathcal{F}^b$ of all bounded functions in $\mathcal{F}$, which is assumed to include the set $\mathcal{F}^s$ of all step functions. First some useful observations are given for the case where the set I is rich. Then, if a function f is unbounded from only one side, the set of step functions is sufficiently rich to determine $\int f$. The following claims are readily derived from the definition of the Choquet integral under the Functional Assumption 1.1

(1.11)   for all f bounded from below: $\int f = \sup\left\{ \int f^s : f \geqslant_p f^s \right\},$

(1.12)   for all f bounded from above: $\int f = \inf\left\{ \int f^s : f \leqslant_p f^s \right\}.$

LEMMA 1.6. *Suppose the Functional Assumption* 1.1 *holds*, V *is a Choquet integral on* $\mathcal{F}^b$, *and* V *is pointwise monotonic. Then* $V(f) \geqslant \int f$ *for every* $f \in \mathcal{F}$ *that is bounded from below, and* $V(f) \leqslant \int f$ *for every* $f \in \mathcal{F}$ *that is bounded from above.*

PROOF. We only consider the case where f is bounded from below. Then

$$V(f) \geqslant \sup\{V(f^s): f \geqslant_p f^s\} = \sup\left\{ \int f^s : f \geqslant_p f^s \right\} = \int f \, dv,$$

the inequality by pointwise monotonicity, the second equality by (1.11). □

The inequalities in the above lemma can be strict. The following example, a variation on (1) of §5.4 in Savage (1954), illustrates.

EXAMPLE 1.7 (truncation-continuity cannot be omitted). Let $\Omega = \mathbb{N}$, $\mathscr{A} = 2^\Omega$, $I = {]}0, 1]$. Let $\mathcal{U}$ be an ultrafilter (for a definition see Kelley (1955) or Wakker (1989b)) on $\mathbb{N} \setminus \{1\}$, containing no finite set. Let P be a probability measure assigning probability $1/2$ to $\{1\}$ and probability $1/2$ to every subset of $\mathbb{N}$ in $\mathcal{U}$. Obviously, P is not countably additive. One may think of P as assigning probability $1/2$ to two atoms, $\{1\}$, and an "invisible atom at infinity" described by $\mathcal{U}$. Let

$$V: f \mapsto \int f \, dP - \lim_{\epsilon \downarrow 0} P\big(f^{-1}({]}0, \epsilon])\big).$$

For every function bounded in the sense defined above, $V(f)$ is the Choquet integral of f, which in this case is a regular additive integral. However, for f': $\omega \mapsto 1/\omega$, $\lim_{\epsilon \downarrow 0} P(f'^{-1}({]}0, \epsilon]))$ is $1/2$, so that $V(f')$ is the Choquet integral of f' minus $1/2$. V is not identical to the Choquet integral. □

For functions that are bounded from one side, a continuity condition could be imposed in the spirit of truncation-continuity as defined below, but with truncations replaced by pointwise dominating, or dominated, step functions. Thus Choquet integrals could be characterized. However, if a function f is unbounded from both sides, then no step function will dominate it pointwise, or be dominated pointwise, and approximations as in (1.11) and (1.12) cannot be invoked to identify the Choquet integral. For these functions an alternative method must be used. For the additive Lebesgue integral, where the integrals of functions bounded from one side are defined through (1.11) and (1.12), the integrals of functions f that are unbounded from both sides are traditionally defined as the sum of the integrals of $f^+ := f_{\vee 0}$ and $f^- := f^{\wedge 0}$, whenever this sum is defined. Such an approach could also be used for the definition of the Choquet integral. It is however, to the best of my knowledge, not suited for the purpose of this paper, where properties of integrals are to be "translated" into conditions in terms of the preferences represented by the integrals.
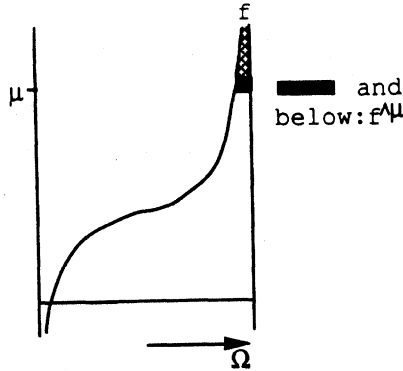
FIGURE 3. Above truncation. The above truncation $f^{\wedge\mu}$ of f at level μ. If the integral of f is finite, then the area dashed xxx tends to 0 if μ tends to infinity. If the integral of f is plus infinite or undefined, then the area dashed xxx is infinite no matter how large μ is.

The reason is that I do not know of a way to translate the addition-operation into conditions for preferences. Hence I use an alternative approach, similar to the definition of improper Riemann integrals, where integrals are obtained as a limit of integrals of truncated functions. The method will also be useful if not I, but Ω, is rich.

For $f \in \mathscr{F}$ and $\mu \in I$, $f^{\wedge\mu}$, the *above truncation* of f at μ, assigns $\min\{f(\omega), \mu\}$ to every $\omega \in \Omega$; $f_{\vee\nu}$, the *below truncation* of f at ν, assigns $\max\{f(\omega), \nu\}$ to every $\omega \in \Omega$. As a preparation for truncation-continuity, note that a finite Choquet integral can be approximated by integrals of truncations, also if the integrand is unbounded. The restriction to $\mu \in f(\Omega)$ below makes the result also suited for approaches where the range of alternatives does not satisfy richness assumptions.

LEMMA 1.8 (*Truncation-continuity of the Choquet integral*). *Let $\int$ denote a Choquet integral, and let $f \in \mathscr{F}$. Suppose $\int f$ is finite. Then*

(1.13a) $$\forall \epsilon > 0 \; \exists \mu \in f(\Omega) : 0 \leqslant \int f - \int f^{\wedge\mu} \leqslant \epsilon \quad and$$

(1.13b) $$\forall \epsilon > 0 \; \exists \nu \in f(\Omega) : 0 \leqslant \int f_{\vee\nu} - \int f \leqslant \epsilon.$$

PROOF. Note that the lemma did not require the truncations to be contained in $\mathscr{F}$. We only prove (1.13a). *First suppose $f(\Omega)$ is unbounded above.* If the integral $\int_{\mathbb{R}_+} v(\{s \in \Omega : f(s) \geqslant \tau\}) \, d\tau$ is finite, then for every positive ϵ there must exist a positive $\mu \in \mathbb{R}$ such that the integral differs by no more than ϵ from $\int_{[0,\mu]} v(\{s \in \Omega : f(s) \geqslant \tau\}) \, d\tau$. The Choquet integral $\int f^{\wedge\mu}$ is

$$\int_{[0,\mu]} v(\{s \in \Omega : f(s) \geqslant \tau\}) \, d\tau + \int_{\mathbb{R}_-} \left[v(\{s \in \Omega : f(s) \geqslant \tau\}) - 1 \right] d\tau,$$

and differs by less than ϵ from $\int f$.

Next suppose $f(\Omega)$ is not unbounded above, so that it has a finite supremum. Then we can take $f^{\wedge\mu}$ with μ in $f(\Omega)$ differing by less than ϵ from the supremum of $f(\Omega)$. Hence $(f^{\wedge\mu}) + \epsilon$ has Choquet integral larger than or equal to that of f. By translation invariance, the Choquet integral of $f^{\wedge\mu}$ differs by no more than ϵ from the Choquet integral of f. □

The above result is characteristic for finite Choquet integrals. As soon as a Choquet integral of f is infinite, the difference between this and the Choquet integral of any above truncation of f is always infinite. To characterize functions, possibly unbounded, that have a finite V value identical to the Choquet integral, the above property of functions f with finite Choquet integral values is mimicked.

DEFINITION 1.9. We call $f \in \mathcal{F}$ *truncation-robust* (*with respect to V*) if for all $g^s \in \mathcal{F}^s$

$$(1.14a) \quad V(f) > V(g^s) \Rightarrow \exists \mu \in I: f^{\wedge \mu} \in \mathcal{F}, \quad V(f^{\wedge \mu}) > V(g^s) \quad \text{and}$$

$$(1.14b) \quad V(f) < V(g^s) \Rightarrow \exists \nu \in I: f_{\vee \nu} \in \mathcal{F}, \quad V(f_{\vee \nu}) < V(g^s).$$

V is *truncation-continuous* if the above conditions hold for all $f \in \mathcal{F}$. □

Truncation-continuity imposes some richness condition on $\mathcal{F}$; loosely speaking, sufficiently many truncations must be available for each $f \in \mathcal{F}$. In all main results below, truncation-continuity implies that $\mathcal{F}$ is *weakly truncation-closed*: for every $\mu \in I$ there must exist $\mu' \geqslant \mu$ such that $f^{\wedge \mu'} \in \mathcal{F}$, and for every $\nu \in I$ there must exist $\nu' \leqslant \nu$ such that $f_{\vee \nu'} \in \mathcal{F}$. Given that, for a Choquet integral functional, the function $f \in \mathcal{F}$ has a finite Choquet integral value if and only if f is truncation-robust with respect to that Choquet integral functional. Most results in the literature require, more restrictively, that $\mathcal{F}$ be *truncation-closed*, i.e., that all truncations are contained in $\mathcal{F}$.

To establish integral representations for unbounded utility functions, Wakker (1989b, Chapter V) assumed that for every $f \in \mathcal{F}$ there exists a "certainty equivalent", or "fair price", in the terminology of de Finetti. A *certainty equivalent* is an element $\alpha \in I$ such that $V(\alpha) = V(f)$. To achieve maximal unity between the several approaches of this paper, including those where richness is imposed on Ω, the assumption is formulated in a slightly more general way.

DEFINITION 1.10. V satisfies the *step equivalent assumption* if

$$\forall f \in \mathcal{F} \, \exists f^s \in \mathcal{F}^s: V(f) = V(f^s). \quad □$$

As constant functions are special cases of step functions, the above condition is indeed implied by, so less restrictive than, the assumption that certainty equivalents exist. In the presence of convexity of I, and continuity of V on $\mathcal{F}^s$ as implied by the other conditions in all main results below, the two conditions are actually equivalent (the certainty equivalent is between the maximum and minimum of the step equivalent). Step equivalence serves to exclude infinitesimal differences between equivalence classes. As truncation-continuity, it can as well serve to exclude, in the presence of the other conditions, functions with integral value plus or minus infinity. Actually, such functions do not pose serious problems for representation theorems. The functions that are really problematic are those that are unbounded from both sides, and have an undefined integral; i.e., their positive part gives integral plus infinity, their negative part minus infinity. These functions can satisfy the step equivalent assumption, as well as the other usual decision-theoretic assumptions. Only to discard these, the full strength of truncation-continuity is needed.

LEMMA 1.11. *Under the Functional Assumption 1.1, V is a Choquet integral whenever V is a Choquet integral on $\mathcal{F}^b$, V is pointwise monotonic, V satisfies the step equivalent assumption, and V is truncation-continuous.*

PROOF. V satisfies *above-truncation-continuity* if (1.14a) holds for all f. We first establish (without using truncation-continuity in full force):

(1.15)   if V satisfies above-truncation-continuity, then for all $f \in \mathscr{F}$:

$$\int f \, dv \text{ is well defined and } V(f) \leqslant \int f \, dv.$$

As a preparation we show that $\int f$ is well defined. Suppose, for contradiction, it were not. Let $V(f) = V(f^s)$ for $f^s \in \mathscr{F}^s$. Because $\int f$ is undefined, I is unbounded from below (in fact from both sides) so $\nu \in I$ can be found with $\nu = V(\nu) < V(f^s)$, i.e., $V(\nu) < V(f)$. By (1.14a), $V(\nu) < V(f^{\wedge \mu})$ for some μ. Contradiction results from $-\infty < V(\nu) < V(f^{\wedge \mu}) \leqslant \int f^{\wedge \mu} = -\infty$, where the weak inequality follows from Lemma 1.6, the equality holds because $\int f$ is undefined.

So $\int f$ is well defined. Suppose now, for contradiction, that $V(f) > \int f$. Let $V(g^s) = V(f)$ for a $g^s \in \mathscr{F}^s$. The just-assumed inequality implies the existence of a $\xi \in I$ such that $\xi < V(f)$. By this, and convexity of I, g^s can be modified slightly into f^s, such that $V(f) > V(f^s) > \int f$. By above-truncation-continuity, there is an above truncation $f^{\wedge \mu}$ such that still

(1.16) $$V(f^{\wedge \mu}) > V(f^s) > \int f \, dv.$$

However, by Lemma 1.6 and pointwise monotonicity of the Choquet integral,

$$V(f^{\wedge \mu}) \leqslant \int f^{\wedge \mu} \, dv \leqslant \int f \, dv,$$

contradicting (1.16). So (1.15) has been proved.

V satisfies *below-truncation-continuity* if (1.14b) holds. Similarly to (1.15) we get

(1.17)   if V satisfies below-truncation-continuity, then for all $f \in \mathscr{F}$:

$$\int f \, dv \text{ is well defined and } V(f) \geqslant \int f \, dv.$$

The theorem follows from (1.15) and (1.17).   □

1.6. *Results for the functional approach, with functions as alternatives.* In this subsection the results of the previous subsections are combined into a method for extending characterizations of Choquet integrals from step functions to more general functions, as long as the latter have a finite Choquet integral.

THEOREM 1.12. *Under the Functional Assumption 1.1, V is a Choquet integral whenever the following four conditions are satisfied:*
  (1) *V is a Choquet integral on $\mathscr{F}^s$,*
  (2) *V is pointwise monotonic,*
  (3) *V satisfies the step equivalent assumption,*
  (4) *V is truncation-continuous.*

PROOF.   This follows from Lemmas 1.5 and 1.11.   □

The following theorem extends the main result in Schmeidler (1986) to unbounded functions with finite Choquet integral.

THEOREM 1.13. *Suppose the Functional Assumption 1.1 holds with $I = \mathbb{R}$.*
$V: \mathscr{F} \to \mathbb{R}$ is a Choquet integral whenever the following five conditions are satisfied:
  (1) $V(1_\Omega) = 1$,
  (2) *comonotonic additivity: if f, g are comonotonic and $f, g, f + g \in \mathscr{F}$, then*
$V(f + g) = V(f) + V(g)$,
  (3) *V is pointwise monotonic,*
  (4) *V satisfies the step equivalent assumption,*
  (5) *V is truncation-continuous.*

PROOF.   By Schmeidler (1986, second half of p. 257), Conditions (1), (2) and (3)
imply that V is a Choquet integral on $\mathscr{F}^s$. The result now follows from Theorem
1.12.   □
  Condition (1) is necessary by Formula (1.2), Condition (2) is necessary by Theorem
1.3, Condition (3) is necessary by Formula (1.5), and necessity of Condition (4) follows
mainly from convexity of I. If $\mathscr{F}$ is weakly truncation-closed, then necessity of
Condition (5) follows from Lemma 1.8. So, if $\mathscr{F}$ is weakly truncation-closed, then all
conditions are necessary. Comonotonic additivity was only used for step functions in
the proof.
  Note that the above theorem applies to any set $\mathscr{F}$ that contains all step functions
together with any arbitrary set of unbounded functions with finite Choquet integral,
as long as weak truncation-closedness is satisfied. The corollary below is the result as
given in Schmeidler (1986). It makes use of the fact that in the above theorem step
equivalence and truncation-continuity have not been used on $\mathscr{F}^b$.

COROLLARY 1.14.   *Suppose the Functional Assumption 1.1 holds with $I = \mathbb{R}$. Let $\mathscr{F}$*
be the set of all bounded measurable functions from Ω to $\mathbb{R}$. Then V is a Choquet
integral if and only if $V(1_\Omega) = 1$ and V satisfies comonotonic additivity and pointwise
monotonicity.

PROOF.   Sufficiency of the conditions follows from Theorem 1.13 because trunca-
tion-continuity is trivially satisfied, and step equivalence is only needed for un-
bounded functions. Necessity was demonstrated above.   □
  The following generalization of Theorem 1.13 is more in accordance with the setup
of the following sections. It allows I to be any nondegenerate interval.

THEOREM 1.15 (*Choquet integral for unbounded functions*).   *Under the Functional*
Assumption 1.1, $V: \mathscr{F} \to \mathbb{R}$ is a Choquet integral whenever the following five conditions
are satisfied:
  (1) $\exists \nu < \mu \in I$ *such that $V(\nu \times 1_\Omega) = \nu$, $V(\mu \times 1_\Omega) = \mu$,*
  (2) *comonotonic mixability: if $f, g \in \mathscr{F}^s$ are comonotonic then*

$$V(\tfrac{1}{2}f + \tfrac{1}{2}g) = \tfrac{1}{2}V(f) + \tfrac{1}{2}V(g),$$

  (3) *V is pointwise monotonic,*
  (4) *V satisfies the step equivalent assumption,*
  (5) *V is truncation-continuous.*

PROOF.   By Theorem 1.12 it suffices to show that V is a Choquet integral on $\mathscr{F}^s$.
By appropriate translation and change of scale of both V and I, we may assume that
$\nu = 0$ and $\mu = 1$. Then for all $f \in \mathscr{F}^s$ also $\tfrac{1}{2}f \in \mathscr{F}^s$, and, with $g \equiv 0$ in Condition
(2), $V(\tfrac{1}{2}f) = \tfrac{1}{2}V(f)$. For all $f, g, f + g \in \mathscr{F}^s$, by Condition (2), it follows that
$V(\tfrac{1}{2}f + \tfrac{1}{2}g) = \tfrac{1}{2}V(f) + \tfrac{1}{2}V(g)$. These two observations imply comonotonic additivity
for all step functions $f, g, f + g \in \mathscr{F}^s$. Now the reasoning of Schmeidler (1986,
second half of p. 257) can be repeated, showing that V is indeed a Choquet integral
on $\mathscr{F}^s$.   □

2. Representations of preferences in DMUU, with acts as alternatives. This section studies *decision making under uncertainty* (DMUU). Here an alternative is an *act*, i.e., a function from a nonempty set Ω, the *state space*, to a nonempty set $\mathscr{C}$, the *consequence space*. Exactly one of the states is true, the others are untrue, and a decision maker is uncertain which state is true. Hence he is uncertain which consequence will result from a chosen act. The *preference relation* $\succcurlyeq$ of the decision maker is a binary relation on $\mathscr{F}$, the set of acts under consideration. We write $f \succ g$ if $f \succcurlyeq g$ and not $g \succcurlyeq f$, $f \preccurlyeq g$ if $g \succcurlyeq f$, $f \prec g$ if $g \succ f$, and $f \sim g$ if $f \succcurlyeq g$ and $g \succcurlyeq f$. Further, $\succcurlyeq$ is a *weak order* if it is complete ($f \succcurlyeq g$ or $g \succcurlyeq f$ for all f, g) and transitive, and $\succcurlyeq$ is *trivial* if $f \succcurlyeq g$ for all f, g. Elements of $\mathscr{C}$ are identified with the associated constant acts. Thus, for $\alpha, \beta \in \mathscr{C}$, we write $\alpha \succcurlyeq \beta$ if the associated constant acts are contained in $\mathscr{F}$ and the same relation applies to these. Following Fishburn (1982, p. 24), $A \subset \mathscr{C}$ is a *preference interval* if

$$\forall \alpha, \beta, \gamma \in \mathscr{C} : [\alpha, \gamma \in A, \alpha \succcurlyeq \beta \succcurlyeq \gamma] \Rightarrow \beta \in A.$$

For example, if $\mathscr{C}$ is the set of rational numbers and $\succcurlyeq = \geqslant$ on $\mathscr{C}$, then $[0, \sqrt{2}] \cap \mathscr{C}$ is a preference interval. Given an algebra $\mathscr{A}$ of *events* on Ω, a *step act* is an act that is constant on every element of a finite partition of Ω consisting of elements from $\mathscr{A}$. $\mathscr{F}^s$ denotes the set of step acts. Again, step acts are measurable whatever the algebra on $\mathscr{C}$ is, and $\mathscr{F}^s$ is a subset of the set of measurable simple acts. As illustrated in Fishburn (1982, p. 122), it may be an essential subset. In Wakker (1989b, Chapter V) the term step should have been used instead of simple.

Under the following assumption it can be seen that for every measurable simple act f^s there is a step act g^s such that $\forall \omega : f^s(\omega) \sim g^s(\omega)$. *We assume throughout this section*:

ASSUMPTION 2.1 (DMUU Assumption). Ω is a nonempty set endowed with an algebra $\mathscr{A}$. $\mathscr{C}$ is a nonempty set, endowed with an algebra $\mathscr{D}$. $\mathscr{F}$ is the set of acts, i.e., $\mathscr{F} \subset \mathscr{C}^{\Omega}$; $\succcurlyeq$ is a nontrivial binary relation on $\mathscr{F}$. $\mathscr{D}$ contains all preference intervals. All acts are assumed measurable. $\mathscr{F}$ contains $\mathscr{F}^s$. □

2.1. *The Central Representation Theorem.* This subsection presents a theorem by means of which many representation theorems from the literature can be extended, i.e., those where the range for (utility values of) alternatives is an interval (i.e., $u(\mathscr{C})$ below). The extensions are then a matter of routine. As an illustration, two important extensions are made explicit in the following two subsections.

V represents $\succcurlyeq$, or $\succcurlyeq$ *maximizes* V, if V is a functional from $\mathscr{F}$ to $\mathbb{R}$ such that $[\forall f, g \in \mathscr{F} : f \succcurlyeq g \Leftrightarrow V(f) \geqslant V(g)]$. We say that $\succcurlyeq$ maximizes *Choquet expected utility* (CEU) if there exists a capacity v on $\mathscr{C}$, and a function $u : \mathscr{C} \to \mathbb{R}$ such that $f \mapsto \int (u \circ f) \, dv$ represents $\succcurlyeq$; $\int (u \circ f) \, dv$ is the *Choquet expected utility* of f (CEU(f)). We use the term *subjective expected utility* (SEU) instead of Choquet expected utility if the capacity is additive. Note that measurability of u is ensured if $\mathbb{R}$ is endowed with the usual algebra generated by intervals, because then all sets $u^{-1}(]-\infty, \mu])$, $u^{-1}([\nu, \infty[)$ are preference intervals. In proofs below, measurability of sets of the form $u^{-1}(](j - 1)/m, j/m])$ is often used without explicit mention.

Next we adapt some definitions from §1, mainly by replacing $\geqslant$ on $\mathbb{R}$ by $\succcurlyeq$ on $\mathscr{C}$. Act f *pointwise dominates* act g, *notation* $f \succcurlyeq_p g$, if $[\forall \omega \in \Omega : f(\omega) \succcurlyeq g(\omega)]$. We say $\succcurlyeq$ is *pointwise monotonic* if

(2.1) $\forall f, g \in \mathscr{F} : f \succcurlyeq_p g \Rightarrow f \succcurlyeq g.$

A set C of acts is *comonotonic* if

$$\nexists f, g \in C, \omega, \omega' \in \Omega: f(\omega) \succ f(\omega'), g(\omega) \prec g(\omega').$$

An act f is *bounded* (*from*) *above* if there exists a $\mu \in \mathscr{C}$ such that $\forall \omega \in \Omega$: $\mu \succcurlyeq f(\omega)$. Then μ is an *upper bound* for f. *Boundedness* (*from*) *below*, and a *lower bound*, are defined similarly; f is *bounded* if it is both bounded from above and from below. $\mathscr{F}^b$ *denotes* the set of all acts in $\mathscr{F}$ that are bounded.

DEFINITION 2.2. We say that $\succcurlyeq$ satisfies the *step equivalent assumption* if

$$\forall f \in \mathscr{F} \ \exists f^s \in \mathscr{F}^s: f \sim f^s. \quad \square$$

LEMMA 2.3. *Suppose the DMUU Assumption 2.1 holds, and $\succcurlyeq$ maximizes CEU on $\mathscr{F}^s$ with respect to a capacity v and a function u. Suppose $u(\mathscr{C})$ is a nondegenerate interval. Then $\succcurlyeq$ maximizes CEU on $\mathscr{F}^b$ with respect to the same v and u whenever $\succcurlyeq$ is a weak order on $\mathscr{F}^b$, $\succcurlyeq$ is pointwise monotonic, and $\succcurlyeq$ satisfies the step equivalent assumption.*

PROOF. Because CEU represents $\succcurlyeq$ on $\mathscr{F}^s$, all step acts equivalent to f have the same CEU value. Hence we can define $V: \mathscr{F} \to \mathbb{R}$ by $V: f \mapsto \mathrm{CEU}(f^s)$ where $f^s \in \mathscr{F}^s$, $f^s \sim f$; V represents $\succcurlyeq$. It remains to prove that V is a Choquet integral on $\mathscr{F}^b$. Consider the set $\mathscr{F}'$ of functions from Ω to $u(\mathscr{C})$ of the form $u \circ f, f \in \mathscr{F}$. By two-fold pointwise monotonicity, any two acts f, g such that $[u \circ f = u \circ g]$, are $\sim$ equivalent, so have the same V value; hence we can define $V': \mathscr{F}' \to \mathbb{R}$ according to $V'(u \circ f) = V(f)$ for all f. It suffices to prove that V' is a Choquet integral. This follows from Lemma 1.5, if we verify all conditions of Lemma 1.5, thus of Assumption 1.1, for V'. This is done in the remainder of the proof.

Note that $\mathscr{F}'$ contains the entire $\mathscr{F}'^s$. V' is the Choquet integral on $\mathscr{F}'^s$, and inherits pointwise monotonicity from V. The algebra $\mathscr{D}'$ on I should be chosen such that all functions $u \circ f \in \mathscr{F}'$ are $\mathscr{D}' - \mathscr{A}$ measurable. This is ensured by taking for $\mathscr{D}'$ the minimal algebra that is possible, i.e., the algebra consisting of all finite unions of intervals. Then for each act f and interval A, $(u \circ f)^{-1}(A)$ is the inverse under f of a preference interval, so is measurable. $\square$

For $f: \Omega \to \mathscr{C}$ and $\alpha \in \mathscr{C}, f^{\wedge \alpha}$, the *above truncation* of f at α, assigns α to every $\omega \in \Omega$ s.t. $f(\omega) \succ \alpha$, and $f(\omega)$ to all remaining ω. For $f: \Omega \to \mathscr{C}$ and $\beta \in \mathscr{C}, f_{\vee \beta}$, the *below truncation* of f at β, assigns β to every $\omega \in \Omega$ s.t. $f(\omega) \prec \beta$, and $f(\omega)$ to all remaining ω. Central is the following property: $f \in \mathscr{F}$ is *truncation-robust* (*with respect to* $\succcurlyeq$) if for all $g^s \in \mathscr{F}^s$

(2.2a) $\qquad f \succ g^s \Rightarrow \exists \mu \in \mathscr{C}: f^{\wedge \mu} \in \mathscr{F}, f^{\wedge \mu} \succ g^s, \quad$ and

(2.2b) $\qquad f \prec g^s \Rightarrow \exists \nu \in \mathscr{C}: f_{\vee \nu} \in \mathscr{F}, f_{\vee \nu} \prec g^s.$

We call $\succcurlyeq$ *truncation-continuous* if every $f \in \mathscr{F}$ is truncation-robust. The stronger version where g^s is replaced by a general $g \in \mathscr{F}$, turns out to be equivalent in the presence of the other conditions in the main results. This stronger version was announced in Wakker (1985, p. 123), and actually introduced in Wakker (1989b, Definition V.4.5), for the context of additive representations; see Corollary 2.14. The condition in the literature most similar to truncation-continuity has been formulated for DMUR and will in that context be compared to the condition of this paper; see Example 3.5, and the text above that example. Truncation-continuity implies *weak truncation-closedness* in all main results: for every $\mu \in \mathscr{C}$ there must exist $\mu' \succcurlyeq \mu$

such that $f^{\wedge\mu'} \in \mathscr{F}$, and for every $\nu \in \mathscr{C}$ there must exist $\nu' \preccurlyeq \nu$ such that $f_{\vee\nu'} \in \mathscr{F}$.

LEMMA 2.4.   *Suppose the DMUU Assumption 2.1 holds, and $\succcurlyeq$ maximizes CEU on $\mathscr{F}^b$ with respect to a capacity v and a function u. Suppose $u(\mathscr{C})$ is a nondegenerate interval. Then $\succcurlyeq$ maximizes CEU with respect to the same v and u on the entire $\mathscr{F}$ (with CEU(f) real-valued for all $f \in \mathscr{F}$) whenever $\succcurlyeq$ is a weak order on $\mathscr{F}$, $\succcurlyeq$ is pointwise monotonic, $\succcurlyeq$ satisfies the step equivalent assumption, and $\succcurlyeq$ is truncation-continuous.*

PROOF.   Define $V, V', \mathscr{F}', \mathscr{D}', I$ as in the proof of Lemma 2.3. V' inherits step equivalence and truncation-continuity from $\succcurlyeq$ (for truncation-continuity, note that u represents $\succcurlyeq$ on $\mathscr{C}$). By Lemma 1.11, V' is a Choquet integral. The same follows for V.   □

THEOREM 2.5 (*Central Representation Theorem*).   *Suppose the DMUU Assumption 2.1 holds, and $\succcurlyeq$ maximizes CEU on $\mathscr{F}^s$ with respect to a capacity v and a function u. Suppose $u(\mathscr{C})$ is a nondegenerate interval. Then $\succcurlyeq$ maximizes CEU with respect to the same v and u on the entire $\mathscr{F}$ (with CEU(f) real-valued for all $f \in \mathscr{F}$) whenever the following four conditions hold:*
(1)   *$\succcurlyeq$ is a weak order on $\mathscr{F}$,*
(2)   *$\succcurlyeq$ is pointwise monotonic,*
(3)   *$\succcurlyeq$ satisfies the step equivalent assumption,*
(4)   *$\succcurlyeq$ is truncation-continuous.*

PROOF.   This follows from Lemmas 2.3 and 2.4.   □

2.2.   *The A&A approach to DMUU, with probability distributions as consequences.*
In this subsection results of Schmeidler (1989) and Anscombe and Aumann (1963) are entended. *We assume throughout this subsection, in addition to the DMUU Assumption 2.1:*

ASSUMPTION 2.6 (A&A Assumption).   *$\mathscr{C}$ is the set of all probability distributions with finite support over a nonempty set Γ. For any function $U: \Gamma \to \mathbb{R}$, EU: $\mathscr{C} \to \mathbb{R}$ denotes the expectation of U.*   □

The approach of this subsection is called the *A&A approach*. Elements of Γ are called *prizes*. Acts f, g can be mixed in a pointwise manner, as

$$\sigma f + (1 - \sigma)g : \omega \mapsto \sigma f(\omega) + (1 - \sigma)g(\omega).$$

Results are obtained for general sets of measurable acts $\mathscr{F}$, not necessarily closed under mixture operations. Hence many of the conditions, imposed on all acts in the literature, are weakened here by restricting them to $\mathscr{F}^s$, the set of step acts. This also avoids the measurability problems in the mixing of acts as in Schmeidler (1989) (compare Example 1.2).

The stochastic mechanism determining the prize that will result from a consequence is an auxiliary structure to facilitate the analysis. It makes possible the application of techniques for mixture spaces as started in von Neumann and Morgenstern (Appendix in 1947, 1953), and extensively used in Fishburn (1982), and many other recent works in decision theory. The first appearance of the following condition (without a restriction to step acts) was in von Neumann and Morgenstern (1944, Axioms 3:B:c and 3:B:d).

DEFINITION 2.7.   We say that $\succcurlyeq$ is *s(tep)-vNM-continuous* if $\forall f, g, h \in \mathscr{F}^s$

$$f \succ g \succ h \Rightarrow \exists \sigma, \tau \in \,]0, 1[\quad \text{such that} \quad (1 - \sigma)f + \sigma h \succ g \succ \tau f + (1 - \tau)h.   \square$$

The well-known independence condition, and its restriction to comonotonic acts as introduced in Schmeidler (1989), are restricted to step acts below.

DEFINITION 2.8. We say $\succcurlyeq$ satisfies *s(tep)-comonotonic independence* if, for all comonotonic step acts $\{f, g, h\}$, and $0 < \sigma < 1$,

$$f \succ g \Rightarrow \sigma f + (1 - \sigma)h \succ \sigma g + (1 - \sigma)h.$$

The definition of *s(tep)-independence* is obtained by requiring the above implication also for all noncomonotonic step acts $\{f, g, h\}$. □

The following theorem extends the Theorem of Schmeidler (1989) to acts with unbounded utility. The step equivalent condition is equivalent to the existence of a certainty equivalent, i.e., an equivalent constant act.

THEOREM 2.9 (*Extension of Schmeidler's (1989) CEU-representation to unbounded acts*).   *Suppose the DMUU Assumption 2.1 and the A&A Assumption 2.6 hold. Then* $\succcurlyeq$ *maximizes CEU with respect to a capacity v and a function* EU: $\mathscr{C} \to \mathbb{R}$ *derived from a function* U: $\Gamma \to \mathbb{R}$ *(with all Choquet integrals finite), whenever the following six conditions are satisfied:*

(1) $\succcurlyeq$ *is a weak order on* $\mathscr{F}$,
(2) $\succcurlyeq$ *is s-vNM-continuous,*
(3) $\succcurlyeq$ *is s-comonotonic independent,*
(4) $\succcurlyeq$ *is pointwise monotonic,*
(5) $\succcurlyeq$ *satisfies the step equivalent assumption,*
(6) $\succcurlyeq$ *is truncation-continuous.*

PROOF.   As shown in Schmeidler (1989), or Wakker (1990c, Theorem 6), Conditions (1), (2), (3) and (4) imply the representation on $\mathscr{F}^s$. EU($\mathscr{C}$) being a nondegenerate interval, Theorem 2.5 now gives the desired result.   □

Schmeidler (1989) considered the special case of the above theorem where $\mathscr{F}$ is the set of all bounded acts, and Conditions (2) and (3) are strengthened to hold for all these acts. On $\mathscr{F}^b$, Condition (6) is trivially satisfied, and Condition (5) is straightforwardly implied by the strengthening of s-continuity to all acts (leaving aside some measurability complications).

COROLLARY 2.10 (*Extension of Anscombe and Aumann's (1963) SEU representation to unbounded acts*).   *In the above theorem maximization of SEU holds if Condition (3) is strengthened to s-independence.*

PROOF.   Additivity of the capacity follows straightforwardly, e.g., by applying the theorem of Anscombe and Aumann (1963) to any fixed finite partition of Ω and the set of step acts measurable with respect to that partition.   □

The above corollary has extended the classical result of Anscombe and Aumann to infinite state spaces, and also to unbounded acts; that latter extension has not yet been provided in the literature. Fishburn (1982, Chapter 10) extends the result of Anscombe and Aumann to infinite state spaces, but assumes that $\mathscr{F}$ contains all measurable acts (and even, to obtain a mixture space structure, mixtures of these that can be nonmeasurable). This implies boundedness (with probability one) of all acts in $\mathscr{F}$, and boundedness of utility as soon as there is a denumerable partition of Ω with all elements having positive probability.

2.3.   *The connected topology approach to DMUU, with a connected topological space as consequence space.*   This subsection gives representation theorems where $u(\mathscr{C})$ is

an interval because u is continuous with respect to a connected topology on $\mathscr{C}$. Such an assumption is usually satisfied in the literature, e.g., if $\mathscr{C} = \mathbb{R}$, designating money, or $\mathscr{C} = \mathbb{R}_+^m$, designating commodity bundles. Also, the setup of the previous subsection can be considered a special case of this setup, $\mathscr{C}$ there being a convex subset of a linear space and $(EU =)u$ being linear. See the works of de Finetti or Yaari (1987). The most general approach along this line is Wakker (1989b), where the following assumption is made:

ASSUMPTION 2.11 (Topological Assumption). $\mathscr{C}$ is a connected separable topological space. □

The approach of this section is called the *connected topology approach*. The assumption of connectedness is crucial; the assumption of separability has been made for simplicity of presentation, and can nearly always be omitted. See Remark 4.8. The motivation for the approach of this subsection is as follows. As compared to Savage's approach, the continuity assumption (Condition (6) in Theorem 2.17) requiring fineness of the state space is replaced by continuity requiring fineness of the consequence space. The latter assumption is satisfied anyhow in many analyses. As compared to the A & A approach, we do not require all simple objective probability distributions over prizes to be available as consequences. These objective probability distributions over prizes are to be dealt with in a linear way, i.e., as in EU (see the A & A Assumption 2.6), in the A & A approach. Relaxing this requirement is particularly desirable in nonadditive contexts. This subsection generalizes the approaches of de Finetti to continuous instead of linear utility, and of course to more general consequence spaces than $\mathbb{R}$. Linearity of utility is generally considered to be too restrictive. The price to pay for the great generality is that the characterizing conditions are more complicated than those in the previous subsection. The presentation below will be concise.

For $A \in \mathscr{A}$, $\alpha \in \mathscr{C}$, and $f\colon \Omega \to \mathscr{C}$, $f_{-A}\alpha$ assigns $f(\omega)$ to every $\omega \in A^c$, and α to every $\omega \in A$. Let $\pi = (A_1, \ldots, A_m)$ be an ordered partition, or partition for short, of Ω. *In this paper a partition is always, without explicit mention, assumed to consist of events.* Write $\sum_{j=1}^m \alpha_j 1_{A_j}$ for the act with value α_j on A_j, $j = 1, \ldots, m$. $\mathscr{F}^\pi$ is the set of step acts of the form $\sum_{j=1}^m \alpha_j 1_{A_j}$, and $\mathscr{F}_c^\pi$ is the (comonotonic!) set of step acts of the form $\sum_{j=1}^m \alpha_j 1_{A_j}$ with, furthermore, $\alpha_1 \succcurlyeq \cdots \succcurlyeq \alpha_m$. Note that in the latter definition the ordering of the events in π is essential. The preference relation $\succcurlyeq$ is *s(tep)-continuous* if for any partition $(A_1, \ldots, A_m)$ and any act $f = \sum_{j=1}^m \beta_j 1_{A_j}$, the sets

$$\left\{ (\alpha_1, \ldots, \alpha_m) \in \mathscr{C}^m \colon \sum_{j=1}^m \alpha_j 1_{A_j} \succcurlyeq f \right\} \quad \text{and} \quad \left\{ (\alpha_1, \ldots, \alpha_m) \in \mathscr{C}^m \colon \sum_{j=1}^m \alpha_j 1_{A_j} \preccurlyeq f \right\}$$

are closed with respect to the product topology on $\mathscr{C}^m$. This *s*-continuity is of a finite-dimensional character, thus is weaker than most of the other continuity conditions used in the literature. It is weaker than continuity of $\succcurlyeq$ with respect to the (restriction of the) product topology, and also weaker than the sup-metric continuity as used in Koopmans (1972). Also the step equivalent assumption can be seen to be implied by product-topology-continuity, and by Koopmans' sup-metric-continuity.

From a set $C \subset \mathscr{F}$ it can be concluded that an event A is not impossible if there are acts $f, g \in C$ that coincide outside of A, but are not equivalent. So we call event A *essential* (*with respect to* $\succcurlyeq$) *on* C if there exist $f \succ g \in C$ where f and g coincide outside of A. The opposite of essential is *inessential*. (*In*)*essential* abbreviates (in)essential on $\mathscr{F}$, π(*in*)*essential* abbreviates (in)essential on $\mathscr{F}_c^\pi$. Write, for

$\alpha, \beta, \gamma, \delta \in \mathcal{C}$,

$$\alpha\beta \succ^* \gamma\delta \quad \text{if } \exists A \in \mathcal{A}, f \in \mathcal{F}^s, g \in \mathcal{F}^s \quad \text{such that}$$

(2.3) $f_{-A}\alpha \succcurlyeq g_{-A}\beta \quad \text{and} \quad f_{-A}\gamma \prec g_{-A}\delta.$

The idea is that the above preferences show that the "tradeoff" of getting α instead of β is better than the tradeoff of getting γ instead of δ; (2.5) will further illustrate. More comments on this definition are given in Wakker (1989b, §§II.4 and IV.2) and Wakker and Tversky (1991). We write $\succcurlyeq^*$ instead of $\succ^*$ if in (2.3) we have $\preccurlyeq$ instead of $\prec$ and if furthermore A is essential on $\mathcal{F}^s$.

The following variations were introduced in Wakker (1989b, §VI.4). We write $\succ_c^*$ instead of $\succ^*$ if furthermore the four acts in (2.3) are comonotonic. We write $\succcurlyeq_c^*$ instead of $\succcurlyeq^*$ if the four acts are comonotonic, and furthermore A is π-essential for a π such that the four acts in question are contained in $\mathcal{F}_c^\pi$. (The four acts are comonotonic if and only if they are contained in a set $\mathcal{F}_c^\pi$; see Wakker (1989b), Lemma VI.3.3.) Similarly to Wakker (1989b, Lemma IV.2.6), one shows that if $\succcurlyeq$ maximizes SEU, then

(2.4) $\alpha\beta \succcurlyeq^* \gamma\delta \Rightarrow u(\alpha) - u(\beta) \geqslant u(\gamma) - u(\delta) \quad \text{and}$

(2.5) $\alpha\beta \succ^* \gamma\delta \Rightarrow u(\alpha) - u(\beta) > u(\gamma) - u(\delta).$

Similarly to Wakker (1989b, Formulas VI.5.1, VI.5.5, VI.5.6) it is shown that if $\succcurlyeq$ maximizes CEU, then

(2.6) $\alpha\beta \succcurlyeq_c^* \gamma\delta \Rightarrow u(\alpha) - u(\beta) \geqslant u(\gamma) - u(\delta) \quad \text{and}$

(2.7) $\alpha\beta \succ_c^* \gamma\delta \Rightarrow u(\alpha) - u(\beta) > u(\gamma) - u(\delta).$

DEFINITION 2.12. The preference relation $\succcurlyeq$ satisfies *tradeoff consistency* if there do not exist consequences $\alpha, \beta, \gamma, \delta$ such that [both $\alpha\beta \succcurlyeq^* \gamma\delta$ and $\gamma\delta \succ^* \alpha\beta$]. The preference relation $\succcurlyeq$ satisfies *comonotonic tradeoff consistency* if there do not exist consequences $\alpha, \beta, \gamma, \delta$ such that [both $\alpha\beta \succcurlyeq_c^* \gamma\delta$ and $\gamma\delta \succ_c^* \alpha\beta$]. □

By (2.4) and (2.5), tradeoff consistency is a necessary condition for SEU maximization, and by (2.6) and (2.7) comonotonic tradeoff consistency is necessary for CEU maximization. The following theorem and corollary show that the tradeoff consistency conditions, in the presence of the other usual axioms, are also sufficient.

THEOREM 2.13 (*Extension of Wakker's* (1989b) *CEU representation to unbounded acts*). *Suppose the DMUU Assumption 2.1 and the Topological Assumption 2.11 hold. Then $\succcurlyeq$ maximizes CEU with respect to a capacity v and a continuous function u: $\mathcal{C} \rightarrow \mathbb{R}$ (with all Choquet integrals finite) whenever the following six conditions are satisfied*:
  (1) $\succcurlyeq$ *is a weak order on $\mathcal{F}$,*
  (2) $\succcurlyeq$ *is s-continuous,*
  (3) $\succcurlyeq$ *satisfies comonotonic tradeoff consistency,*
  (4) $\succcurlyeq$ *is pointwise monotonic,*
  (5) $\succcurlyeq$ *satisfies the step equivalent assumption,*
  (6) $\succcurlyeq$ *is truncation-continuous.*

PROOF. By Theorem VI.5.1 in Wakker (1989b), the above theorem holds if Ω is finite. Hence it also holds if a fixed finite partition of Ω is taken, and $\succcurlyeq$ is restricted to the acts measurable with respect to that partition. The resulting capacity v and

utility function u can be taken independent of the partition in question: fit v and u together for any two partitions by taking a common refinement of these partitions, using uniqueness of v and uniqueness up to transformations of u as provided in Observation VI.5.1′ in Wakker (1989b). The fitting-together is completely similar to Theorem V.3.4 in Wakker (1989b). This gives a CEU representation on $\mathscr{F}^s$. (Also the uniqueness results, given in Subsection 4.3, follow at this stage from the uniqueness results in Wakker (1989b, Observation VI.5.1′).) By continuity of $u, u(\mathscr{C})$ is an interval; it is nondegenerate by nontriviality of $\succcurlyeq$. Finally, Theorem 2.5 gives the remaining results. □

The above theorem generalizes the additive Theorem V.6.1 in Wakker (1989b). The latter is given as a corollary. In the remainder of this paragraph we discuss some details. One difference is that Wakker (1989b) used a constant-continuity condition instead of the present step equivalence assumption. The only implication of constant-continuity was, however, the existence of certainty equivalents, as derived in Lemma V.4.4 there (also straightforwardly implied by step equivalence, s-continuity, and connectedness). The second difference is that Wakker (1989b) assumed that $\mathscr{D}$ contains all open subsets, instead of the present assumption that $\mathscr{D}$ contain all preference intervals; the latter can be seen to be somewhat less restrictive.

COROLLARY 2.14 (*Extension of Wakker's* (1989b) *SEU representation to unbounded acts*).   *In the above theorem* $\succcurlyeq$ *maximizes SEU if comonotonic tradeoff consistency is strengthened to tradeoff consistency*. □

Necessary and sufficient conditions for a representation as above (also using a connected topology; compare Remark 4.8) are given in Grodal (1978). Her condition that the set of acts is a "mixture" and that all constant acts should be available implies truncation-closedness, and the availability of all step acts. Her condition of continuity with respect to nullsets implies countable additivity of the probability measure and truncation-continuity; see Subsection 4.1. Her main characterizing condition (instead of tradeoff consistency) requires "independence" of a mean groupoid operation on the set of consequences, derived from the preference relation, and using continuity for its definition. It has the disadvantage of not being elementary in terms of preferences, hence not being directly testable.

2.4.   *Savage's approach to DMUU, with an infinitely divisible state space*.   This subsection shows how to use truncation-continuity to extend the famous result of Savage (1954) to the case of unbounded utility. The approach of this subsection is called *Savage's approach*, and is characterized mainly by Condition (6) in Theorem 2.17, imposing restrictions upon the state space. A nonadditive generalization of Savage's model is very complicated; it has been obtained in Gilboa (1987). There the sure-thing principle no longer holds, and conditional preferences, as used below, can no longer be defined. So there Gilboa's axiom P7*, by itself logically independent of Savage's P7, would have to be generalized in the same spirit as Savage's P7 is generalized below. This paper, however, will follow Savage's setup as closely as possible, to obtain a direct generalization thereof. Hence the nonadditive generalization of Savage's model is not studied here.

Truncation-continuity is used again to exclude acts with unbounded or (more importantly) undefined SEU. The main difference as compared to the setups considered so far is that not the consequence space, but the state space satisfies a fineness restriction. For that reason the restrictive conditions of continuity of utility are, loosely speaking, replaced by a continuity condition for probability.

The notation $f_{-A}\alpha$ from Subsection 2.3 is also used in this subsection, as is the definition of essentiality of events, and the convention that elements of partitions are

events. Some further preparatory *notations* are given. For $f, g: \Omega \rightarrow \mathscr{C}, A \in \mathscr{A}, f_A$ is the restriction of f to A, and $g_{-A}f_A$ is identical to f on A and to g on A^c. Let us recall that, for any fixed $\omega \in \Omega, g_{-A}f(\omega)$ is equal to g on A^c and to the constant $f(\omega)$ on A.

The sure-thing principle, Savage's (1954) most famous condition, is given in the version as it is usually taken nowadays; i.e., it entails only Savage's P2. (Savage himself also included his P3, i.e., Condition (3) in Theorem 2.17, in the term sure-thing principle.) It says that a preference between two acts should only depend on the states of nature at which the acts are different. It is straightforwardly verified that this is a necessary condition for SEU maximization, and that it is implied in Corollaries 2.10 and 2.14. Many related independence-conditions from other decision theories have been listed in Wakker (1989b, §II.5). As throughout this paper, conditions are restricted to $\mathscr{F}^s$ as much as possible; this also holds for the sure-thing principle. It can be seen that the conditions in Theorem 2.17, with the sure-thing principle restricted to step acts, imply the sure-thing principle for all acts.

DEFINITION 2.15.   We say that $\succcurlyeq$ satisfies the *sure-thing principle* if

$$\forall f, g, s, t \in \mathscr{F}, A \in \mathscr{A} \ni f_{-A}s_A, g_{-A}s_A, f_{-A}t_A, g_{-A}t_A$$

$$\in \mathscr{F}: f_{-A}s_A \succcurlyeq g_{-A}s_A \Leftrightarrow f_{-A}t_A \succcurlyeq g_{-A}t_A.$$

We say $\succcurlyeq$ satisfies the *s(tep)-sure-thing principle* if the above is required only on $\mathscr{F}^s$.

□

In the presence of the *s*-sure-thing principle the following definitions are useful. Write $f_A \succcurlyeq_A g_A$ whenever $\exists h: h_{-A}f_A, h_{-A}g_A \in \mathscr{F}^s, h_{-A}f_A \succcurlyeq h_{-A}g_A$. By the *s*-sure-thing principle

$$f_A \succcurlyeq_A g_A \Rightarrow \forall h \in \mathscr{F}^s: h_{-A}f_A \succcurlyeq h_{-A}g_A.$$

Similarly, the other conditional preference relations $\succ_A$, $\preccurlyeq_A$, $\prec_A$ and $\sim_A$ are derived from their versions without the subscript A.

The only nonstraightforward complication as compared to the previous subsections is that we can no more tightly enclose bounded acts from above and below by pointwise dominating/dominated step acts, the needed consequences possibly not being available. This can be inferred from Example 4.10. Hence pointwise dominance is modified by the following condition, that was announced in Subsection 1.4. Note below that $f(\omega)_{A_i}$ denotes the restriction of the constant act $f(\omega)$ to A_i.

DEFINITION 2.16.   We say that $\succcurlyeq$ satisfies *conditional monotonicity* if, for every partition $(A_1, \ldots, A_m)$ of Ω, act $f \in \mathscr{F}$, and step act $f^s \in \mathscr{F}^s$:

$$\left(\forall \text{ essential } A_i, \omega \in A_i: f(\omega)_{A_i} \succcurlyeq_{A_i} f^s_{A_i}\right) \Rightarrow f \succcurlyeq f^s \quad \text{and}$$

$$\left(\forall \text{ essential } A_i, \omega \in A_i: f(\omega)_{A_i} \preccurlyeq_{A_i} f^s_{A_i}\right) \Rightarrow f \preccurlyeq f^s. \quad \square$$

Note that in the antecedents above only (conditional parts of) step acts are used, i.e., the constant act $f(\omega)$, and f^s. In the first implication above, f *conditionally dominates* f^s, in the second f is *conditionally dominated* by f^s. The condition is a variation of Savage's P7. Savage's P7 requires the above for single events A_i and preferences at the right-hand side of " $\Rightarrow$ " conditional on A_i, instead of considering m fold partitions as we do. However, it is elementarily verified that in Savage's setup the above condition is implied by his P7, mainly by the sure-thing principle and the richness of the set of acts in his setup. For our purposes, with fewer assumptions

about availability of acts, Savage's P7 would become too weak. Several constructions that were possible in the proofs of Savage require more caution in the present approach, where fewer auxiliary "intermediate" acts are available. Fishburn (1970, Theorem 14.1) uses a somewhat weaker condition than P7 of Savage by requiring $\succ_{A_i}$ instead of $\succcurlyeq_{A_i}$, and $\prec_{A_i}$ instead of $\preccurlyeq_{A_i}$, in the antecedent. A version in the same spirit for the present setup would meet considerable complications, again because of the generality concerning available acts. Hence such a version is not presented.

We could, as in the A & A approach and the connected topology approach, have started by giving a characterization of functionals which are integrals, and subsequently derive from this a representation result for binary relations. However, the extension of the functional to nonstep functions is now more complicated, and needs ordinal conditions like variations of the (s-)sure-thing principle (and Condition (3) below). Hence the result is derived immediately for representations of preferences, and the underlying functional result is not made explicit. As pointed out in Savage (1954, §4; see also Wakker, 1981) the assumption that $\mathscr{A}$ is a σ-algebra is essential in the theorem below. Let us compare the conditions below with Savage's Postulates. First, of course, the set $\mathscr{F}$ as determined by the DMUU Assumption 2.1 is different. Conditions (1), (4), and (5) are identical to Savage's Postulates P1, P4, and P5, respectively. His P2, P3, and P6 postulates have been weakened in Conditions (2), (3), and (6) to hold only for step acts. His P7 in the present context would be weaker than Condition (7), but in his setup would be equivalent. Conditions (8) and (9) have been added to deal with unbounded acts, and are only needed for these.

THEOREM 2.17 (*Savage for unbounded utility*).   *Under the DMUU Assumption 2.1 with $\mathscr{A}$ a σ-algebra, $\succcurlyeq$ maximizes SEU with respect to a probability measure P and a utility function u (with all expected utilities finite) whenever the following nine conditions are satisfied*:

(1) $\succcurlyeq$ *is a weak order on $\mathscr{F}$,*
(2) $\succcurlyeq$ *satisfies the s-sure-thing principle,*
(3) *if event A is essential, then $\forall \alpha, \beta \in \mathscr{C}$: $\alpha \succcurlyeq \beta \Leftrightarrow \alpha_A \succcurlyeq_A \beta_A$,*
(4) *if $\alpha \succ \beta$ and $\gamma \succ \delta$, then for events A, B: $[\beta_{-A}\alpha \succcurlyeq \beta_{-B}\alpha] \Leftrightarrow [\delta_{-A}\gamma \succcurlyeq \delta_{-B}\gamma]$,*
(5) $\succcurlyeq$ *on $\mathscr{C}$ is nontrivial,*
(6) *if, for $f, g \in \mathscr{F}^s$, $f \succ g$ and $\alpha \in \mathscr{C}$, then there exists a partition $(A_1, \ldots, A_m)$ of Ω such that $f_{-A_j}\alpha \succ g$ and $f \succ g_{-A_j}\alpha$ for all j,*
(7) $\succcurlyeq$ *is conditionally monotonic,*
(8) $\succcurlyeq$ *satisfies the step equivalent assumption,*
(9) $\succcurlyeq$ *is truncation-continuous.*

PROOF. Conditions (1)–(6) imply that there exists a unique finitely additive probability measure P on $\mathscr{A}$ and a function u: $\mathscr{C} \to \mathbb{R}$, such that $\succcurlyeq$ is represented by SEU on $\mathscr{F}^s$. Further $[\forall A \in \mathscr{A}, 0 \leqslant \mu \leqslant P(A), \exists B \subset A$ such that $P(B) = \mu]$. This was essentially proved in Savage (1954, up to §5.4) and Fishburn (1970, §14.1–14.4) and is not repeated here. It remains to show that $\succcurlyeq$ is represented by SEU, with respect to the same P and u, on the entire $\mathscr{F}$ (with $\int (u \circ f)\, dP$ real-valued for all $f \in \mathscr{F}$).

For every act $f \in \mathscr{F}^s$, all step acts equivalent to f have the same SEU value. Hence V: $\mathscr{F} \to \mathbb{R}$ can be defined by V: $f \mapsto \text{SEU}(f^s)$, where $f^s \in \mathscr{F}^s$ is any step act such that $f^s \sim f$. By transitivity of $\succcurlyeq$, this function V represents preferences. It remains to show that V is SEU. The following is an immediate consequence of conditional monotonicity, applied twice, with $m = 1$:

(2.8) $\forall f \in \mathscr{F}, \alpha \in \mathscr{C} \ni [\forall \omega \in \Omega: f(\omega) \sim \alpha]: f \sim \alpha.$

We next show:

(2.9) $f \in \mathscr{F}^b \Rightarrow V(f) = \int (u \circ f)\, dP.$

Say $\nu \preccurlyeq \mu$ are bounds of f. As in the previous subsections (compare Lemma 1.4 and Subsection 1.4) f is tightly enclosed by step acts, but now these will be conditionally, rather than pointwise, dominating/dominated. If $\mu \sim \nu$ then by (2.8), $V(f) = V(\mu) = \int (u \circ f)\, dP$, and we are done immediately. So say $\nu \prec \mu$. By rescaling, we may assume that $u(\nu) = 0$ and $u(\mu) = 1$ for simplicity of notation. Let $m \in \mathbb{N}$. Define

$$A_1 := \{\omega \in \Omega : 0 \leqslant u(f(\omega)) \leqslant 1/m\},$$

and for all $1 < i \leqslant m$,

$$A_i := \left\{\omega \in \Omega : \frac{i-1}{m} < u(f(\omega)) \leqslant \frac{i}{m}\right\}.$$

All these sets are preference intervals, so are measurable. So the step act $\phi^{-,m}$ can be defined as follows. For $i = 1, \ldots, m$, split A_i into two events $A_i^{\mu,-}$ and $A_i^{\nu,-}$. The superscript minus refers to the minus of $\phi^{-,m}$; versions with $+$ appear below. $A_i^{\mu,-}$ has probability $((i-1)/m)P(A_i)$; here $\phi^{-,m}$ is μ, i.e., has u value 1. $A_i^{\nu,-}$ has probability $((m-i+1)/m)P(A_i)$; here $\phi^{-,m}$ is ν, i.e., has u value 0. Note that

$$\int_{A_i} u \circ \phi^{-,m}\, dP = \frac{i-1}{m} P(A_i) \quad \text{for all } i.$$

For all essential

$$A_i : \left[\forall \omega \in A_i : f(\omega)_{A_i} \succcurlyeq_{A_i} \phi_{A_i}^{-,m}\right],$$

as can be inferred from the preference, for any arbitrary $f^s \in \mathscr{F}^s$, $f_{-A_i}^s f(\omega)_{A_i} \succcurlyeq f_{-A_i}^s \phi_{A_i}^{-,m}$; this preference follows from the representation by SEU that has already been established on $\mathscr{F}^s$. Hence, by conditional monotonicity, $f \succcurlyeq \phi^{-,m}$.

The step act $\phi^{+,m}$ is defined similarly to $\phi^{-,m}$. The only difference is that now, for $i = 1, \ldots, m$, A_i is split into two events $A_i^{\mu,+}$ and $A_i^{\nu,+}$, with probabilities $(i/m)P(A_i)$, respectively $((m-i)/m)P(A_i)$. Similarly, now $\phi^{+,m} \succcurlyeq f$. Note that

$$\int_{A_i} u \circ \phi^{+,m}\, dP = \frac{i}{m} P(A_i) = \left(\int_{A_i} u \circ \phi^{-,m}\, dP\right) + \frac{1}{m} P(A_i) \quad \text{for all } i.$$

It follows that $V(\phi^{-,m}) \leqslant V(f) \leqslant V(\phi^{+,m})$.
Also

$$V(\phi^{-,m}) = \int (u \circ \phi^{-,m})\, dP \leqslant \int (u \circ f)\, dP \leqslant \int (u \circ \phi^{+,m})\, dP = V(\phi^{+,m})$$

(the inequalities by elementary properties of integration) and

$$0 \leqslant \int (u \circ \phi^{+,m}) \, dP - \int (u \circ \phi^{-,m}) \, dP = \frac{1}{m}.$$

So both $V(f)$ and $\int (u \circ f) \, dP$ can be enclosed arbitrarily tightly within the same sequences, and must be identical: (2.9) follows. Note that we have used the step equivalent assumption in this derivation. Formula (2.9) can also be derived without step equivalence; this was more or less what Savage did.

Next we derive the following analogue of Lemma 1.6:

(2.10)   if f is bounded below, then $V(f) \geqslant \int (u \circ f) \, dP$.

Let ν be the lower bound of f. Define, for any $\mu \in f(\Omega)$ for which $\mu \succ \nu$ (if no such μ exists, apply (2.8)), $m \in \mathbb{N}$, $A_1, \ldots, A_m$ as below (2.9). Add $A_{m+1} := \{\omega: f(\omega) \succ \mu\}$. Define $\phi^{-,m}$ as before, with $\phi^{-,m} = \mu$ on A_{m+1} added. By conditional monotonicity,

$$V(f) \geqslant V(\phi^{-,m}) = \int (u \circ \phi^{-,m}) \, dP.$$

The supremum of these values over μ and m is $\int (u \circ f) \, dP$. Hence (2.10) follows.

In the same way one demonstrates:

(2.11)   if f is bounded above, then $V(f) \leqslant \int (u \circ f) \, dP$.

We say that $\succcurlyeq$ satisfies *above-truncation-continuity* if (2.2a) holds for all $f \in \mathscr{F}$. Next we prove (without using truncation-continuity in full force):

(2.12)   if $\succcurlyeq$ satisfies above-truncation-continuity, then for all $f \in \mathscr{F}$:

$$\int (u \circ f) \, dP \text{ is well defined and } V(f) \leqslant \int (u \circ f) \, dP.$$

We first show that $\int (u \circ f) \, dP$ is well defined. Suppose it were undefined. Let $f \sim f^s$ for $f^s \in \mathscr{F}^s$. Because $\int (u \circ f) \, dP$ is undefined, $u(\mathscr{C})$ is unbounded from below (in fact from both sides) so $g^s \in \mathscr{F}^s$ can be found with $g^s \prec f^s$, i.e., $g^s \prec f$. By above-truncation-continuity, $g^s \prec f^{\wedge \mu}$ for some μ. Contradiction follows from

$$-\infty < V(g^s) < V(f^{\wedge \mu}) \leqslant \mathrm{EU}(f^{\wedge \mu}) = -\infty,$$

where the second inequality follows since V is representing, the weak inequality follows from (2.11), and the equality follows because $\mathrm{EU}(f)$ is undefined.

So $\int (u \circ f) \, dP$ is well defined. Suppose, for contradiction, that $V(f) > \int (u \circ f) \, dP$. By step equivalence, $V(g^s) \sim V(f)$ for a $g^s \in \mathscr{F}^s$. Because $\int (u \circ f) \, dP \geqslant \inf(u \circ f(\Omega))$, we have $[\exists \omega' \in \Omega: u \circ f(\omega') < V(f)]$. By means of this consequence $f(\omega')$, and the availability of all probabilities, we can modify $g^s \in \mathscr{F}^s$ into f^s such that

$$V(f) > V(f^s) > \int (u \circ f) \, dP.$$

By above-truncation-continuity there is an above truncation $f^{\wedge\mu}$ such that still

$$(2.13) \qquad V(f^{\wedge\mu}) > V(f^s) = \int (u \circ f^s)\, dP > \int (u \circ f)\, dP.$$

However, by (2.11) and pointwise monotonicity of the integral,

$$V(f^{\wedge\mu}) \leqslant \int u \circ (f^{\wedge\mu})\, dP \leqslant \int (u \circ f)\, dP,$$

contradicting (2.13). Hence (2.12) has been proved.

Similarly $\succcurlyeq$ satisfies *below-truncation-continuity* if (2.2b) holds for all $f \in \mathscr{F}$, and we get

$$(2.14) \qquad \text{if } \succcurlyeq \text{ satisfies below-truncation-continuity, then for all } f \in \mathscr{F}:$$

$$\int (u \circ f)\, dP \text{ is well defined and } V(f) \geqslant \int (u \circ f)\, dP.$$

The theorem follows from (2.12) and (2.14). □

I think the above result shows that boundedness of utility, as implied in Savage's model, is not a consequence of mainly his postulates P6 (Condition (6) above) and/or P7 (conditional monotonicity). Each of these postulates, as well as all other postulates of Savage, hold in the above theorem. I prefer to ascribe boundedness of utility in Savage's setup to his assumption that the preference relation is defined (i.e., complete) over all acts (i.e., measurable functions from Ω to $\mathscr{C}$), thus under unbounded utility also for the acts having infinite or undefined expected utility. Note that, as a corollary of the above theorem, the used conditions do imply pointwise monotonicity.

3. DMUR, with probability distributions as alternatives.

This section studies *decision making under risk* (*DMUR*), where alternatives are probability distributions over a set $\mathscr{C}$, the consequence space. This setup should be distinguished from the A&A approach in Subsection 2.2. As compared to DMUU it assumes that an ("objective") probability distribution over Ω is given. Then any act can be associated with the probability distribution over the consequences generated by the act, and the state space Ω can be forgotten. Results of DMUR and DMUU can be related by the techniques described in Wakker (1990a). Thus it can be seen that the results of this section and Subsection 2.4 are closely related. The same appears from the proof of Savage's theorem in Fishburn (1970) and Arrow (1971). The definitions of weak order and representing function, or function to be maximized, from §2 are also used in this section.

Note below that probability distributions are assumed to be finitely additive, not necessarily countably additive. The extension to nonadditive probability measures, with a setup as sketched in Jaffray (1989), meets complications like those for the nonadditive version of Savage's approach, and is not pursued here. A *simple* probability distribution has finite support. In this paper it is also called *step*-probability distribution, for unity of terminology. Note that $\mathscr{P}$ below is not supposed to be closed under mixing. Probability distributions which assign probability one to a single element of $\mathscr{C}$ are identified with the associated element of $\mathscr{C}$; hence $\succcurlyeq$ below induces again a binary relation on $\mathscr{C}$, also denoted as $\succcurlyeq$.

ASSUMPTION 3.1 (DMUR Assumption). $\mathscr{C}$ is a nonempty set, endowed with an algebra $\mathscr{D}$ of subsets containing all one-point subsets. $\mathscr{P}$ is a set of probability

distributions on $\mathscr{D}$, containing the set $\mathscr{P}^s$ of all step probability distributions; $\succeq$ is a nontrivial binary relation on $\mathscr{P}$. For the induced binary relation $\succeq$ on $\mathscr{C}$, $\mathscr{D}$ contains all preference intervals. □

The following condition adapts the s-sure-thing principle and s-independence from Definition 2.8 to the present setup.

DEFINITION 3.2. We say that $\succeq$ satisfies $s(tep)$-independence if, for all $P, Q, R \in \mathscr{P}^s$ and $0 < \sigma < 1$,

$$P \succ Q \Rightarrow \sigma P + (1 - \sigma)R \succ \sigma Q + (1 - \sigma)R. \quad \square$$

Definition 2.7 (s-vNM-continuity) is straightforwardly adapted to the present context, by replacing $f, g, h, \mathscr{F}^s$ by $P, Q, R, \mathscr{P}^s$.

DEFINITION 3.3. We say that $\succeq$ satisfies the *step equivalent assumption* if

$$\forall P \in \mathscr{P} \, \exists P^s \in \mathscr{P}^s \colon P \sim P^s. \quad \square$$

Because the definition below is the natural adaptation of Definition 2.16, the same term is used. Note that the probability measures P_i are not required to be contained in $\mathscr{P}$, and that the sets occurring below are preference intervals, hence measurable. In the context studied mostly in the literature, e.g., in Fishburn (1982, Chapter 3), $\mathscr{P}$ is rich and is closed under convex combinations; then the condition below can be restricted to the simpler case $m = 1$. Together with mainly independence required for all probability distributions instead of just for step-probability distributions, and richness of $\mathscr{P}$, it then straightforwardly implies the condition below. Also note that all preferences in antecedents are between step-probability distributions.

DEFINITION 3.4. We say that $\succeq$ satisfies *conditional monotonicity* if, for all $P_i^s \in \mathscr{P}^s$, positive $\sigma_1, \ldots, \sigma_m$ summing to one, and probability measures P_i on $\mathscr{D}$ for which $\sum_{i=1}^m \sigma_i P_i \in \mathscr{P}$, we have

$$\forall i \colon P_i(\{\alpha \in \mathscr{C} \colon \alpha \succeq P_i^s\}) = 1 \Rightarrow \sum_{i=1}^m \sigma_i P_i \succeq \sum_{i=1}^m \sigma_i P_i^s,$$

$$\forall i \colon P_i(\{\alpha \in \mathscr{C} \colon \alpha \preceq P_i^s\}) = 1 \Rightarrow \sum_{i=1}^m \sigma_i P_i \preceq \sum_{i=1}^m \sigma_i P_i^s. \quad \square$$

In the first line above, $\sum_{i=1}^m \sigma_i P_i$ *conditionally dominates* $\sum_{i=1}^m \sigma_i P_i^s$, in the second line $\sum_{i=1}^m \sigma_i P_i$ is *conditionally dominated* by $\sum_{i=1}^m \sigma_i P_i^s$. Fishburn (1982, Chapter 3) uses a condition that is somewhat weaker by requiring $\succ$ instead of $\succeq$, and $\prec$ instead of $\succeq$, in the antecedents above. A version in the same spirit for the present setup would meet considerable complications because of the generality concerning available probability distributions (see Fishburn's reasoning on p. 29, lines 22–26). Hence such a version is not given.

It is elementarily verified that in the theorem below the conditions of s-independence and s-vNM-continuity, do imply the same conditions for all probability distributions in the presence of the other conditions. EU(P) denotes the expectation of a measurable function $u \colon \mathscr{C} \to \mathbb{R}$ under probability measure P; it is infinite respectively undefined if the expectation is.

Similarly to previous definitions, for $P \in \mathscr{P}$ and $\mu \in \mathscr{C}$, $P^{\wedge \mu}$, the *above truncation* of P at μ, is as P on $\{\alpha \in \mathscr{C} \colon \alpha \preceq \mu\} \setminus \{\mu\}$, and assigns all remaining probability $P(\{\mu\}) + P(\{\alpha \in \mathscr{C} \colon \alpha \succ \mu\})$ to μ. So $P^{\wedge \mu}(\{\alpha \in \mathscr{C} \colon \alpha \succ \mu\}) = 0$. In words, all probability mass from $\{\alpha \in \mathscr{C} \colon \alpha \succ \mu\}$ is shifted to μ. Similarly, for $\nu \in \mathscr{C}$, $P_{\vee \nu}$, the *below truncation* of P at ν, shifts all probability mass from $\{\alpha \in \mathscr{C} \colon \alpha \prec \nu\}$ to ν.

$P \in \mathscr{P}$ is *truncation-robust* (*with respect to* $\succeq$) if for all $Q^s \in \mathscr{P}^s$

(3.1a) $P \succ Q^s \Rightarrow \exists \mu \in \mathscr{C}: P^{\wedge \mu} \in \mathscr{P}, P^{\wedge \mu} \succ Q^s$ and

(3.1b) $P \prec Q^s \Rightarrow \exists \nu \in \mathscr{C}: P_{\vee \nu} \in \mathscr{P}, P_{\vee \nu} \prec Q^s$.

We call $\succeq$ *truncation-continuous* if every $P \in \mathscr{P}$ is truncation-robust.

In Theorem 3.6 below, truncation-continuity will imply *weak truncation-closedness*, i.e., $\forall P \in \mathscr{P}, \mu, \nu \in \mathscr{C} \ \exists \mu' \succeq \mu, \nu' \preccurlyeq \nu$ such that $P^{\wedge \mu'}, P_{\vee \nu'} \in \mathscr{P}$. The conditions in the literature most similar to truncation-continuity may be Axioms A5 and A5* in Fishburn (1982, §§3.2 and 3.3). Axiom A5 is, like truncation-continuity, necessary and sufficient for the EU contribution of upper and lower tails of probability distributions to tend to zero; thus it is also necessary and sufficient for finiteness of EU. It is, however, more complicated and uses (and requires available) several conditional probability distributions, and the results of "gluings-together" of these. More in the spirit of truncation-continuity is Fishburn's axiom A5*, with comparable simplicity. However, as indicated by Fishburn, his axiom A5* is necessary and sufficient for finiteness of EU only if the probability distributions are countably additive, not if they are finitely additive. The main difference between Fishburn's latter axiom and ours, and between Fishburn's definition of above (or, similarly, below) truncation and ours is, loosely speaking, as follows. We assign the probability mass taken from above the level of truncation entirely to that level of truncation, whereas Fishburn spreads that probability mass evenly over the remaining consequences. This is done by taking the conditional distribution given the event that the obtained consequence does not exceed the level of truncation. Fishburn's truncation is more drastic, leading to a more restrictive continuity condition. Even more drastic is the amputating as used by de Finetti (1974), assigning, with $\mathscr{C} = \mathbb{R}_+$, for above truncations, all truncated probability mass to 0. Next we discuss approaches to DMUR where the outcome set is rich, and utility has an interval as range. Let us consider the Menger (1934) version of the St. Petersburg paradox, where the utility function is assumed unbounded from above, with connected range. A lottery is evaluated that yields utility 2^k if after k tosses of a coin for the first time heads comes up, so has infinite expected utility. It has been observed several times that infinity of expected utility corresponds with divergence, to infinity, of the certainty equivalents of truncations of the lottery in question; truncation is obtained here by stopping after a finite number of tosses of the coin. See for instance Arrow (1971, p. 23), or Samuelson (1977, p. 34). Also Chew and Epstein (1989, end of §4) consider certainty equivalents of truncations, and convergence of these to a real number. They formulate this condition for general lotteries, and for theories that generalize expected utility, but do not make explicit their method of truncation (denoted F_{K_n}). Addition of the certainty equivalent condition and truncation continuity completes their idea. At the very end of §4 they suggest that completeness of preferences over unbounded probability measures must probably be abandoned. The following example illustrates some of the conditions discussed above:

EXAMPLE 3.5 (comparing methods of truncation). Let $\mathscr{C} = \left]0, 1\right]$, u is the identity, and $\mathscr{D}$ is the usual Borel σ-algebra. $\mathscr{P}$ contains all finitely additive probability measures on $\mathscr{D}$; they obviously have finite expectation. Suppose $\succeq$ maximizes EU. Let P be any step-probability distribution such that $EU(P) = 3/4$, say P assigns probability 1 to $\{3/4\}$. Let Q be a discrete probability measure which assigns probability 1 to $\{1/n: n \in \mathbb{N}\}$, probability $1/2$ to $\{1\}$, and probability 0 to every finite subset of $\{1/n: n \geq 2\}$; Q can for instance be generated by the act f' in Example 1.7. Obviously, Q is not countably additive; $EU(Q) = 1/2$, and $P \succ Q$. $\mathscr{P}$ is truncation-

closed, hence, as claimed at the beginning of Subsection 4.4, $\succeq$ satisfies truncation-continuity. Indeed, the EU values of below truncations of Q tend to $1/2$ and, e.g., for any $\mu \geqslant 3/4$, $\nu < 1/2$, $P \succ Q_{\vee \nu}$. However, for any $\nu \in \,]0,1[$, the conditional distribution of P on $]\nu, 1]$ assigns probability 1 to 1, so Q is preferred to none of these and Fishburn's analogue of truncation-continuity is not satisfied; it is not a necessary condition for EU maximization. $\quad \square$

THEOREM 3.6 (*unbounded utility for probability distributions*). *Under the DMUR Assumption 3.1, $\succeq$ maximizes EU with respect to a utility function u: $\mathscr{C} \to \mathbb{R}$ (with all integrals finite) whenever the following six conditions are satisfied:*

(1) $\succeq$ *is a weak order on $\mathscr{P}$,*
(2) $\succeq$ *satisfies s-independence,*
(3) $\succeq$ *satisfies s-vNM-continuity,*
(4) $\succeq$ *is conditionally monotonic,*
(5) $\succeq$ *satisfies the step equivalent assumption,*
(6) $\succeq$ *is truncation-continuous.*

PROOF. It is well known that Conditions (1) (when restricted to $\mathscr{P}^s$), (2) and (3) are necessary and sufficient for the existence of a function u: $\mathscr{C} \to \mathbb{R}$ such that $\succeq$ maximizes EU with respect to u on $\mathscr{P}^s$. Let $P \in \mathscr{P}$. All P^s equivalent to P have the same EU value, hence we can define the representing $V(P) = \mathrm{EU}(P^s)$ for any such P^s. It remains to prove that $V(P) = \mathrm{EU}(P)$. The following is an immediate consequence of conditional monotonicity, applied twice, with $m = 1$:

$$(3.2) \qquad \forall P \in \mathscr{P}, \alpha \in \mathscr{C}: P(\{\beta \in \mathscr{C}: \beta \sim \alpha\}) = 1 \Rightarrow P \sim \alpha.$$

P is *bounded (from) below* if $\exists \nu \in \mathscr{C}$: $P(\{\alpha \in \mathscr{C}: \alpha \succeq \nu\}) = 1$; then ν is a *lower bound* of P. P is *bounded (from) above* if $\exists \mu \in \mathscr{C}$: $P(\{\alpha \in \mathscr{C}: \mu \succeq \alpha\}) = 1$; then μ is an *upper bound* of P. P is *bounded* if it is both bounded from above and below. $\mathscr{P}^b$ denotes the set of all probability distributions in $\mathscr{P}$ that are bounded. We first show:

$$(3.3) \qquad P \in \mathscr{P}^b \Rightarrow V(P) = \mathrm{EU}(P).$$

Suppose ν is a lower bound and μ is an upper bound for P. Note that $\mathrm{EU}(P)$ is well defined and is contained in $[u(\nu), u(\mu)]$, so is finite. We tightly enclose P by step probability distributions; these will not be pointwise dominating/dominated, but conditionally dominating/dominated. If $\mu \sim \nu$ then, by (3.2), $V(P) = V(\mu) = \mathrm{EU}(P)$ and we are done immediately. So suppose $\mu \succ \nu$. Assume, for simplicity of notation, that $u(\nu) = 0$, $u(\mu) = 1$; this can always be obtained by rescaling. Let $m \in \mathbb{N}$. Define $A_1 := u^{-1}([0, 1/m])$, and for all $1 < i \leqslant m$, $A_i := u^{-1}([(i-1)/m, i/m])$. Define the step probability distribution $P^{-,m}$ as a mixture of step probability distributions $P^{i,-,m}$ as follows. Define $P^{i,-,m}(\{\mu\}) = (i-1)/m$, $P^{i,-,m}(\{\nu\}) = (m-i+1)/m$, and $\sigma_i := P(A_i)$ for all $1 \leqslant i \leqslant m$. $P^{-,m} := \sum_{i=1}^m \sigma_i P^{i,-,m}$. For every i with $P(A_i) > 0$ write P_{A_i} for the conditional probability distribution P given A_i. For all such i, according to the EU representation already obtained for step probability distributions, $P_{A_i}(\{\alpha \in \mathscr{C}: \alpha \succeq P^{i,-,m}\}) = 1$, hence by conditional monotonicity (with i's omitted for which $P(A_i) = 0$) $P \succeq P^{-,m}$.

The step probability distribution $P^{+,m}$ is defined similarly as $P^{-,m}$. The only difference is that now, for $i = 1, \ldots, m$,

$$P^{i,+,m}(\{\mu\}) = \frac{i}{m}, \qquad P^{i,+,m}(\{\nu\}) = (m-i)/m.$$

Now $P^{+,m} \succcurlyeq P$ is derived similarly as above. It follows that $V(P^{-,m}) \leqslant V(P) \leqslant V(P^{+,m})$. Note that

$$\mathrm{EU}(P^{i,+,m}) = \frac{i}{m} = \mathrm{EU}(P^{i,-,m}) + \frac{1}{m} \quad \text{for all } i.$$

Also

$$V(P^{-,m}) = \mathrm{EU}(P^{-,m}) \leqslant \mathrm{EU}(P) \leqslant \mathrm{EU}(P^{+,m}) = V(P^{+,m})$$

(the inequalities by elementary properties of integration) and $0 \leqslant \mathrm{EU}(P^{+,m}) - \mathrm{EU}(P^{-,m}) = 1/m$. The same inequalities hold with $V(P)$ instead of $\mathrm{EU}(P)$, by the preferences $P^{-,m} \preccurlyeq P \preccurlyeq P^{+,m}$ as established above. Hence (3.3) follows.

Next we show:

(3.4)   if P is bounded below, then $V(P) \geqslant \mathrm{EU}(P)$.

Say ν is a lower bound for P. Define, for any $\mu \succ \nu$ (if no such μ exists the case is straightforward by (3.2)), and $m \in \mathbb{N}$, $A_1, \ldots, A_m, \sigma_1, \ldots, \sigma_m$ as below (3.3), and add

$$A_{m+1} := \{\alpha \in \mathscr{C} : \alpha \succ \mu\}, \qquad \sigma_{m+1} = P(A_{m+1}).$$

Let $P^{m+1,-,m}$ assign probability 1 to μ, and $P^{-,m} := \sum_{i=1}^{m+1} \sigma_i P^{i,-,m}$. Again, by conditional monotonicity, $V(P) \geqslant V(P^{-,m}) = \mathrm{EU}(P^{-,m})$. $\mathrm{EU}(P)$ is the supremum of the latter values over μ, m. Hence (3.4) is proved. In particular we see that $\mathrm{EU}(P)$ is finite. In the same way one demonstrates:

(3.5)   if P is bounded above, then $V(P) \leqslant \mathrm{EU}(P)$.

By boundedness from above, $\mathrm{EU}(P)$ in (3.5) is well defined but might in general be $-\infty$; the above inequality shows, however, that $\mathrm{EU}(P)$ is finite. We say that $\succcurlyeq$ satisfies *above-truncation-continuity* if (3.1a) holds for all $P \in \mathscr{P}$. We prove next (without using truncation-continuity in full force):

(3.6)   if $\succcurlyeq$ satisfies above-truncation-continuity, then for all $P \in \mathscr{P}$:

$$\mathrm{EU}(P) \text{ is well defined and } V(P) \leqslant \mathrm{EU}(P).$$

We first show that $\mathrm{EU}(P)$ is well defined. Suppose it were undefined. Let $P \sim P^s$ for $P^s \in \mathscr{P}^s$. $\mathrm{EU}(P)$ being undefined, $u(\mathscr{C})$ is unbounded from below (in fact from both sides) so $Q^s \in \mathscr{P}^s$ can be found with $Q^s \prec P^s$, i.e., $Q^s \prec P$. By above-truncation-continuity, $Q^s \prec P^{\wedge \mu}$ for some μ. Contradiction follows from

$$-\infty < V(Q^s) < V(P^{\wedge \mu}) \leqslant \mathrm{EU}(P^{\wedge \mu}) = -\infty,$$

where the weak inequality follows from (3.5), the equality because $\mathrm{EU}(P)$ is undefined.

So $\mathrm{EU}(P)$ is well defined. Suppose, for contradiction, $V(P) > \mathrm{EU}(P)$. Because $\mathrm{EU}(P) \geqslant \inf(u(\mathscr{C}))$, $[\exists \alpha \in \mathscr{C} : u(\alpha) < V(P)]$. Mix α with R^s ($R^s \in \mathscr{P}^s, R^s \sim P$) to obtain $P^s \in \mathscr{P}^s$ such that $V(P) > V(P^s) > \mathrm{EU}(P)$. By above-truncation-continuity, there is an above truncation $P^{\wedge \mu}$ such that still

(3.7) $V(P^{\wedge \mu}) > V(P^s) = \mathrm{EU}(P^s) > \mathrm{EU}(P).$

However, by (3.5) and pointwise monotonicity of the integral,

$$V(P^{\wedge\mu}) \leqslant \text{EU}(P^{\wedge\mu}) \leqslant \text{EU}(P),$$

contradicting (3.7). Hence (3.6) has been proved. Similarly one defines below-truncation-continuity and gets

(3.8)   if $\succcurlyeq$ satisfies below-truncation-continuity, then for all $P \in \mathscr{P}$:

$$\text{EU}(P) \text{ is well defined and } V(P) \geqslant \text{EU}(P).$$

The theorem follows from (3.6) and (3.8).   □

Many papers have dealt with the topic of the above theorem, i.e., expected utility for unbounded probability distributions. Attention is mostly restricted to countably additive probability distributions. One of the first contributions may have been Blackwell and Girshick (1954), dealing with all step and discrete probability distributions; the utility function must be bounded in their approach.

DeGroot (1970) does deal with probability distributions with unbounded utility. He uses a two-stage approach. In the first stage utilities are derived to represent preferences on $\mathscr{P}^s$, and on $\mathscr{P}^b$ (denoted $\mathscr{P}_B$ there). In the second stage these utilities are used as primitive to define the class ($\mathscr{P}_E$, see end of §7.9 there) of all probability distributions with finite expected utility. Hence DeGroot's approach does not yield a representation in the strict sense that all conditions are directly given in elementary properties of preferences, i.e., can be directly tested in principle. Here truncation-continuity would have provided a proper means to define $\mathscr{P}_E$ directly in terms of elementary properties for preferences. Strangely enough, in the continuation of the second stage, in §7.10, DeGroot characterizes EU maximization on $\mathscr{P}_E$ through conditions for preferences; with utility available as primitive, the characterization can be obtained as a tautology!

The disadvantage of DeGroot's approach also applies to Ledyard (1971). He uses the utility function derived from the step probability distributions to define a pseudo-metric, needed in the central Axiom 5. Also he allows the utility function to be unbounded; through Axiom 4.(b) he restricts attention to bounded probability distributions.

Ryan (1974) (see also Arrow 1974) observed that in special cases, where one has restricted the set of utility functions to be considered, it is possible to at least identify subclasses of the set of distributions with finite expected utility in an empirically meaningful way. For instance if $\mathscr{C} = [0, \infty[$ and risk aversion has been established, so that the utility function must be concave (and increasing), then the distributions with finite first moment have finite expected utility.

Grandmont (1972) studies, under countable additivity, topological restrictions, and assumes closedness under countable convex combinations; part (2) in the proof of his Lemma 2 shows how the latter closedness implies boundedness of utility.

Foldes (1972) considers, for several special cases of $\mathscr{P}$ all included in the setup of this paper, several topological conditions leading to continuous utilities.

Fishburn made several contributions. Fishburn (1970, Chapters 8 and 10) dealt with bounded utility functions. Fishburn (1982, Chapter 3) extended the results to unbounded utility. Fishburn's assumptions that $\mathscr{P}$ is closed under convex combinations, contains all conditional probability measures, and contains all one-point measures, implies truncation-closedness and the inclusion of all step probability distributions. Fishburn (1976), in reaction to the exchange between Ryan (1974) and Arrow (1974), reformulated axioms from his other papers in a way suited for the case where

consequences are monetary and utility is increasing. In Fishburn (1982, Chapter 3), Axioms A4 and A0.1 in Theorem 1 (through boundedness of utility in Theorem 1) and Axiom A5 in Theorem 2 imply truncation-continuity.

A general result has been provided in Nielsen (1984). He restricts attention to the case where $\mathcal{A}$ is a σ-algebra containing all one-point subsets, and all probability measures contained in $\mathcal{P}$ are countably additive. For this case he uses topological conditions to characterize expected utility. His approach, contrary to ours, does not need (any analogue of) truncation-continuity. Compare the remark at the end of Subsection 4.1 below.

4. Further results and remarks.

4.1. *Countable additivity and set-continuity*.   In the literature the following continuity conditions are often imposed on capacities.

DEFINITION 4.1.   We call v *set-continuous* if

$$A_j \uparrow A \Rightarrow v(A_j) \uparrow v(A) \quad \text{and} \quad A_j \downarrow A \Rightarrow v(A_j) \downarrow v(A). \quad \square$$

Set-continuous abbreviates "continuous with respect to increasing and decreasing sets". It is well known that a probability measure P is *countably additive* if and only if it is set-continuous. The characterization of set-continuity of a capacity through the generated Choquet integral is straightforward. A functional $V: \mathcal{F} \to \mathbb{R}$ is *set-continuous* if

$$\left[A_j \uparrow A \Rightarrow V(1_{A_j}) \uparrow V(1_A)\right] \quad \text{and} \quad \left[A_j \downarrow A \Rightarrow V(1_{A_j}) \downarrow V(1_A)\right].$$

One directly observes:

LEMMA 4.2.   *Suppose V is the Choquet integral. Then the associated capacity v is continuous if and only if V is set-continuous.* $\square$

Hence set-continuous or countably additive versions of the previous theorems can be obtained as follows.

PROPOSITION 4.3.   *In Theorem 1.12, Theorem 1.13, Corollary 1.14, and Theorem 1.15, the capacity in question is continuous if and only if V is set-continuous.* $\square$

The characterization of set-continuity of a capacity in terms of a preference relation is through the following condition, given for σ-algebras and additive probabilities in Fishburn (1982, F7 in §10.3). We say that $\succcurlyeq$ is *set-continuous* if

(4.1) $\forall f^s \in \mathcal{F}^s, \alpha \succ \beta, A_j \uparrow A: \beta_{-A}\alpha \succ f^s \Rightarrow \exists j: \beta_{-A_j}\alpha \succ f^s$ and

(4.2) $\forall f^s \in \mathcal{F}^s, \alpha \succ \beta, A_j \downarrow A: \beta_{-A}\alpha \prec f^s \Rightarrow \exists j: \beta_{-A_j}\alpha \prec f^s.$

By substitution one straightforwardly verifies:

PROPOSITION 4.4 (*countable additivity or set-continuity*).   *The capacities in Theorems 2.5, 2.9, and 2.13 are set-continuous if and only if $\succcurlyeq$ is set-continuous. The probability measures in Corollaries 2.10 and 2.14, and in Theorem 2.17, are countably additive if and only if $\succcurlyeq$ is set-continuous.* $\square$

In the context of countable additivity, several parts in the analysis of this paper could have been abbreviated and simplified. Compare §3.3 in Fishburn (1982). One can then adopt a monotone continuity condition as in Arrow (1971) (or continuity with respect to nullsets in the terminology of Grodal 1978), to imply both truncation-continuity and set-continuity.

4.2. *Rank-dependent utility.* This subsection extends Quiggin's (1982) rank-dependent utility to unbounded probability distributions. Although this result falls somewhat outside the scope of this paper, which only considered the additive case for DMUR, the result is nevertheless given in view of its importance. The extension is only formulated for intervals $\mathscr{C}$ and countably additive probability distributions. The extension to any connected topological space and finitely additive probability distributions is straightforward, but omitted for brevity. Under the DMUR Assumption 3.1, $\succcurlyeq$ maximizes *rank-dependent utility* if there exists a function $u \colon \mathscr{C} \to \mathbb{R}$, and a nondecreasing (probability transformation) function $\phi \colon [0, 1] \to [0, 1]$ with $\phi(0) = 0$ and $\phi(1) = 1$, such that $\succcurlyeq$ is represented by

$$P \mapsto \int_{\mathbb{R}_+} \phi\big(P(u^{-1}[\tau, \infty[)\big)\, d\tau + \int_{\mathbb{R}_-} \big[\phi\big(P(u^{-1}[\tau, \infty[)\big) - 1\big]\, d\tau.$$

Stochastic dominance means that $P \succcurlyeq Q$ whenever

$$\forall \alpha \in \mathbb{R} \colon P(]-\infty, \alpha]) \leqslant Q(]-\infty, \alpha]).$$

The step equivalent assumption, and truncation-continuity, have been defined in Definition 3.3 and below Formula (3.1).

COROLLARY 4.5 (*unbounded utility for Quiggin's (1982) rank-dependent utility*). *Suppose the DMUR Assumption 3.1 holds, all elements of $\mathscr{P}$ are countably additive, and $\mathscr{C} \subset \mathbb{R}$ is an interval. Then $\succcurlyeq$ maximizes rank-dependent utility on $\mathscr{P}$ whenever the following five conditions hold*:
(1) $\succcurlyeq$ *maximizes rank-dependent utility on $\mathscr{P}^s$, with u continuous,*
(2) $\succcurlyeq$ *is a weak order on $\mathscr{P}$,*
(3) $\succcurlyeq$ *satisfies stochastic dominance,*
(4) $\succcurlyeq$ *satisfies the step equivalent assumption,*
(5) $\succcurlyeq$ *satisfies truncation-continuity.*

PROOF. The corollary follows from Theorem 2.5 as follows. Define $\Omega = [0, 1]$, endowed with the Lebesgue measure (i.e., the uniform distribution), $\mathscr{A}$ and $\mathscr{D}$ are the Borel σ-algebras, and $\mathscr{F}$ is the set of random variables from Ω to $\mathscr{C}$. Identify random variables with the probability distributions generated over $\mathscr{C}$. Then rank-dependent utility maximization over $\mathscr{P}$ is equivalent to Choquet expected utility maximization over $\mathscr{F}$ with capacity $v := \phi \circ \lambda$ (λ denotes the Lebesgue measure), and stochastic dominance is equivalent to pointwise monotonicity. The latter is most easily seen by relating probability distributions to the generalized inverses of their distribution functions. $\square$

Axioms to characterize maximization of rank-dependent utility on $\mathscr{P}^s$ have been provided in Quiggin (1982; here $\phi(1/2) = 1/2$), Yaari (1987; here u is the identity), Chew (1989), Nakamura (1992), and Wakker (1990d).

4.3. *Uniqueness results.* Below, *cardinal* abbreviates "unique up to a positive affine transformation", and *ordinal* abbreviates "unique up to a strictly increasing

transformation". The uniqueness results for the above theorems are standard. In Theorem 1.12, 1.13, Corollary 1.14, and Theorem 1.15, the capacity v is uniquely determined. In Theorem 2.9, v is unique and U is cardinal, as in Schmeidler (1989); in Corollary 2.10 the same holds with now v a probability measure. For Theorem 2.13 the same uniqueness results hold as in Observation VI.5.1' in Wakker (1989b), for Corollary 2.14 the same uniqueness results hold as in Observation V.3.4' in Wakker (1989b); for brevity these results are not repeated here. In Theorem 2.17, P is unique and u is cardinal as in Savage (1954), in Theorem 3.6 u is cardinal as in Fishburn (1982, Chapter 3).

4.4. *Necessity of the conditions.* Many of the conditions used in several theorems were not only sufficient, but also necessary, for the derived implications. Pointwise monotonicity is necessary whenever used, as is truncation-continuity whenever weak truncation-closedness holds. For representations of binary relations it is always necessary that the binary relations are weak orders. Leaving aside trivial observations, we further mention the following necessary conditions:

(1) in Theorem 1.13 and Corollary 1.14: Condition (1) and comonotonic additivity (the latter by Theorem 1.3),

(2) in Theorem 1.15: Condition (1) and comonotonic mixability,

(3) in Theorem 2.9 and Corollary 2.10: s-vNM-continuity and s-comonotonic (or s-) independence,

(4) in Theorem 2.13 and Corollary 2.14: s-continuity, (comonotonic) tradeoff consistency,

(5) in Theorem 2.17: the s-sure-thing principle, Conditions (3) and (4), conditional monotonicity,

(6) in Theorem 3.6: s-independence, s-vNM-continuity, and conditional monotonicity.

The conditions above that were restricted to step acts hold as well if they are extended to nonstep acts. In particular, all conditions above with s-prescript hold as well if this prescript is deleted (and sometimes alternatives are restricted to those that are in the domain of $\succcurlyeq$). Not necessary are structural or richness conditions, such as I being an interval, $\mathscr{C}$ being a connected separable topological space or a set containing "many" probability distributions, Condition (6) in Theorem 2.17, or the availability of all step probability distributions over $\mathscr{C}$; further weak truncation-closedness, as implied by truncation-continuity, is nonnecessary; given weak truncation-closedness, truncation-continuity is necessary. Step equivalence in full generality is not necessary, but in the presence of the other richness assumptions of this paper it is.

4.5. *Further remarks and results.*

REMARK 4.6 (mixture spaces). The only property of $\mathscr{C}$ used in Theorem 2.9 and Corollary 2.10 is that $\mathscr{C}$ is a mixture space (defined in Fishburn 1982 and Wakker 1989b), and that EU is affine. Another special case of mixture spaces is the case where $\mathscr{C}$ is an interval and consequences designate money, or $\mathscr{C}$ is $\mathbb{R}_+^m$ and consequences are commodity bundles, each time with linear/affine utility. The first case concerns the works of de Finetti (1937, 1972, 1974), Lindley (1982), Yaari (1987), and Chateauneuf (1991). □

The following simple example ensures that the results of this paper do not fall victim to the same problem as Savage (1954), i.e., there is no "hidden" way in which boundedness of utility would be implied.

EXAMPLE 4.7 (unbounded utility; examples). The results of the functional approach apply to the functionals V both in (a) and (b) below. The results of the A & A

approach apply to $\succcurlyeq$ in (b) below, the results of the connected topology approach, as well as those of Savage's approach, apply to $\succcurlyeq$ in (a) below. The results of DMUR apply to $\succcurlyeq$ in (c) below.

(a) Let $\Omega = [0, 1]$, $\mathscr{A}$ is the Borel-σ-algebra, $\mathscr{C} = \mathbb{R}$, $\mathscr{D}$ is the Borel-σ-algebra, P on $\mathscr{A}$ is the Lebesgue measure, and $u: \mathscr{C} \to \mathbb{R}$ is the identity. $\mathscr{F}$ contains the step acts together with any arbitrary subset of the set of all measurable acts from Ω to $\mathscr{C}$ with finite expectation, and their truncations; $\succcurlyeq$ is represented by V = expectation (= expected utility).

(b) Ω, $\mathscr{A}$, P are as above; $\mathscr{C}$ consists of all (step =) simple probability distributions over $\mathbb{R}$, $U: \mathbb{R} \to \mathbb{R}$ is the identity, and $u = \mathrm{EU}: \mathscr{C} \to \mathbb{R}$ is expected value. $\mathscr{D}$ contains $\mathrm{EU}^{-1}(E)$ for all Borel-measurable sets E. $\mathscr{F}^s$ consists, as usual, of all step acts from Ω to $\mathscr{C}$. $\mathscr{F}$ contains $\mathscr{F}^s$, together with any arbitrary subset of the set of all measurable acts from Ω to $\mathscr{C}$ with finite expectation of EU, and their truncations; $\succcurlyeq$ is represented by V = the expectation of EU.

(c) Let $\mathscr{C} = \mathbb{R}$, $\mathscr{D}$ is the Borel-σ-algebra, $u: \mathscr{C} \to \mathbb{R}$ is the identity. $\mathscr{P}$ consists of $\mathscr{P}^s$, together with any arbitrary subset of the set of all finitely additive probability measures on $\mathscr{D}$ with finite expectation, and their truncations; $\succcurlyeq$ is represented by V = the expectation of u.   □

REMARK 4.8 (deleting topological separability).   The requirement of topological separability in Subsection 2.3 has been added mainly for simplicity of formulation. It can nearly always be omitted, with the exception of the case of exactly one essential state, which comes down to the case that there is no uncertainty. Compare Remark A3.1 in Wakker (1989b).   □

REMARK 4.9 (circularity in the definition of truncation-robust alternatives).   There is some circularity in the identification of alternatives with finite value in this paper: To define $\succcurlyeq$ (or the function V) its domain should first be defined; however, in the definition of the elements of its domain (the truncation-robust alternatives) already $\succcurlyeq$ (or V) was used! This presentation was chosen for conciseness and accessibility. Below a procedure is sketched that is more constructive by circumventing the circularity:

Stage 1. $\mathscr{F}^s$ is defined.

Stage 2. $\mathscr{F}^b$ is defined.

Stage 3.   An alternative f, bounded from below, is incorporated if and only if there is a step alternative that is preferred to all step alternatives that are pointwise, or conditionally, dominated by f. An alternative f, bounded from above, is incorporated if and only if there is a step alternative that is dispreferred to all step alternatives that pointwise, or conditionally, dominate f.

Stage 4.   An alternative f, unbounded from both sides, is incorporated if and only if its truncations (or a subset of truncations, related to a subset of consequences unbounded from above and below) can be incorporated already at Stage 3.

By this procedure, in each stage the extension of domain either did not use $\succcurlyeq$ (or V) at all, or only used it as already defined on domains that could be established in previous stages. All truncation-robust alternatives can be incorporated this way.   □

EXAMPLE 4.10 (necessity of conditional monotonicity in Theorem 3.6; stochastic dominance instead does not suffice; similarly, pointwise monotonicity instead of conditional monotonicity is not sufficient in Theorem 2.17, i.e., Savage's approach). This example will be formulated in the most popular context, DMUR. It can be adapted to the other contexts by constructing an underlying, sufficiently rich, state space Ω, as described in detail in Wakker (1990a).

Let $\mathscr{C} = \{-1/j: j \in \mathbb{N}\} \cup \{1 + 1/j: j \in \mathbb{N}\}$, $\mathscr{D} = 2^{\mathscr{C}}$. $\mathscr{P}$ contains all finitely additive probability distributions over $\mathscr{C}$, and u denotes the identity function; $\succcurlyeq$

maximizes expected value on $\mathscr{P}^s$. To describe examples, let

$$\sum_{j=1}^{\infty} P\left(-\frac{1}{j}\right) = \alpha, \qquad P\left(\left\{-\frac{1}{j} : j \in \mathbb{N}\right\}\right) = \alpha + \beta,$$

$$P\left(\left\{1 + \frac{1}{j} : j \in \mathbb{N}\right\}\right) = \gamma + \delta, \qquad \sum_{j=1}^{\infty} P\left(1 + \frac{1}{j}\right) = \delta,$$

where $\alpha, \beta, \gamma, \delta$ are nonnegative and sum to one. For $P \in \mathscr{P}$, F_P denotes the distribution function, i.e., $F_P(t) = P(-\infty, t]$ for all $t \in \mathbb{R}$. Note that, for all $P \in \mathscr{P}$, F_P is constant on $[0, 1]$. Now

$$\lim_{j \to \infty} F_P\left(-\frac{1}{j}\right) = \alpha, \quad P\left(\left\{-\frac{1}{j} : j \in \mathbb{N}\right\}\right) = F_P(0) = F_P(1) = \alpha + \beta,$$

$$\lim_{j \to \infty} F_P\left(1 + \frac{1}{j}\right) = \alpha + \beta + \gamma.$$

For all countably additive P, $\beta = 0$ because $P(\{0\}) = 0$, and $\gamma = 0$ by countable additivity, which implies right-continuity of F_P. For finitely additive P, β and γ may very well be positive. For any step probability distribution P^s that stochastically dominates P, $F_{P_s}(-1/j) \leqslant \alpha$ for all j, so that also $F_{P_s}(0) = F_{P_s}(1) \leqslant \alpha$. For any step probability distribution P^s, stochastically dominated by P, $F_{P_s}(1 + 1/j) \geqslant \alpha + \beta + \gamma$ for all j, so that also $F_{P_s}(1) = F_{P_s}(0) \geqslant \alpha + \beta + \gamma$. The gap in expected value between dominating and dominated step probability distributions is at least $\beta + \gamma$. The expectation of P as above simply is undetermined if β or γ is nonzero. Suppose $\succcurlyeq$ is represented by the functional

$$V : P \mapsto \sum_{j=1}^{\infty} \left(-\frac{1}{j}\right) P\left(-\frac{1}{j}\right) + \sum_{j=1}^{\infty} \left(1 + \frac{1}{j}\right) P\left(1 + \frac{1}{j}\right) + \frac{\gamma}{2}.$$

It is straightforwardly verified that all conditions of Theorem 3.6 other than conditional monotonicity are satisfied, and that strict stochastic dominance is satisfied as well. Independence and vNM-continuity even hold in full strength, by linearity and continuity of the functional in probability. But conditional monotonicity is not satisfied: Let P be such that $\gamma = 1$; let Q^s assign probability $1/3$ to -1 and probability $2/3$ to 2. So the expected value of Q^s, which is equal to $V(Q^s)$, is 1. Every positive consequence $1 + 1/j$ is preferred to Q^s. P assigns probability 1 to such consequences, i.e., $P(\{1 + 1/j : j \in \mathbb{N}\}) = 1$. By conditional monotonicity (with $m = 1$) we should find $P \succcurlyeq Q^s$. However, $V(P) = \gamma/2 = 1/2 < 1 = V(Q^s)$, so $P \prec Q^s$: conditional monotonicity is violated. $\square$

4.6. *The unity in the different approaches.* Let us shortly repeat the approach to integral representations, to point out the unity in the different results. In a first stage one obtains the representation for step alternatives. For the functional approach this is mainly obtained through (comonotonic) additivity or mixability, in the A & A approach the theorem of Anscombe and Aumann (1963) (additive) or Schmeidler (1989) (nonadditive) is used, in the connected topology approach Theorem IV.2.7 (additive) or VI.5.1 (nonadditive) of Wakker (1989b) is used, in Savage's approach the proof up to §5.4 of Savage (1954) is used (which does not use his condition P7, i.e., conditional monotonicity), and for DMUR the result of von Neumann and

Morgenstern (1944) is used. It can be seen that there are close relations between the different approaches. For instance, there is a close relationship between (comonotonic) additivity/mixability in the functional approach, the sure-thing principle, and (comonotonic) independence in the A & A approach and DMUR. Tradeoff consistency in the connected topology approach can be seen to be a strengthening of the s-sure-thing principle that (see Wakker, 1989a, §10) in the presence of mixture operations is straightforwardly implied by the independence conditions.

The extendability of the integral representation to bounded alternatives is ensured by monotonicity, because every bounded alternative can be approximated from above and below by step alternatives. Here monotonicity is pointwise in Figure 2(a) and, for the functional approach, in Lemmas 1.4 and 1.5; based on these are Theorems 1.12, 1.13, and 1.15 for the functional approach, Theorem 2.5 (through Lemma 2.3), and then, building on the latter theorem, Theorem 2.9 for the A & A approach, Theorem 2.13 for the connected topology approach, and Corollary 4.5 for rank-dependent utility. Monotonicity is conditional in Figure 2(b) and Subsection 1.4; based on this are Theorem 2.17 (through (2.8) and (2.9) for Savage's approach, and Theorem 3.6 (through (3.2) and (3.3)) for DMUR.

The extension to any unbounded alternative with finite expected utility is obtained by truncation continuity. For the functional approach this is done in the results of Subsection 1.6 (based on Lemma 1.11). Theorem 2.5 obtains the extension through Lemma 2.4, and prepares for Theorems 2.9, 2.13, and Corollary 4.5. In Savage's approach the extension is obtained at the end of the proof of Theorem 2.17, and in DMUR at the end of the proof of Theorem 3.6. Again, the reasonings are very similar.

5. Conclusion. It is hoped that, after study of this paper, the reader will not only understand the specific theorems of the paper, but also the general principle to derive infinite-dimensional integral representations from finite-dimensional ones.

References

Allais, M. (1953). Fondements d'une Théorie Positive des Choix Comportant un Risque et Critique des Postulats et Axiomes de l'Ecole Américaine. *Colloques Internationaux du Centre National de la Recherche Scientifique* 40, Econométrie, Centre National de la Recherche Scientifique, Paris; 257–332. Translated into English, with additions, as The Foundations of a Positive Theory of Choice Involving Risk and a Criticism of the Postulates and Axioms of the American School, in M. Allais and O. Hagen (1979) Eds., *Expected Utility Hypotheses and the Allais Paradox*, Reidel, Dordrecht, The Netherlands, 27–145.

Anscombe, F. J. and Aumann, R. J. (1963). A Definition of Subjective Probability. *Ann. Math. Statistics* **34** 199–205.

Arrow, K. J. (1971). *Essays in the Theory of Risk-Bearing*. North-Holland, Amsterdam.

_____ (1974). The Use of Unbounded Utility Functions in Expected-Utility Maximization: Response. *Quarterly J. Econ.* **88** 136–138.

Aumann, R. J. (1962). Utility Theory without the Completeness Axiom. *Econometrica* **30** 445–462.

Bhaskara Rao, K. P. S. and Bhaskara Rao, M. (1983). *Theory of Charges*. Academic Press, London.

Blackwell, D. and Girshick, M. A. (1954). *Theory of Games and Statistical Decisions*. Wiley, New York.

Chateauneuf, A. (1991). On the Use of Non-additive Probabilities in Modeling Uncertainty Aversion and Risk Aversion. *J. Math. Econ.* (forthcoming).

Chew, S. H. (1989). The Rank-Dependent Quasilinear Mean. *Ann. Statistics* (tentatively accepted).

_____ and Epstein, L. G. (1989). Axiomatic Rank-Dependent Means. *Ann. Oper. Res.* **19** 299–309.

Choquet, G. (1953–1954). Theory of Capacities. *Ann. l'Institut Fourier* **5** (Grenoble) 131–295.

de Finetti, B. (1937). La Prévision: Ses Lois Logiques, ses Sources Subjectives. *Ann. l'Institut Henri Poincaré* **7** 1–68. Translated into English by H. E. Kyburg, Foresight: Its Logical Laws, Its Subjective Sources, in H. E. Kyburg and H. E. Smokler (1964) Eds., *Studies in Subjective Probability*, Wiley, New York, 53–118; 2nd edition (1980), Krieger, New York.

_____ (1972). *Probability, Induction and Statistics*. Wiley, New York.

_____ (1974). *Theory of Probability*. Vol. I, Wiley, New York.

DeGroot, M. H. (1970). *Optimal Statistical Decisions*. McGraw-Hill, New York.

Dellacherie, C. (1970). Quelques Commentaires sur les Prolongements de Capacités. *Seminaire de Probabilités V Strasbourg*, Lecture Notes in Math., Vol. 191, Springer-Verlag, Berlin.

Dunford, N. and Schwartz, J. T. (1958). *Linear Operators, Part I*. Interscience Publishers, New York.

Fishburn, P. C. (1970). *Utility Theory for Decision Making*. Wiley, New York.

_____ (1976). Unbounded Utility Functions in Expected Utility Theory. *Quart. J. Econ.* **90** 163–168.

_____ (1982). *The Foundations of Expected Utility*. Reidel, Dordrecht.

_____ (1988). *Nonlinear Preference and Utility Theory*. Johns Hopkins University Press, Baltimore, MD.

Foldes, L. (1972). Expected Utility and Continuity. *Review of Econ. Studies* **39** 407–421.

Gilboa, I. (1987). Expected Utility with Purely Subjective Non-Additive Probabilities. *J. Math. Econ.* **16** 65–88.

Grandmont, J.-M. (1972). Continuity Properties of a von Neumann-Morgenstern Utility. *J. Econ. Theory* **4** 45–57.

Grodal, B. (1978). Some Further Results on Integral Representation of Utility Functions. Institute of Economics, University of Copenhagen, Copenhagen, Denmark. Forthcoming in Vind, K. *Independent Preferences*.

Jaffray, J. Y. (1989). Linear Utility Theory for Belief Functions. *Oper. Res. Letters* **8** 107–112.

Kelley, J. L. (1955). *General Topology*. Van Nostrand, London.

Koopmans, T. C. (1972). Representations of Preference Orderings with Independent Components of Consumption and Representations of Preference Orderings over Time, in C. B. McGuire and R. Radner (Eds.), *Decision and Organization*, North-Holland, Amsterdam, 57–100.

Ledyard, J. O. (1971). A Pseudo-Metric Space of Probability Measures and the Existence of Measurable Utility. *Ann. Math. Statistics* **42** 794–798.

Lindley, D. V. (1982). Scoring Rules and the Inevitability of Probability. *International Statistical Review* 1–26.

Lippman, S. A. (1975). On Dynamic Programming with Unbounded Rewards. *Management Sci.* **21** 1225–1233.

Machina, M. J. (1982). Expected Utility Analysis without the Independence Axiom. *Econometrica* **50** 277–323.

_____ (1987). Choice under Uncertainty: Problems Solved and Unsolved. *Economic Perspectives* **1** 121–154.

Mak, K.-T. (1987). Coherent Continuous Systems and the Generalized Functional Equation of Associativity. *Math. Oper. Res.* **12** 597–625.

Menger, K. (1934). Das Unsicherheitsmoment in der Wertlehre. *Zeitschrift für National-ökonomie* **51** 459–485. Translated into English by W. Schoellkopf as The Role of Uncertainty in Economics, in Shubik, M. (Ed.) (1967), *Essays in Math. Econ. in Honor of Oskar Morgenstern*, Princeton University Press, Princeton, NJ.

Nakamura, Y. (1992). Multi-Symmetric Structures and Non-Expected Utility. *J. Math. Psychology* (forthcoming).

Nielsen, L. T. (1984). Unbounded Expected Utility and Continuity. *Math. Social Sciences* **8** 201–216. (See also Nielsen, L. T. (1987). Corrigenda. *Math. Social Sciences* **14** 193–194.)

Quiggin, J. (1982). A Theory of Anticipated Utility. *J. Econ. Behaviour and Organization* **3** 323–343.

Ramsey, F. P. (1931). Truth and Probability, in *The Foundations of Mathematics and Other Logical Essays*, 156–198, Routledge and Kegan Paul, London. Reprinted in H. E. Kyburg and H. E. Smokler (Eds.), (1964), *Studies in Subjective Probability*, Wiley, New York, 61–92.

Ryan, T. M. (1974). The Use of Unbounded Utility Functions in Expected-Utility Maximization: Comment. *Quarterly J. Econ.* **88** 133–135.

Samuelson, P. A. (1977). St. Petersburg Paradoxes: Defanged, Dissected and Historically Described. *J. Econ. Literature* **15** 24–55.

Savage, L. J. (1954). *The Foundations of Statistics*. Wiley, New York (second edition 1972, Dover, New York).

Schmeidler, D. (1986). Integral Representation without Additivity. *Proc. Amer. Math. Soc.* **97** 255–261.

_____ (1989). Subjective Probability and Expected Utility without Additivity. *Econometrica* **57** 571–587.

Shapley, L. S. (1977). The St. Petersburg Paradox: A Con Game? *J. Economic Theory* **14** 439–442.

Toulet, C. (1986). An Axiomatic Model of Unbounded Utility Functions. *Math. Oper. Res.* **11** 81–94.

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

von Weizsäcker, C. (1965). Existence of Optimal Programs of Accumulation for an Infinite Time Horizon. *Review of Econ. Studies* **32** 85–104.

Wakker, P. P. (1981). Agreeing Probability Measures for Comparative Probability Structures. *Ann. Statistics* **9** 658–62.

_____ (1985). Continuous Expected Utility for Arbitrary State Spaces. *Methods of Oper. Res.* **50** 113–129.

_____ (1989a). Continuous Subjective Expected Utility with Nonadditive Probabilities. *J. Math. Econ.* **18** 1–27.

_____ (1989b). *Additive Representations of Preferences, A New Foundation of Decision Analysis*. Kluwer Academic Publishers, Dordrecht, The Netherlands.

_____ (1990a). Under Stochastic Dominance Choquet-Expected Utility and Anticipated Utility Are Identical. *Theory and Decision* **29** 119–132.

_____ (1990b). Characterizing Optimism and Pessimism Directly through Comonotonicity. *J. Econ. Theory* **52** 453–463.

_____ (1990c). A Behavioral Foundation for Fuzzy Measures. *Fuzzy Sets and Systems* **37** 327–350.

_____ (1990d). Separating Marginal Utility and Probabilistic Risk Aversion. Working Paper FSB-9005, Duke University, Fuqua School of Business, Durham, NC.

_____ and Tversky, A. (1991). An Axiomatization of Cumulative Prospect Theory. University of Nijmegen, NICI, Nijmegen, The Netherlands.

Weibull, J. W. (1985). Discounted-Value Representations of Temporal Preferences. *Math. Oper. Res.* **10** 244–250.

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

_____ (1988). A Controversial Proposal Concerning Inequality Measurement. *J. Econ. Theory* **44** 381–397.

P. Wakker: University of Leiden, Medical Decision-Making Unit, Leiden, The Netherlands