Econ Theory (2009) 39:157–176
DOI 10.1007/s00199-007-0302-z

RESEARCH ARTICLE

A new integral for capacities

Ehud Lehrer

Received: 18 June 2006 / Accepted: 10 October 2007 / Published online: 30 October 2007

Abstract A new integral for capacities is introduced and characterized. It differs from the Choquet integral on non-convex capacities. The main feature of the new integral is concavity, which might be interpreted as uncertainty aversion. The integral is extended to fuzzy capacities, which assign subjective expected values to random variables (e.g., portfolios) and may assign subjective probability only to a partial set of events. An equivalence between the minimum over sets of additive capacities (not necessarily probability distributions) and the integral w.r.t. fuzzy capacities is demonstrated. The extension to fuzzy capacities enables one to calculate the integral also in cases where the information available is limited to a few events.

Keywords Capacities · Non-additive probability · Decisions under uncertainty · Uncertainty aversion · Concave integral · Choquet integral · Fuzzy capacities · Large core

JEL Classification C71 · D80 · D81 · D84

1 Introduction

In many economic activities individuals often face risks and uncertainties concerning future events. The probabilities of these events are rarely known, and individuals are

I wish to thank Eran Hanany, David Schmeidler, Eilon Solan and especially Yaron Azrieli and the anonymous referee of *Economic Theory* for their helpful comments.

E. Lehrer (✉)
School of Mathematical Sciences, Tel Aviv University, Tel Aviv 69978, Israel
e-mail: lehrer@post.tau.ac.il

E. Lehrer
INSEAD, Bd. de Constance, 77305 Fontainebleau Cedex, France

left to act on their subjective beliefs. Since the work of Ellsberg (1961), the conventional theory based on (additive) expected utility has become somewhat controversial, both on descriptive and normative grounds. There is a cumulative indication that individuals often do not use regular (additive) subjective probability. Rather, they exhibit what is referred to as an uncertainty aversion.[1]

Schmeidler (1989) proposed one of the most influential alternative theories to that of additive subjective probabilities. In Schmeidler's model, individuals make assessments that fail to be additive across disjoint events. The expected value of utility with respect to a non-additive probability distribution is defined according to the Choquet integral. The decision maker chooses the act that maximizes the expected utility. Following Choquet, a possible non-additive probability is referred to as a capacity.

Since Schmeidler's breakthrough, the Choquet integral has been extensively used in decision theory (see Gilboa 1987; Wakker 1989; Sarin and Wakker 1992). Dow and Werlang (1992, 1994) applied the Choquet integral to game theory and finance. Schmeidler (1986) and Groes et al. (1998) provided a few characterizations of the Choquet integral.

Another prominent integral is the Sugeno (or fuzzy) integral (see Sugeno 1974). It is expressed in maximum–minimum terms and it corresponds to the notion of the median, rather than to that of the average. As opposed to the Choquet integral and the one introduced here, the Sugeno integral does not coincide with the regular integral when the capacity is additive.[2]

This paper presents a new integral with respect to capacities, which differs from the Choquet integral on non-convex capacities. The new integral makes use of the concavification of a cooperative game that appeared in Weber (1994) and later in Azrieli and Lehrer (2007b). It is axiomatically characterized in two ways.

The key property of the new integral is concavity. This means that the sum of the integrals of two functions is less than or equal to the integral of the sum. In the context of decision under uncertainty this property might be interpreted as uncertainty aversion.

Three more axioms are necessary in order to characterize the integral. The first requires that when the underlying probability space consists of one point, the integral coincides with the conventional integral. The second is an axiom of monotonicity with respect to capacities. It states that an additive capacity P assigns to every subset a value which is greater than or equal to that assigned by v, if and only if the integral of any non-negative function with respect to P is greater than or equal to the integral taken with respect to v.

The last axiom states that when integrating an indicator of a set S, the integral depends only on the values that the capacity takes on the subsets of S. In other words, the integral of an indicator of S does not depend on the values that the capacity ascribes to any event outside of S.

[1] A myriad of empirical evidence of choices, that are not consistent with conventional subjective probability and expected utility, have been documented in the literature (see, Camerer and Weber 1992; Starmer 2000).

[2] For further discussion of this issue the reader is referred to Murofushi and Sugeno (1991).

In Sect. 9 we introduce an integral w.r.t. fuzzy capacities. Fuzzy capacities assign subjective expected values to some, but not all, random variables (e.g., portfolios). In particular, a fuzzy capacity may assign subjective probabilities only to some events and not to all. The new integral aggregates all available information, and enables one to calculate an average value also when there is partial information and the capacity does not provide the likelihood of every possible event.

The integral w.r.t.fuzzy capacities is inspired by Azrieli and Lehrer (2007a) who used the operational technique (concavification and alike) extensively and employed it to investigate cooperative population games.

It turns out that a strong relation exists between the minimum over additive capacities and the new integral. A full equivalence between the representation of an order over random variables as a minimum over additive capacities[3] and a representation by the integral w.r.t. fuzzy capacities is shown in Sect. 9.

It might be that a capacity is specified over a subset of events and not over all of them. The definition of fuzzy capacities also covers this case. That is, the integral w.r.t. fuzzy capacities enables one to define the integral of a partially-specified capacity. This is particularly important when the capacity is additive (i.e., a regular probability distribution) and the decision maker is not informed of the probability of all events.

The paper is organized as follows. Section 2 presents the new integral and Sect. 3 illustrates it through a few examples including Ellsberg's paradox. Section 4 compares the new integral and Choquet integral. Example 3 is meant to convince the reader that sometimes the new integral results in more intuitive decisions than those derived from Choquet integral. This section also shows that the new integral is an extension of Lebesgue integral. Section 5 provides the two characterization theorems whose proofs appear in Sect. 6. Section 7 refers to capacities with large core where the new integral is obtained as the minimum over the core's members of the corresponding expectations. Section 8 discusses first order stochastic dominance. Section 9 extends the integral to fuzzy capacities. Final comments are found in Sect. 10. The first comment refers to an example by Machina (2007), the second to risk measures and their relation to the new integral, and the third comment is about the extension of the integral to large spaces.

2 The new integral

A *capacity* is a function v that assigns a non-negative real number to every subset of a finite set N ($|N| = n$) and satisfies (i) $v(\emptyset) = 0$; and (ii) if $S \subseteq T \subseteq N$, then $v(S) \leq v(T)$. Such a capacity is said to be defined over N. A capacity P defined over N is *additive* if for any two disjoint subsets $S, T \subseteq N$, $P(S) + P(T) = P(S \cup T)$.

A random variable over N is a function $X : N \to \mathbb{R}$. A random variable is non-negative if $X(i) \geq 0$ for every $i \in N$. The following definition introduces the new integral. As will be discussed in Sect. 4.2, the definition is analogous to that of Lebesgue integral.

[3] See Gilboa and Schmeidler (1989) for the case of probability distributions.

Definition 1 Let v be a capacity defined over N. Fix a non-negative random variable X. Define,

$$\int^{\text{cav}} X \mathrm{d}v = \min\{f(X)\}, \tag{1}$$

where the minimum is taken over all concave and homogeneous functions $f : \mathbb{R}^n_+ \to \mathbb{R}$ such that [4] $f(\mathbb{1}_R) \geq v(R)$ for every $R \subseteq N$.

Remark 1 Since the minimum of a family of concave and homogeneous functions over $\mathbb{R}^n_+$ is concave and homogeneous, so is $\int^{cav} X \mathrm{d}v$, as a function of X.

Let v and w be two capacities. We say that $v \geq w$ if $v(S) \geq w(S)$ for every $S \subseteq N$. The following lemma provides an explicit formula for the new integral.

Lemma 1 (i) *For every non-negative X defined over N,*

$$\int^{\text{cav}} X \mathrm{d}v = \max\left\{\sum_{R\subseteq N} \alpha_R v(R);\ \sum_{R\subseteq N} \alpha_R \mathbb{1}_R = X,\ \alpha_R \geq 0\right\}. \tag{2}$$

(ii) $\int^{\text{cav}} X \mathrm{d}v = \min_{P \text{ is additive and } P\geq v} \int X \mathrm{d}P.$

The proof of the lemma is based on the fact that, as a function of X, $\int^{\text{cav}} X \mathrm{d}v$ is concave. This is rather standard and is therefore omitted. Note that in Lemma 1 (ii) the capacities P need not be probability distributions, nor do they satisfy $P(N) = v(N)$.

Zhang et al. (2002) discussed expressions similar to that on the right-hand side of Eq. (2) with a further restriction that all the sets are required to be mutually disjoint. With this restriction the integral becomes analogous to a Riemann integral.

The sum $\sum_{R\subseteq N} \alpha_R \mathbb{1}_R$ is a *decomposition* of X, if $\alpha_R \geq 0$ for every $R \subseteq N$ and $\sum_{R\subseteq N} \alpha_R \mathbb{1}_R = X$. It is an *optimal decomposition* of X w.r.t. v if it is a decomposition of X w.r.t. v and $\int^{\text{cav}} X \mathrm{d}v = \sum_{R\subseteq N} \alpha_R v(R)$. When talking about decompositions, the reference to v will be often dropped.

3 Examples

Example 1 Let $N = \{1, 2, 3\}$, $v(N) = 1$, $v(12) = v(23) = \frac{2}{3}$, $v(13) = \frac{1}{4}$ and $v(i) = 0$ for every $i \in N$. A function over N is a 3-dimensional vector. Consider $X = (1, 2, 1)$. Note that $(1, 1, 0) + (0, 1, 1)$ is a decomposition of X. Furthermore, it is an optimal decomposition of X: $\int^{\text{cav}} X \mathrm{d}v = \frac{2}{3} + \frac{2}{3} = \frac{4}{3}$.

Example 2 (resolving Ellsberg paradox) Suppose that an urn contains 30 red balls and 60 other balls that are either green or blue. A ball is randomly drawn from the urn and a decision maker is given a choice between the two gambles.

[4] $\mathbb{1}_R$ is the indicator of R: $\mathbb{1}_R = (\mathbb{1}^1_R, \ldots, \mathbb{1}^n_R)$, where $\mathbb{1}^i_R$ equals 1 if $i \in R$ and 0, otherwise.

Gamble **X**:	to receive \$100 if a red ball is drawn.
Gamble **Y**:	to receive \$100 if a green ball is drawn.
	In addition, the decision maker is also given the choice between these two gambles:
Gamble **Z**:	to receive \$100 if a red or blue ball is drawn.
Gamble **W**:	to receive \$100 if a green or blue ball is drawn.

It is well documented that most people strongly prefer Gamble **X** to Gamble **Y** and Gamble **W** to Gamble **Z**. This is a violation of the expected utility theory.

There are three states of nature in this scenario: R, G and B, one for each color. Denote by N the set containing these states. Each of the gambles corresponds to a real function (a random variable) defined over N. For instance, Gamble **X** corresponds to the random variable X, defined as $X(R) = 100$ and $X(G) = X(B) = 0$.

The probability of four events are known: $p(\emptyset) = 0$, $p(N) = 1$, $p(\{R\}) = \frac{1}{3}$ and $p(\{G, B\}) = \frac{2}{3}$. The probability p is partially specified: it is defined only on a sub-collection of events and not on all events. Although the new integral has been introduced so far in relation to capacities defined over all events, the same idea may be used for what will be later called fuzzy capacities (among which capacities that may be defined only over a sub collection of events). This is explained here in order to resolve Ellsberg paradox and will be elaborated on later in Sect. 9.

The integral of a function X is defined in a fashion similar to Eq. (1). When p is defined only over familiar events, X is allowed to be written as a positive linear combination of characteristic functions of *familiar* events only. Using only the four familiar events, X is optimally decomposed as $X = 100 \cdot \mathbb{1}_{\{R\}}$. And thus, $\int^{\text{cav}} X \mathrm{d}p = 100 \cdot \frac{1}{3}$. When doing the same for Y (the random variable that corresponds to Gamble **Y**), one cannot obtain a precise decomposition of Y. The maximal non-negative function which is lower than or equal to Y and can be written only in terms of the four familiar events is $0 \cdot \mathbb{1}_N$. The integral of Y is therefore equal to 0. Since, $100 \cdot \frac{1}{3} > 0$, X is preferred to Y.

A similar method applied to Z and W yields: $Z \geq 100 \cdot \mathbb{1}_{\{R\}}$ and the right-hand side is the greatest of its kind. Thus, $\int^{\text{cav}} Z \mathrm{d}p = 100 \cdot \frac{1}{3}$, while W is optimally decomposed as $100 \cdot \mathbb{1}_{\{G,B\}}$. Therefore, $\int^{\text{cav}} W \mathrm{d}p = 100 \cdot \frac{2}{3}$. Since $100 \cdot \frac{1}{3} < 100 \cdot \frac{2}{3}$, Gamble **W** is preferred to Gamble **Z** .

The intuition is that the decision maker bases her evaluation of unknown random variables on known figures: the probabilities of the familiar events. Using simple functions that can be expressed by these events, the decision maker approximates from below any unknown random variable and the maximal simple function of this kind in the one used by the new integral.

4 The new integral and Choquet integral

4.1 The new integral and Choquet integral

Let v be a capacity defined over N. The Choquet integral of non-negative X w.r.t. v, denoted $\int^C X \mathrm{d}v$, is defined by $\sum_{i=1}^{n} (X_{\sigma(i)} - X_{\sigma(i-1)}) v(R_i)$, where σ is a permutation over N that satisfies $X_{\sigma(1)} \leq \cdots \leq X_{\sigma(n)}$ and $R_i = \{\sigma(i), \ldots, \sigma(n)\}$ ($X(\sigma(0)) = 0$, by convention).

Note that,

$$X = \sum \alpha_i \mathbb{1}_{R_i}, \tag{3}$$

where $\alpha_i = X_{\sigma(i)} - X_{\sigma(i-1)}$. Thus, $\sum \alpha_i \mathbb{1}_{R_i}$ is a decomposition of X. This means that a particular decomposition of X is used for the calculation of the Choquet integral. In contrast, the new integral allows all possible decompositions, and as in the definition of the Lebesgue integral (see next section), the one that achieves the maximum of the respective summation is chosen.

This implies, in particular, that always $\int^C X dv \leq \int^{\text{cav}} X dv$. Lovasz (1983) (see also Azrieli and Lehrer 2007b) imply that $\int^C X dv = \int^{\text{cav}} X dv$ for every X if and only if v is convex (i.e., $v(S) + v(T) \leq v(S \cup T) + v(S \cap T)$ for every $S, T \subseteq N$).

The next example demonstrates a case where the new integral results in a more reasonable outcome than the Choquet integral.

Example 3 Let $N = \{1, 2, 3, 4\}$. The capacity v is defined as the minimum of probability distributions as follows. Denote $p_1 = (\frac{1}{8}, \frac{1}{8}, \frac{1}{4}, \frac{1}{2})$, $p_2 = (\frac{1}{2}, \frac{1}{8}, \frac{1}{8}, \frac{1}{4})$, $p_3 = (\frac{1}{8}, \frac{1}{2}, \frac{1}{4}, \frac{1}{8})$, $p_4 = (\frac{1}{8}, \frac{1}{4}, \frac{1}{2}, \frac{1}{8})$ and $p_5 = (\frac{1}{8}, \frac{1}{2}, \frac{1}{8}, \frac{1}{4})$. For every $S \subseteq N$ define $v(S) = \min_{1 \leq i \leq 5} p_i(S)$. Thus, $v(j) = \frac{1}{8}$ for every $j = 1, 2, 3, 4$, $v(12) = v(13) = v(23) = v(14) = \frac{1}{4}$, $v(34) = v(24) = \frac{3}{8}$, $v(S) = \frac{1}{2}$ if $|S| = 3$ and $v(N) = 1$.

Consider $X = (0, 1, 2, 3)$ and $Y = (1, 0, 2, 3)$. X and Y differ in the values of the first two coordinates. While X assigns the value 0 to the first state and 1 to the second, Y assigns the value 1 to the first state and 0 to the second. The Choquet integral of X coincides with that of Y: $\int^C X dv = \int^C Y dv = \frac{1}{2} + \frac{3}{8} + \frac{1}{8} = 1$. On the other hand, the valuations of X and Y by the new integral differ. Since $X = (0, 1, 0, 1) + 2(0, 0, 1, 1)$, $\int^{\text{cav}} X dv = \frac{3}{8} + 2 \cdot \frac{3}{8} = \frac{9}{8}$. Moreover, $\int^{\text{cav}} Y dv = 1$. In particular, $\int^{\text{cav}} X dv > \int^{\text{cav}} Y dv$.

Recall that X and Y differ only on the first two coordinates. State 2 is more likely than state 1 in the sense that for every S that does not contain these states, $v(S \cup \{1\}) \leq v(S \cup \{2\})$, with a strict inequality when $S = \{4\}$. It therefore seems reasonable to evaluate X more than Y, as implied by the new integral and not by the Choquet integral. Technically speaking, the reason why $\int^{\text{cav}} X dv > \int^{\text{cav}} Y dv$ is that $\{2, 4\}$ and $\{1, 4\}$ take part in the optimal decompositions of X and Y, respectively, and $v(2, 4) > v(1, 4)$.

Since both integrals are homogeneous, $\int^{\text{cav}} \frac{10}{9} Y dv = \int^C \frac{10}{9} Y dv = \frac{10}{9}$. A decision maker whose preferences are determined by the Choquet integral, would prefer $\frac{10}{9} Y$ to X, while a decision maker whose preferences are determined by the new integral, would prefer X to $\frac{10}{9} Y$.

The following proposition generalizes this example and provides a new characterization of convex capacities.

Proposition 1 *Let v be a capacity. Then, v is convex if and only if for every nonnegative X and Y, $\int^{\text{cav}} X dv \geq \int^{\text{cav}} Y dv$ whenever $\int^C X dv \geq \int^C Y dv$.*

The proof is postponed to the Appendix.

4.2 The new integral as an extension of Lebesgue integral

The Lebesgue integral of functions over general probability spaces is defined in a fashion similar to that of (2). For the sake of explanation consider functions over the interval [0, 1]. A function f is *simple* if it can be written as $f = \sum_{i=1}^{k} \alpha_i \mathbb{1}_{R_i}$, where R_i is a measurable set in [0, 1] and $\alpha_i \in \mathbb{R}$. For a simple function, the integral of f with respect to a measure μ is defined as $\sum_{i=1}^{k} \alpha_i \int \mathbb{1}_{R_i}\, \mathrm{d}\mu = \sum_{i=1}^{k} \alpha_i \mu(R_i)$. And for a non-negative function f it is defined as

$$\int f \mathrm{d}\mu := \sup \left\{ \int h \, \mathrm{d}\mu;\ h \text{ is simple and } h \leq f \right\}.$$

Lemma 1 (i) implies that the definition of $\int^{\mathrm{cav}} X \mathrm{d}v$ is similar to this definition.

5 Characterization

In this section we characterize the new integral. In what follows $\int X \mathrm{d}v$ should be thought of as a function from pairs (X, v) to the real numbers. The goal is to find a set of plausible properties of such a function that characterizes it uniquely as the new integral.

The first property is a weak version of an axiom, called "Accordance for Additive Measures", that appears in Groes et al. (1998). They required that if v is additive, then $\int X \mathrm{d}v$ is a regular integral. Here, the axiom is restricted to the case where N is a singleton.

Singleton Accordance for Additive Measures (SAAM): If $|N| = 1$, then $\int \mathbb{1}_N \mathrm{d}v = v(N)$.

The following property states that the integral is co-variant with a positive linear re-scaling.

Homogeneity (HO): For any v, X and $\beta \geq 0$, $\int \beta X \mathrm{d}v = \beta \int X \mathrm{d}v$.

The next axiom is the paramount property of the new integral. In order to explain it consider a situation where a bet of one dollar on horse i yields two dollars if horse i wins the race. A vector $\$\frac{1}{2}X(i)$, $i = 1, \ldots, n$, of bettings [i.e., $X(i)$ on horse i] will be referred to as a *bet*. The bet $\frac{1}{2}X$ corresponds to the variable X, which represents the prizes corresponding to all possible horse winnings. Suppose that the gambler's assessments about the likelihood of all possible winnings is given by a capacity (non-additive probability) v. The integral of X attempts to capture the notion of "expected" return from the bet $\frac{1}{2}X$ when the probability considered is v.

Suppose now that X and Y are two bets and $\beta \in (0, 1)$. Then, $\beta\frac{1}{2}X$, $(1-\beta)\frac{1}{2}Y$ and $\beta\frac{1}{2}X + (1-\beta)\frac{1}{2}Y$ are also bets. The decision maker tries to evaluate the bet $\beta\frac{1}{2}X + (1-\beta)\frac{1}{2}Y$. By splitting it into the bets, $\beta\frac{1}{2}X$ and $(1-\beta)\frac{1}{2}Y$ he can ensure an "expected" return of $\int \beta X \mathrm{d}v + \int (1-\beta) Y \mathrm{d}v$. By splitting $\beta\frac{1}{2}X + (1-\beta)\frac{1}{2}Y$ differently the gambler might guarantee even a higher "expected" return. The particular split into $\beta\frac{1}{2}X$ and $(1-\beta)\frac{1}{2}Y$ ensures that the "expected" return from the bet $\beta\frac{1}{2}X + (1-\beta)\frac{1}{2}Y$ [i.e., $\int \beta X + (1-\beta) Y \mathrm{d}v$] is at least $\int \beta X \mathrm{d}v + \int (1-\beta) Y \mathrm{d}v$.

The following concavity axiom captures this idea.

Concavity (CAV): For any v, X , Y and $\beta \in (0,1)$, $\int \beta X + (1-\beta)Y \mathrm{d}v \geq \int \beta X \mathrm{d}v + \int (1-\beta)Y \mathrm{d}v$.

The next axiom refers to two capacities, one of which is additive. It states that P is additive and $P \geq v$, if and only if the integral w.r.t. to P is greater than or equal to that w.r.t. v. It implies that the integral is monotonic with respect to the capacity in the restrictive sense that if P is additive and it is greater than v, then the integral of any non-negative X w.r.t. P is at least as high as the integral of the same X taken w.r.t. v. Furthermore, the axiom requires that if P is not greater than v (meaning that there is S such that $v(S) > P(S)$), then there is a non-negative function whose integral w.r.t. v is greater than that w.r.t. P.

Monotonicity w.r.t. capacity (M): For every additive P, $P \geq v$ if and only if $\int X \mathrm{d}P \geq \int X \mathrm{d}v$ for every non-negative X.

Let S be a subset of N. The *sub-capacity* v_S is a capacity defined over S: $v_S(T) = v(T)$ for every $T \subseteq S$. The next axiom requires that the integral of the indicator of the subset S with respect to v is equal to the integral with respect to v_S, the sub-capacity restricted to S. It suggests that the integral of a function depends on the values that v takes on the subset of N over which the function is not vanishing.

The following axiom equates two integrals: one w.r.t. v over the domain N, and another w.r.t. v_S over a restricted domain, S.

Independence of irrelevant events (IIE): For every S, $\int \mathbb{1}_S \mathrm{d}v = \int \mathbb{1}_S \mathrm{d}v_S$.

Theorem 1 (First Characterization) *The integral* $\int X \mathrm{d}v$ *satisfies* (SAAM), (CAV), (HO), (M), *and* (IIE) *if and only if* $\int X \mathrm{d}v = \int^{\mathrm{cav}} X \mathrm{d}v$ *for every non-negative* X.

The following axiom requires only one of the two implications contained in (M).

Weak monotonicity w.r.t. capacity (WM): For every additive P, if $P \geq v$, then $\int X \mathrm{d}P \geq \int X \mathrm{d}v$ for every non-negative X.

Schmeidler (1986) and Groes et al. (1998) employ the indicator property which is a strong version of the following axiom. They required that $\int \mathbb{1}_S \mathrm{d}v = v(S)$.

Weak Indicator property (WIP): For every S, $\int \mathbb{1}_S \mathrm{d}v \geq v(S)$.

Theorem 2 (Second characterization) *The integral* $\int X \mathrm{d}v$ *satisfies* (SAAM), (CAV), (HO), (WM), *and* (WIP) *if and only if* $\int X \mathrm{d}v = \int^{\mathrm{cav}} X \mathrm{d}v$ *for every non-negative* X.

Remark 2 Properties (SAAM), (HO), (M), (IIE) and (WIP) are also shared by the Choquet integral.

6 The proof of the Theorems

Proof of Theorem 1 The fact that $\int X \mathrm{d}v$ satisfies (SAAM), (CAV), (HO), (M), and (IIE) is easy to check. As for the inverse direction, (M) implies[5] that for every additive

[5] In fact, at this point the 'only if' direction of (M) suffices.

capacity P that satisfies $P \geq v$, $\int X\mathrm{d}P \geq \int X\mathrm{d}v$. Thus, $\min_{P\geq v} \int X\mathrm{d}P \geq \int X\mathrm{d}v$. Lemma 1 (ii) implies that $\int^{\mathrm{cav}} X\mathrm{d}v \geq \int X\mathrm{d}v$. (CAV) and (HO) imply that $\int X\mathrm{d}v$ is concave and homogeneous. As a function of X, $\int^{\mathrm{cav}} X\mathrm{d}v$ is the smallest concave function that is greater than or equal to v. Thus, it remains to show that $\int \mathbb{1}_S\mathrm{d}v \geq v(S)$ for every $S \subseteq N$.

We proceed by induction on the size of S. For S such that $|S| = 1$, (IIE) and (SAAM) imply $\int \mathbb{1}_S\mathrm{d}v_S = \int \mathbb{1}_S\mathrm{d}v = v(S)$. Assume that $\int \mathbb{1}_S\mathrm{d}v \geq v(S)$ for every $S \subseteq N$ with $|S| < \ell$ and we prove it for S of size ℓ.

Fix $S \subseteq N$ with $|S| = \ell$, let Δ be the set of all non-negative variables $X : N \to \mathbb{R}$ such that $X(i) = 0$ for every $i \notin S$ and $\sum_{i\in S} X(i) = 1$ and define $\phi(X) = \int^{\mathrm{cav}} X\mathrm{d}v_S$. Due to (CAV) the function ϕ defined over Δ is concave.

Assume to the contrary that $\phi(\mathbb{1}_S/|S|) = \int \mathbb{1}_S/|S|\mathrm{d}v_S = \int \mathbb{1}_S/|S|\mathrm{d}v < v(S)/|S|$ [the second equality is due to (IIE)]. Then, there is a linear function $g(X)$ defined on Δ that supports the graph of ϕ at $\mathbb{1}_S/|S|$. This function has the form, $g(X) = \langle a, X\rangle + b$, with $\langle\cdot,\cdot\rangle$ being the inner product, $a = (a(i))_{i\in S} \in \mathbb{R}^{|S|}$ and $b \in \mathbb{R}$. The fact that g is a supporting function at $\mathbb{1}_S/|S|$ means that $g(X) \geq \phi(X)$ for every $X \in \Delta$ and $g(\mathbb{1}_S/|S|) = \phi(\mathbb{1}_S/|S|)$.

Define $P(i) = a(i) + b$. Note that since $\sum_{i\in S} X(i) = 1$ for every $X \in \Delta$, $g(X) = \langle P, X\rangle$. By the induction hypothesis, for every $i \in S$, $\phi(\mathbb{1}_{\{i\}}) \geq v(\{i\}) \geq 0$. Since, $P(i) = g(\mathbb{1}_{\{i\}}) \geq \phi(\mathbb{1}_{\{i\}})$ we obtain $P(i) \geq 0$ for every $i \in S$. Thus, one may refer to P as an additive capacity defined over S. This capacity satisfies $g(X) = \int X\mathrm{d}P \geq \phi(X)$ for every non-negative X. By (IIE) and the induction hypothesis, $P(T)/|T| = g(\mathbb{1}_T/|T|) \geq \phi(\mathbb{1}_T/|T|) = \int \mathbb{1}_T/|T|\mathrm{d}v_S = \int \mathbb{1}_T/|T|\mathrm{d}v_T = \int \mathbb{1}_T/|T|\mathrm{d}v \geq v(T)/|T|$ for every non-empty strict subset T of S. As for S itself, $P(S)/|S| = g(\mathbb{1}_S/|S|) = \phi(\mathbb{1}_S/|S|)$, which is by assumption strictly smaller than $v(S)/|S| = v_S(S)/|S|$. Therefore, P is greater than or equal to v_S on every strict subset T of S, while $P(S) < v_S(S)$. In particular, $P \ngeq v_S$.

The axiom[6] (M) ensures that there is X defined over S such that $\int X\mathrm{d}v_S > \int X\mathrm{d}P$. By (HO) $X \neq 0$ and it can be assumed without loss of generality that $X \in \Delta$. This contradicts $\int X\mathrm{d}P = g(X) \geq \phi(X) = \int X\mathrm{d}v_S$ for every $X \in \Delta$. □

Proof of Theorem 2 The first part of the proof of Theorem 1 uses (SAAM), (CAV), (HO), and only (WM). The second part is devoted to showing what (WIP) explicitly assumes. One therefore obtains Theorem 2. □

7 Minimum over the core

The capacity v has a *large core* (Sharkey 1982) if and only if for every $S \subseteq N$ and for every additive capacity Q that satisfies $v \leq Q$, there is P in the core of v such that $P \leq Q$. The capacity v is *exact* (Schmeidler 1972) if and only if for every $S \subseteq N$, there is P in the core[7] of v such that $P(S) = v(S)$. If v is convex, then v has a large core (see Sharkey 1982) and if v has a large core and each of its sub-capacities has a

[6] At this point the "if" part of (M) is being used.

[7] The *core* of v consists of all additive capacities P such that $P \geq v$ and $P(N) = v(N)$.

non-empty core, then it is exact (see, Azrieli and Lehrer 2007b). No two of these three notions are equivalent.

The connection between the largeness of the core and the integral is provided in the following statement.

Proposition 2 (Azrieli and Lehrer 2007b) *v has a large core if and only if*

$$\int^{\mathrm{cav}} X \mathrm{d}v = \min_{P \text{ in the core of } v} \int X \mathrm{d}P \tag{4}$$

for every non-negative X.

Proposition 2 implies that a decision maker that uses a capacity with a large core and $\int^{\mathrm{cav}} X \mathrm{d}v$ to evaluate a random variable X abides to the model of Gilboa and Schmeidler (1989). In this model, preference orders over random variables are represented by a minimum over a compact and convex set of probability distributions. When v has a large core the compact and convex set of probability distributions is the core of v.

Azrieli and Lehrer (2007b) show that

Corollary 1

$$\int^{cav} X + c\mathrm{d}v = \int^{cav} X\mathrm{d}v + \int^{cav} c\mathrm{d}v = \int^{cav} X\mathrm{d}v + c \cdot v(N)$$

for every non-negative X and a constant c if and only if v has a large core.

Remark 3 It is important to note that when the capacity v has a large core Corollary 1 enables one to extend the domain of the integral from the non-negative variables to all variables. Let v be a capacity with a large core and let X be any random variable. Then, there is a constant c such that $X + c$ is non-negative. One may then define

$$\int X\mathrm{d}v = \int X + c\mathrm{d}v - cv(N).$$

Capacities with a large core are important to decisions under uncertainty primarily when the decision maker is partially informed of the true distribution (see Lehrer 2006). In Example 2 the decision maker is informed only of two events: the probability of Red is $\frac{1}{3}$, and the probability of Green or Blue is $\frac{2}{3}$. Thus, the decision maker ought to take a decision having only a partial information about the underlying distribution.

In a more general setting, for instance when the distribution of the balls in Ellsberg's urn is dynamic, the decision maker might be informed only of the expectation of some random variables. Suppose, for instance, that the green balls multiply once a day, and the decision maker should take a decision at the second day. The probability of Red is no longer $\frac{1}{3}$. In fact, at the second day the decision maker knows the probability of no non-trivial event.

A simple calculation shows that the decision maker can deduce that at the second day the expectation of the random variable $(1, \frac{1}{6}, 0)$ (i.e., the one that takes the values $1, \frac{1}{6}, 0$ on Red, Green and Blue, respectively) is $\frac{1}{3}$.

In general, the decision maker might be informed of the expectation of every random variable in a set $\mathcal{Y}$. That is, the decision maker is informed of $\mathbb{E}(Y)$ for every $Y \in \mathcal{Y}$, where $\mathbb{E}(\cdot)$ is the expectation with respect to the real distribution, $\mathbb{P}$. It should be emphasized that the real distribution $\mathbb{P}$ is not fully revealed to the decision maker; the latter is informed only of the expectations of some, but not all, random variables. Note that a random variable $Y \in \mathcal{Y}$ might be an indicator variable, in which case the decision maker is informed of the probability of the corresponding event.

A conservative decision maker would like to use the partial information he obtained in order to get an estimation of the probability of all events. The lower bound of the probability of an event S is then,

$$v(S) = \max\left\{\sum_Y \alpha_Y \mathbb{E}(Y);\ \sum_Y \alpha_Y Y \le \mathbb{1}_S, \alpha_Y \in \mathbb{R} \text{ and } Y \in \mathcal{Y}\right\}.$$

The capacity v has a large core and is typically not convex.

The analogous statement of Proposition 2 for the Choquet integral is due to Schmeidler (1986). He showed that v is convex if and only if

$$\int^C X \mathrm{d}v = \min_{P \text{ in the core of } v} \int X \mathrm{d}P$$

for every non-negative X.

8 First order stochastic dominance and concavity

Let (v, N) be a capacity, and X, X' be two non-negative functions over N. We say that X' (first order) *stochastically dominates* X w.r.t. v, denoted $X' \succeq^v X$, if for every number t, $v(X' \ge t) \ge v(X \ge t)$.

The Choquet integral is monotonic w.r.t. stochastic dominance. That is, if $X' \succeq^v X$, then $\int^C X' \mathrm{d}v \ge \int^C X \mathrm{d}v$.

Example 4 Let $N = \{1, 2, 3\}$, $v(N) = 1$, $v(12) = v(13) = \frac{3}{4}$, $v(23) = 1$ and $v(i) = 0$ for every $i \in N$. Consider $X = (1, 1, 1)$ and $X' = (0, \frac{6}{5}, \frac{6}{5})$. $\int^{\text{cav}} X \mathrm{d}v = \frac{5}{4}$, while $\int^{\text{cav}} X' \mathrm{d}v = \frac{6}{5}$. In this example $X' \succeq^v X$ and nevertheless, $\int^{\text{cav}} X' \mathrm{d}v < \int^{\text{cav}} X \mathrm{d}v$. □

Example 4 shows that $\int^{\text{cav}}$ is not monotonic w.r.t. stochastic dominance. The question arises whether there is a reasonable integral which is monotonic w.r.t. stochastic dominance and concave [i.e., satisfies (CAV)] at the same time. The following example shows that there is no homogeneous (non-trivial) integral which possesses these two properties.

Example 5 Let $N = \{1, 2, 3\}$, $v(S) = 1$ if $|S| \geq 2$ and otherwise, $v(S) = 0$. If $|S| = 2$, then $\mathbb{1}_S \succeq^v \mathbb{1}_N$, and if the integral $\int \cdot \, \mathrm{d}v$ is monotonic w.r.t. stochastic dominance, then $\int \mathbb{1}_S \mathrm{d}v \geq \int \mathbb{1}_N \mathrm{d}v$. However, $\mathbb{1}_N = \sum_{S;\ |S|=2} \frac{1}{2} \mathbb{1}_S$, and if $\int \cdot \, \mathrm{d}v$ is concave and homogeneous, then $\int \mathbb{1}_N \mathrm{d}v \geq \sum_{S;\ |S|=2} \frac{1}{2} \int \mathbb{1}_S \mathrm{d}v \geq \frac{3}{2} \int \mathbb{1}_N \mathrm{d}v$. Therefore, an homogeneous integral cannot be both, monotonic w.r.t. stochastic dominance and concave, unless $\int \mathbb{1}_N \mathrm{d}v \leq 0$.

The set N can be thought of as a state space and the function $\frac{2}{3}\mathbb{1}_N$ can be thought of as a portfolio that ensures a payoff of $\frac{2}{3}$ at any state. However, $\frac{2}{3}\mathbb{1}_N$ can be decomposed as an average of three portfolios: $\frac{2}{3}\mathbb{1}_N = \sum_{S;\ |S|=2} \frac{1}{3}\mathbb{1}_S$. Thus, if each of the portfolios $\mathbb{1}_S$, $|S| = 2$ (i.e., a payoff of 1 is guaranteed if a state in S is realized) is selected with probability $\frac{1}{3}$, then, on average, a payoff of $\frac{2}{3}$ is guaranteed at any state. The idea behind concavity is that the value of $\frac{2}{3}\mathbb{1}_N$ should be at least the average of the values of the portfolios forming it. That is, $\int^{\mathrm{cav}} \frac{2}{3}\mathbb{1}_N \mathrm{d}v \geq \sum_{S;\ |S|=2} \frac{1}{3} \int^{\mathrm{cav}} \mathbb{1}_S \mathrm{d}v$. □

9 An integral w.r.t. a fuzzy capacity

9.1 Fuzzy capacity

Let $I = [0, 1]^n$ be the unit square. For every $a \in I$ let $|a|$ be the sum of its coordinates. Any subset of N can be identified with its indicator, which is an extreme point of I. For every $(x_1, \dots, x_n)$ and $(y_1, \dots, y_n)$ in I we say that $(x_1, \dots, x_n) \geq (y_1, \dots, y_n)$ if $x_i \geq y_i$, $i = 1, \dots, n$. A function f over a subset of I is said to be *monotonic* if for every X, Y in the range of f, $X \geq Y$ implies $f(X) \geq f(Y)$. Thus, a capacity is a monotonic function v defined over the extreme points of I and $v(0, \dots, 0) = 0$. The notion of capacity is extended here as follows:

Definition 2 (1) The pair (v, A) is a *fuzzy capacity* if $(1, \dots, 1) \in A \subseteq I$, $v : A \to \mathbb{R}_+$ is monotonic, continuous, and there is a positive K such that $v(a) \leq K|a|$ for every $a \in A$.

(2) (P, A) is an *additive fuzzy capacity* if there are non-negative constants, $p_1, \dots, p_n$, such that for every $a = (a_1, \dots, a_n) \in A$, $P(a) = \sum_{i=1}^n a_i p_i$.

While a capacity v assigns values (subjective probabilities) to events, a fuzzy capacity assigns values (subjective expected value) to random variables. The database of an agent might enable her to evaluate the expected values of some random variables (e.g., portfolios, bets) and not of others. Furthermore, it might enable her to assess the probability of some but not of all events. The set of variables about which the agent has firm assessments is represented by A. Note that A might contain only points of the form $\mathbb{1}_S$, where $S \subseteq N$. In this case v is a partially-specified non-additive probability: it evaluates only the probability of events, and not necessarily all of them.

The integral aggregates all available information, including individual assessments of the likelihood of events and expected values of variables, into a comprehensive picture. Upon observing the comprehensive picture the agent might re-evaluate the likelihood of events or the expected values she assigns to random variables and change her mind.

Similar to the definition in Sect. 2 we define the integral of a non-negative X w.r.t. a fuzzy capacity (v, A). Let $\mathcal{L}$ be the set of all concave, monotonic and homogeneous functions $f : \mathbb{R}^n_+ \to \mathbb{R}$ such that $f(a) \geq v(a)$ for every $a \in A$. The integral w.r.t. (v, A) is defined as[8]

$$\int^{\text{cav}} X \mathrm{d}v = \min_{f \in \mathcal{L}} f(X),$$

for every non-negative X. The minimum of all concave, monotonic and homogeneous functions is well defined and possesses the same properties. Similarly to Lemma 1 one obtains,

$$\int^{\text{cav}} X \mathrm{d}v = \max \left\{ \sum_{i=1}^{k} \alpha_i v(a_i) \right\}, \tag{5}$$

where the maximum is taken over all $a_i \in A, \alpha_i \geq 0, \ i = 1, \ldots, k$ that satisfy $\sum_{i=1}^k \alpha_i a_i \leq X$. Denote by cone$A$ the convex cone generated by A. That is, cone$A = \{\sum \alpha_i a_i;\ a_i \in A \text{ and } \alpha_i \geq 0\}$. Note that in Eq. (5) $\sum_{i=1}^k \alpha_i a_i$ is allowed to be less than or equal to and not necessarily equal to X as in Lemma 1. Inequality is allowed since cone$A$ might be a strict subset of $\mathbb{R}^n_+$. Note also that if (P, A) is additive, then $\int^{\text{cav}} X \mathrm{d}P = \int X \mathrm{d}P$ (the regular integral of X) for every $X \in$ cone$A$.

Example 6 Let $N = \{1, 2\}$. Thus, $I = [0, 1] \times [0, 1]$. Define the fuzzy capacity (v, A) as follows: $A = \{(1, 1), (\frac{1}{2}, \frac{1}{4})\}$, $v(1, 1) = 1$ and $v(\frac{1}{2}, \frac{1}{4}) = \frac{1}{3}$. Consider $X = (1, \frac{3}{4})$. $X = \frac{1}{2}(1, 1) + (\frac{1}{2}, \frac{1}{4})$ and this is an optimal decomposition of X. Thus, $\int^{\text{cav}} X \mathrm{d}v = \frac{1}{2} \cdot 1 + \frac{1}{3} = \frac{5}{6}$. Now let $Y = (2, 3)$. $Y = (2, 3) \geq 2(1, 1)$ while $2(1, 1)$ attains the maximum of the right-hand side of Eq. (5). Therefore, $\int^{\text{cav}} Y \mathrm{d}v = 2$.

Example 2 revisited Ellsberg's paradox was analyzed in Example 2. In order to phrase this analysis in terms of fuzzy capacities, let $N = \{R, G, B\}$. Thus, $I = [0, 1]^3$, the three dimensional unit cube. Set $A = \{(1, 1, 1), (1, 0, 0), (0, 1, 1)\}$ and $v(1, 1, 1) = 1$, $v(1, 0, 0) = \frac{1}{3}$ and $v(0, 1, 1) = \frac{2}{3}$. Let X be $(100, 0, 0)$. Since $100(1, 0, 0)$ is an optimal decomposition of X, $\int^{\text{cav}} X \mathrm{d}v = 100 \cdot \frac{1}{3}$. Also define, $Y = (0, 100, 0)$ the right-hand side of Eq. (5) is attained by $0(1, 1, 1)$, and therefore, $\int^{\text{cav}} Y \mathrm{d}v = 0$.

The *core* of (v, A) (see also[9] Aubin 1979; Azrieli and Lehrer 2007a) consists of all the additive fuzzy capacities P such that $P(1, \ldots, 1) = v(1, \ldots, 1)$ and for every $a \in A$, $P(a) \geq v(a)$. The fuzzy capacity (v, A) is *exact* if for every $a \in A$ there is P in the core of v such that $P(a) = v(a)$.

[8] The set A is dropped from the notation.

[9] Both referred to the special case where $A = I$.

9.2 Minimum over additive capacities and the integral

Let $\mathcal{P}$ be a compact set of additive capacities defined over the extreme points of I. Denote the fuzzy capacity $(v_{\mathcal{P}}, I)$ as follows:

$$v_{\mathcal{P}}(a) = \min_{P\in\mathcal{P}} \int a\mathrm{d}P \quad \text{for} \quad \text{every } a \in I. \tag{6}$$

Remark 4 For any compact set of additive capacities, $\mathcal{P}$, denote by conv$\mathcal{P}$ the convex hull of $\mathcal{P}$. For any $a \in A$, the value $v_{\mathrm{conv}\mathcal{P}}(a)$ is attained at an extreme point of conv$\mathcal{P}$, which is in $\mathcal{P}$. Therefore, $v_{\mathcal{P}} = v_{\mathrm{conv}\mathcal{P}}$.

The following example illustrates the main idea demonstrated in this section.

Example 7 Let $N = \{1, 2, 3\}$ and consider the set $\mathcal{P}$ which consists of the probability distributions $P_1 = (\frac{1}{2}, \frac{1}{4}, \frac{1}{4})$, $P_2 = (\frac{1}{4}, \frac{1}{2}, \frac{1}{4})$ and $P_3 = (\frac{1}{4}, \frac{1}{4}, \frac{1}{2})$. Denote by w the capacity $v_{\mathcal{P}}$ restricted to $A = \{\mathbb{1}_S;\ S \subseteq N\}$. Thus,[10] $w(N) = 1$ and $w(S) = |S|\frac{1}{4}$ for $|S| \le 2$. In this case for every non-negative X, $\min_{P\in\mathcal{P}} \mathbb{E}_P(X) = \int^{\mathrm{cav}} X\mathrm{d}w$.

Now consider $P_4 = (\frac{2}{16}, \frac{7}{16}, \frac{7}{16})$ and $\mathcal{P}' = \{P_1, P_2, P_3, P_4\}$. Denote by u the capacity $v_{\mathcal{P}'}$ restricted to A. Thus, $u(N) = 1$, $u(S) = \frac{1}{2}$ if $|S| = 2$, $u(1) = \frac{1}{8}$, and $u(2) = u(3) = \frac{1}{4}$. In order to show that $\min_{P\in\mathcal{P}'} \mathbb{E}_P(X) \neq \int^{\mathrm{cav}} X\mathrm{d}u$ for some non-negative X, consider $X = (\frac{3}{5}, \frac{2}{5}, 0)$. On one hand, $\min_{P\in\mathcal{P}'} \mathbb{E}_P(X) = \frac{1}{4}$, and on the other, $\int^{\mathrm{cav}} X\mathrm{d}u = \frac{1}{5}u(1,0,0) + \frac{2}{5}u(1,1,0) = \frac{1}{5}\frac{1}{8} + \frac{2}{5}\frac{1}{2} = \frac{9}{40} < \frac{1}{4}$. In other words, in order to get equality between $\int^{\mathrm{cav}} X\mathrm{d}v_{\mathcal{P}'}$ and $\min_{P\in\mathcal{P}} \mathbb{E}_P(X)$, one cannot restrict oneself to A.

We enlarge A: let $A' = A\cup\{(\frac{3}{5}, \frac{2}{5}, 0), (\frac{3}{5}, 0, \frac{2}{5})\}$. Define the fuzzy capacity (w', A') as follows: it coincides with u on A, and $w'(\frac{3}{5}, \frac{2}{5}, 0) = w'(\frac{3}{5}, 0, \frac{2}{5}) = \frac{1}{4}$. For every non-negative X we obtain, $\min_{P\in\mathcal{P}'} \mathbb{E}_P(X) = \int^{\mathrm{cav}} X\mathrm{d}w'$. For instance, let $X = (\frac{3}{5}, \frac{1}{5}, \frac{1}{5})$. $\min_{P\in\mathcal{P}'} \mathbb{E}_P(X) = \mathbb{E}_{P_4}(X) = \frac{2}{16}\frac{3}{5} + \frac{7}{16}\frac{1}{5} + \frac{7}{16}\frac{1}{5} = \frac{1}{4}$ and $\int^{\mathrm{cav}} X\mathrm{d}w' = \frac{1}{2}w'(\frac{3}{5}, \frac{2}{5}, 0) + \frac{1}{2}w'(\frac{3}{5}, 0, \frac{2}{5}) = \frac{1}{4}$.

The information embedded in $\mathcal{P}'$ cannot be compressed into a capacity defined only over the extreme points of I (i.e., to subsets on N). The values of w' over the points $(\frac{3}{5}, \frac{2}{5}, 0)$ and $(\frac{3}{5}, 0, \frac{2}{5})$ are necessary. On the other hand, the values of w' on A' are sufficient to provide all the information needed to obtain $\min_{P\in\mathcal{P}} \mathbb{E}_P(X)$ through the integral.

The following lemma (stated without a proof) connects between the minimum over a set of capacities and exactness.

Lemma 2 *If for any $P, P' \in \mathcal{P}$, $P(1,\ldots,1) = P'(1,\ldots,1)$, then $v_{\mathcal{P}}$ is exact.*

Recall that in Gilboa and Schmeidler (1989) the minimum is taken over a compact and convex set of probability distributions. It turns out that in the current context the representation as a minimum over additive capacities (not necessarily probability

[10] In this example we identify a subset of N with its indicator.

distributions) and the representation as an integral w.r.t. a fuzzy capacity are equivalent. Formally,

Proposition 3 (1) *Let $\mathcal{P}$ be a compact set of additive capacities. Then,*

$$\int^{cav} X \mathrm{d}v_{\mathcal{P}} = \min_{P\in\mathcal{P}} \int X \mathrm{d}P$$

for every non-negative X. Furthermore, if $\mathcal{P}$ is either finite or a polygon, then there is a fuzzy capacity (v, A) with A being finite such that $\min_{P\in\mathcal{P}} \int X\mathrm{d}P = \int^{cav} X\mathrm{d}v$.

(2) *For every fuzzy capacity (v, A), if (v, A) is Lipschitz (i.e., there is a constant $L > 0$ such that for every $a, a' \in A$, $|v(a) - v(a')| \le L\|a - a'\|_2$), then there is a compact and convex set of additive capacities (not necessarily probability distributions), $\mathcal{P}$, such that*

$$\int^{cav} X \mathrm{d}v = \min_{P\in\mathcal{P}} \int X \mathrm{d}P.$$

Moreover, if (v, A) is exact, then $P(1, \ldots, 1) = P'(1, \ldots, 1)$ for every $P, P' \in \mathcal{P}$.

The proof[11] is rather standard and is therefore omitted.

The following example shows that in Proposition 3(2) the Lipschitz condition is necessary.

Example 8 Let $I = [0, 1]^2$ and $v(x, y) = \sqrt{xy}$. The fuzzy capacity (v, I) is concave. However, at the boundary point $(0, 1)$ there is no supporting hyper-plane to the graph of v. Therefore, there is no (non-trivial) additive capacity P such that $P(0, 1) = v(0, 1) = 0$ and at the same time $P(x, y) \ge v(x, y) = \sqrt{xy}$ for every $(x, y) \in I$. However, for every $(a, b) \in I$ with $a, b > 0$ let $P_{(a,b)} = \frac{1}{2}(\sqrt{\frac{b}{a}}, \sqrt{\frac{a}{b}})$. On one hand, $P_{(a,b)}(a, b) = \sqrt{ab} = v(a, b)$ and on the other, $P_{(a,b)}(x, y) = \frac{1}{2}\sqrt{\frac{b}{a}}x + \frac{1}{2}\sqrt{\frac{a}{b}}y \ge \sqrt{xy} = v(x, y)$ for every $(x, y) \in I$. In other words, $P_{(a,b)}$ corresponds to a supporting hyper-plane of the graph of v at the point (a, b). Finally notice that

$$\begin{aligned}\int^{\text{cav}} (0, 1)\mathrm{d}v &= v(0, 1) = 0 \\ &= \inf_{\substack{a,b>0 \text{ and} \\ (a,b)\to(0,1)}} \sqrt{\frac{a}{b}} = \inf_{\substack{a,b>0 \text{ and} \\ (a,b)\to(0,1)}} P_{(a,b)}(0, 1).\end{aligned}$$

In this case the infimum cannot be replaced by a minimum. The reason is that v does not satisfy the Lipschitz condition stated in Proposition 3(2).

[11] It is based on the fact that any concave function over a compact and convex set D, that can be extended as a concave function to an open set that contains D, is the minimum of all its supporting linear functions.

10 Final comments

10.1 Tail-separability

The Choquet expected utility model does not satisfy the sure-thing principle but it retains a reminiscence of it, called tail-separability (see Machina 2007). The latter means that the preference order between two acts that coincide on the lowest (or highest) reward remains unchanged if the size of this reward changes while staying the lowest (or highest). In other words, if two acts coincide on a tail event (where the reward is either the lowest or the highest) and they change over this tail event (while staying such), then the preference order between them does not change.

Machina (2007) introduces a variation of Ellsberg's urn which poses considerable difficulty for Choquet expected utility model. This difficulty arises due to tail-separability. Expected utility model based on the concave integral presented here does not satisfy tail-separability and may resolve the difficulty demonstrated by Machina's example (see Lehrer 2007).

10.2 Relations with risk measurement

The quest for concave or convex and homogenous functionals has been a theme of extensive research in the last years. A functional ρ is a *coherent risk measure* if it is sub-additive, homogenous of degree 1 [e.g., satisfies (HO)], monotonic (as in 11.1.3 below) and satisfies translation invariance, that is

$$\rho(X + c) = \rho(X) - c \tag{7}$$

for every constant c. Coherent risk measures defined over bounded random variables have been introduced by Artzner et al. (1998) and axiomatized by Delbaen (2002). They can be typically represented by the maximum of minus the expectations with respect to priors in a set $\mathcal{P}$. That is,

$$\rho(X) = \max_{P \in \mathcal{P}} \mathbb{E}_P(-X), \tag{8}$$

where $\mathcal{P}$ is a set of priors.

Let P be an additive probability distribution. A coherent risk measure is P-law invariant if two random variables that have the same cumulative distribution functions w.r.t. P share the same risk measure. Kusuoka (2001) characterized the risk measures that are P-law invariant when P is non-atomic. It turns out that these measures are also monotonic with second order stochastic dominance. Leitner (2005) showed that Kusuoka's representation actually characterizes all coherent risk measures that are monotonic with second-order stochastic dominance.

There are a few similarities and differences between the existing analysis of risk measures and the current discussion on the concave integral. Risk measures analysis requires the functional to be convex and translation invariant (Eq. (7)) means that the risk measure of an asset translated by a constant is the risk measure of the asset *minus* the constant. Here, the integral is concave and translation invariance means that the integral of an asset translated by a constant is the integral of the asset *plus* that constant (as in Corollary 1). These differences can be easily reconciled.

Instead of considering the integral itself, one should consider minus the integral, which becomes a convex function. Corollary 1 implies that v, with $v(N) = 1$, has a large core if and only if $-\int^{\text{cav}} X\mathrm{d}v$ is a coherent risk measure.

As for fuzzy capacities, Eqs. (6) and (8), combined with Lemma 2 and Gilboa and Schmeidler (1989) imply that $\rho(X)$ is a coherent risk measure if and only if $\rho(X) = -\int^{\text{cav}} X\mathrm{d}v$ with v being an exact fuzzy capacity and $v(1,\ldots,1) = 1$. It implies that when v with $v(1,\ldots,1) = 1$ is not exact $-\int^{\text{cav}} X\mathrm{d}v$ is not a coherent risk measure. This is so because it does not satisfy translation invariance.

The studies of Kusuoka (2001) and Leitner (2005) assume an underlying additive probability distribution, while the underlying capacity here is typically non-additive. Finally, Sect. 8 deals with *first* order stochastic dominance while Kusuoka's representation respects *second* order stochastic dominance.

10.3 Extension to general spaces

In this paper N is assumed to be finite. However, the integral can be generalized, using precisely the same Eq. (1), to any space. This is done in Lehrer and Teper (2007).

11 Appendix

11.1 Properties

Properties of the new integral that are not mentioned explicitly in the axioms are listed in this section. Proofs will be provided only to the non-obvious properties. In what follows X and X' are non-negative functions over N, or equivalently, points in $\mathbb{R}^n_+$.

11.1.1 Continuity

$\int^{\text{cav}} X\mathrm{d}v$ is continuous in both, X and v.

11.1.2 Monotonicity w.r.t. capacities

If $v \geq v'$, then $\int^{\text{cav}} X\mathrm{d}v \geq \int^{\text{cav}} X\mathrm{d}v'$ for every non-negative X. Note that this property is not implied by axiom (M) that refers only the case where the greater capacity is additive.

11.1.3 Monotonicity w.r.t. functions

If $X \geq X'$, then $\int^{\text{cav}} X\mathrm{d}v \geq \int^{\text{cav}} X'\mathrm{d}v$.

11.1.4 Characteristic functions

By definition, for every $S \subseteq N$, $\int^{\text{cav}} \mathbb{1}_S \mathrm{d}v \geq v(S)$. If $\int^{\text{cav}} \mathbb{1}_S \mathrm{d}v > v(S)$, then there are scalars $\alpha_i > 0$ and R_i which are proper subsets of N, $i = 1, \ldots, k$, such that $\int^{\text{cav}} \mathbb{1}_S \mathrm{d}v = \sum_{i=1}^k \alpha_i v(R_i)$ and $\int^{\text{cav}} \mathbb{1}_{R_i} \mathrm{d}v = v(R_i),\ i = 1, \ldots, k$.

11.1.5 Totally balanced capacity

Bondareva–Shapley theorem (see Bondareva 1962; Shapley 1967) implies that for any $R \subseteq N$, the core of the sub-capacity v_R is not empty if and only if $\int^{\text{cav}} \mathbb{1}_R \mathrm{d}v = v(R)$. Thus, $\int^{\text{cav}} \mathbb{1}_R \mathrm{d}v = v(R)$ for every $R \subseteq N$ if and only if the capacity is totally balanced (i.e., the core of each of its sub-capacities is not empty).

11.1.6 The integral and the totally balanced cover

Let $S \subseteq N$. Define the capacity v^S as follows: $v^S(R) = v(R)$ if $R \neq S$ and $v^S(S) = \int^{\text{cav}} \mathbb{1}_S \mathrm{d}v$. Then, $\int^{\text{cav}} X \mathrm{d}v = \int^{\text{cav}} X \mathrm{d}v^S$. Thus, increasing the value of the capacity from $v(S)$ to $\int^{\text{cav}} \mathbb{1}_S \mathrm{d}v$ would not change the integral.

Let v be a capacity. Define the capacity B_v as follows: $B_v(S) = \int^{\text{cav}} \mathbb{1}_S \mathrm{d}v$ for every $S \subseteq N$. The capacity B_v is the *totally balanced cover* of v. Then, $\int^{\text{cav}} X \mathrm{d}v = \int^{\text{cav}} X \mathrm{d}B_v$ for every non-negative X.

11.1.7 The integral and the maximum of a function

It might be that $\int^{\text{cav}} X \mathrm{d}v > \max(X)$. However, X can be expressed as a positive linear combination of (characteristic) functions whose integral is between their minimum and their maximum. Furthermore,

Lemma 3 (i) $\int^{cav} X \mathrm{d}v \leq \max(X)$ *for every non-negative* X *if and only if* $B_v(N) \leq 1$.
(ii) *If* $v(N) = 1$, *then* $\int^{\text{cav}} X \mathrm{d}v \leq \max(X)$ *for every non-negative* X *if and only if the core of* v *is non-empty.*

The proof is deferred to the second part of the Appendix.

11.1.8 The integral and the minimum of a function

As stated in Sect. 4, the new integral is always greater than or equal to the Choquet integral. When $v(N) = 1$, $\int^C X \mathrm{d}v \geq \min(X)$, and therefore $\int^{\text{cav}} X \mathrm{d}v \geq \min(X)$.

11.1.9 Piecewise linearity

$\int^{\text{cav}} X \mathrm{d}v$ is piecewise linear in X. That is, the set $\mathbb{R}^n_+$ can be divided into finitely many closed cones $F_1, \ldots, F_\ell$ such that $\int^{\text{cav}} X \mathrm{d}v$ is linear in each one: for every $X, X' \in F_i$, $\int^{\text{cav}} X + X' \mathrm{d}v = \int^{\text{cav}} X \mathrm{d}v + \int^{\text{cav}} X' \mathrm{d}v$.

11.1.10 Local additivity

The previous property implies that $\int^{\text{cav}} X\text{d}v$ is locally additive. That is, every X is included in an open cone, say U_X, such that for every $X' \in U_X$, $\int^{\text{cav}} X + X'\text{d}v = \int^{\text{cav}} X\text{d}v + \int^{\text{cav}} X'\text{d}v$. (It is not true that for every $X', X'' \in U_X$, $\int^{\text{cav}} X' + X''\text{d}v = \int^{\text{cav}} X'\text{d}v + \int^{\text{cav}} X''\text{d}v$.)

11.1.11 Minimum over a set of capacities

Let $\mathcal{C}$ be a set of capacities. Denote $m(\mathcal{C})(S) = \inf_{v\in\mathcal{C}} v(S)$ for every $S \subseteq N$. It turns out that for every $\mathcal{C}$, $\int^{\text{cav}} X\text{d}m(\mathcal{C}) \le \min_{v\in\mathcal{C}} \int^{\text{cav}} X\text{d}v$. However, if $\mathcal{C}$ is the set of all additive capacities that are greater than or equal to v, then $\int^{\text{cav}} X\text{d}m(\mathcal{C}) = \min_{v\in\mathcal{C}} \int^{\text{cav}} X\text{d}v$.

11.2 Proofs

Proposition 1 Let v be a capacity. Then, v is convex if and only if for every non-negative X and Y, $\int^{\text{cav}} X\text{d}v \ge \int^{\text{cav}} Y\text{d}v$ whenever $\int^{C} X\text{d}v \ge \int^{C} Y\text{d}v$.

Proof If v is convex, then $\int^{\text{cav}} X\text{d}v = \int^{C} X\text{d}v$ for every non-negative X. Conversely, if v is not convex, then in particular v is not identically 0. Moreover, by Lovasz (1983, Proposition 4.1, p. 249) there is a non-negative X such that $\int^{\text{cav}} X\text{d}v \neq \int^{C} X\text{d}v$. Since $\int^{\text{cav}} X\text{d}v \ge \int^{C} X\text{d}v$, $\int^{\text{cav}} X\text{d}v > \int^{C} X\text{d}v$. By the definition of the new integral, there is $S \subseteq N$ such that $\int^{\text{cav}} \mathbb{1}_S\text{d}v = v(S) > 0$. There is a constant $c > 0$ such that $\int^{\text{cav}} X\text{d}v > \int^{\text{cav}} c\mathbb{1}_S\text{d}v > \int^{C} X\text{d}v$. Since $\int^{\text{cav}} c\mathbb{1}_S\text{d}v = \int^{C} c\mathbb{1}_S\text{d}v$, we obtain, $\int^{\text{cav}} X\text{d}v > \int^{\text{cav}} c\mathbb{1}_S\text{d}v$ and $\int^{C} c\mathbb{1}_S\text{d}v > \int^{C} X\text{d}v$, as desired. □

Lemma 3 (i) $\int^{\text{cav}} X\text{d}v \le \max(X)$ for every non-negative X if and only if $B_v(N) \le 1$.
(ii) If $v(N) = 1$, then $\int^{\text{cav}} X\text{d}v \le \max(X)$ for every non-negative X if and only if the core of v is non-empty.

Proof (i) Suppose first that $B_v(N) \le 1$ and suppose to the contrary that there is a non-negative X such that $\int^{\text{cav}} X\text{d}v > \max(X)$. Since the integral is homogeneous, it can be assumed without loss of generality that $\max(X) = 1$. In particular, $\mathbb{1}_N \ge X$. By monotonicity w.r.t. functions, $\int^{\text{cav}} \mathbb{1}_N\text{d}v \ge \int^{\text{cav}} X\text{d}v$ and therefore, $\int^{\text{cav}} \mathbb{1}_N\text{d}v > 1$. However, $\int^{\text{cav}} \mathbb{1}_N\text{d}v = \int^{\text{cav}} \mathbb{1}_N\text{d}B_v = B_v(N)$. Thus, $B_v(N) > 1$, which contradicts the assumption.
Conversely, suppose that $\int^{\text{cav}} X\text{d}v \le \max(X)$ for every non-negative X. It implies in particular that $\int^{\text{cav}} \mathbb{1}_N\text{d}v \le 1$. However, $\int^{\text{cav}} \mathbb{1}_N\text{d}v = B_v(N)$, which implies that $B_v(N) \le 1$.
(ii) When $v(N) = 1$, $B_v(N) \le 1$ means that $B_v(N) = v(N)$, which by Bondareva–Shapley theorem (see Bondareva, 1962; Shapley 1967) is equivalent to the non-emptiness of the core.

□

References

Artzner, P., Delbaen, F., Eber, J.-M., Heath, D.: Coherent measures of risk. Financ Stoch **9**, 203–228 (1998)
Aubin, J.P.: Mathematical Methods of Game and Economic Theory. Amsterdam: North-Holland (1979)
Azrieli, Y., Lehrer, E.: Market games in large economies with a finite number of types. Econ Theory **31**, 327–342 (2007a)
Azrieli, Y., Lehrer, E.: Extendable cooperative games. J Public Econ Theory **9**, 1069–1078 (2007b)
Bondareva, O.: The theory of the core in an n-person game (in Russian). Vestn Leningr Univ **13**, 141–142 (1962)
Camerer, C.F., Weber, M.: Recent developments in modelling preferences: uncertainty and ambiguity. J Risk Uncertain **5**, 325–370 (1992)
Delbaen, F.: Coherent risk measures on general probability spaces. In: Sandmann, K. et al. (eds.) Advances in Finance and Stochastics. Essays in Honour of Dieter Sondermann, pp. 1–37. Heidelberg: Springer (2002)
Dow, J., Werlang, S.: Uncertainty aversion, risk aversion, and the optimal choice of portfolio. Econometrica **60**, 197–204 (1992)
Dow, J., Werlang, S.: Nash equilibrium under Knightian uncertainty. J Econ Theory **64**, 305–324 (1994)
Gilboa, I.: Expected utility with purely subjective non-additive probabilities. J Math Econ **16**, 65–88 (1987)
Gilboa, I., Schmeidler, D.: Maxmin expected utility with non-unique prior. J Math Econ **18**, 141–153 (1989)
Groes, E., Jacobsen, H.J., Sloth, B., Tranæs, T.: Axiomatic characterizations of the Choquet integral. Econ Theory **12**, 441–448 (1998)
Kusuoka, S.: On law-invariant coherent risk measures. Adv Math Econ **3**, 83–95 (2001)
Lehrer, E.: Partially specified probabilities: decisions and games. http://www.math.tau.ac.il/~lehrer/Papers/Partially-web.pdf (2006)
Lehrer, E.: A comment on an example by Machina. http://www.math.tau.ac.il/~lehrer/Papers/comment-web.pdf (2007)
Lehrer, E., Teper, R.: The concave integral over large spaces. http://www.math.tau.ac.il/~lehrer/Papers/integral-large-spaces-web.pdf (2007)
Leitner, J.: A short note on second-order stochastic dominance preserving coherent risk measures. Math Financ **15**, 649–651 (2005)
Lovasz, L.: Submodular functions and convexity. In: Bachem, A., et al. (eds.) Mathematical Programming: The State of the Art, pp. 235–257. Heidelberg: Springer (1983)
Machina, M.: Risk, ambiguity and the rank-dependence axioms, mimeo (2007)
Murofushi, T., Sugeno, M.: Fuzzy t-conorm integral with respect to fuzzy measures: generalization of Sugeno integral and Choquet integral. Fuzzy Sets Syst **42**, 57–71 (1991)
Sarin, R., Wakker, P.P.: A simple axiomatization of non-additive expected utility. Econometrica **60**, 1255–1272 (1992)
Schmeidler, D.: Cores of exact games I. J Math Anal Appl **40**, 214–225 (1972)
Schmeidler, D.: Integral representation without additivity. Proc Am Math Soc **97**, 255–261 (1986)
Schmeidler, D.: Subjective probabilities without additivity. Econometrica **57**, 571–587 (1989)
Shapley, L.S.: On balanced sets and cores. Nav Res Logist Q **14**, 453–460 (1967)
Sharkey, W.W.: Cooperative games with large cores. Int J Game Theory **11**, 175–182 (1982)
Starmer, C.: Developments in non-expected utility theory: the hunt for a descriptive theory of choice under risk. J Econ Lit **38**, 332–382 (2000)
Sugeno, M.: Theory of fuzzy integrals and its applications, PhD Dissertation, Tokyo Institute of Technology, Tokyo, Japan (1974)
Wakker, P.P.: Continuous subjective expected utility with nonadditive probabilities. J Math Econ **18**, 1–27 (1989)
Weber, R.J.: Games in coalitional form. In: Aumann, R.J., Hart, S. (eds.) Handbook of Game Theory, vol. II. Amsterdam: Elsevier (1994)
Zhang, G., Yonghong, W., Jie, L.: Lattice-valued Zp-pan-integrals I: For lattice-valued simple functions on lattice. J Fuzzy Math **10**, 213–226 (2002)