ELSEVIER Journal of Mathematical Economics 30 (1998) 405–420

JOURNAL OF Mathematical ECONOMICS

Additivity with multiple priors

Paolo Ghirardato [a,*], Peter Klibanoff [b,1], Massimo Marinacci [c,2]

[a] Division of the Humanities and Social Sciences, California Institute of Technology, Pasadena, CA 91125, USA

[b] Department of Managerial Economics and Decision Sciences, J.L. Kellogg Graduate School of Management, Northwestern University, Evanston, IL 60208, USA

[c] Department of Economics, University of Toronto, Toronto, Canada, M5S 3G7

Received 31 July 1996; revised 31 December 1996; accepted 31 July 1997

Abstract

The functional defined as the 'min' of integrals with respect to probabilities in a given non-empty closed and convex class appears prominently in recent work on uncertainty in economics. In general, such a functional violates the additivity of the expectations operator. We characterize the types of functions over which additivity of this functional is preserved. This happens exactly when 'integrating' functions which are positive affine transformations of each other (or when one is constant). We show that this result is quite general by restricting the types of classes of probabilities considered. Finally, we prove that with a very peculiar exception, all the results hold more generally for functionals which are linear combinations of the 'min' and the 'max' functional. © 1998 Elsevier Science S.A. All rights reserved.

JEL classification: D81

Keywords: Additivity; Expected utility functional; Priors; Probability measures

* Corresponding author. E-mail: paolo@hss.caltech.edu.

1 E-mail: peterk@nwu.edu.

2 E-mail: massimo@chass.utoronto.ca.

PII: S0304-4068(97)00047-5

1. Introduction and motivation

The additivity of the expected utility functional is one of the key properties that has made it attractive for use in economic applications. Recent attempts to improve the decision-theoretic basis of economics have relaxed this additivity in order to capture a broader range of economically relevant behavior. In particular, in the wake of the seminal experiments of Ellsberg (1961), decision theorists have produced a number of axiomatic models of behavior aimed at explaining his results and, more generally, capturing ambiguous beliefs. Schmeidler (1989) suggested extending the classical Subjective Expected Utility (SEU) model of Anscombe and Aumann (1963) and Savage (1954) by allowing the preferences of the decision maker (DM) to be represented by Choquet integrals with respect to beliefs which are not necessarily additive, technically known as *capacities*. This is what is known as the Choquet Expected Utility (CEU) model. A related model, the so-called 'multiple priors' model extends SEU by representing the DM's beliefs by a *set* of probability measures, and models her as choosing the act which maximizes the *minimal* expected utility with respect to beliefs in this set. This is the case in the papers of Gilboa and Schmeidler (1988) and Chateauneuf (1991). [3]

Each of these models results in a relaxing, to some extent, of the additivity of SEU. It is precisely this relaxation that yields behavior reflecting attitudes towards ambiguity, as observed by Ellsberg (1961) and the large literature that followed. To better understand and apply these newer theories, it is important to know exactly when additivity and its resulting behavioral implications are maintained and when they are not. Furthermore, as many standard mathematical results rely on additivity, their use is delimited by the extent to which additivity holds in these models. While the additivity of the Choquet integral has been carefully studied (it is additive when integrating the sum of comonotonic functions, defined in Section 2), that of the functional defined as the 'min' of integrals over a (closed and convex) set of probabilities (as obtained in multiple priors model) has not. This paper addresses exactly that issue. More specifically, we ask the following questions: Suppose that we are modelling a DM who can be described by the multiple priors model. Hence, her preferences will be described by a closed and convex set C of probabilities over a state space Ω, and a utility function, u, over outcomes. Assume that we know her preferences under risk (i.e., u), however we do not know anything more about her, in particular, we do not know anything more about the set of probabilities, C. Given two acts f and g, [4] we want to know

[3] As is well-known (see Schmeidler, 1989) the two models we have described have a non-empty intersection, which corresponds to the case in which the set of probabilities representing the DM's beliefs is the core of a supermodular (or convex) capacity. However, neither model is nested in the other as there are DM's whose preferences obey the multiple priors model but not the CEU model and vice versa.

[4] From here on, we fix the utility function u and write simply f and g for $u \circ f$ and $u \circ g$.

whether, for all C, the 'min' of the expected utilities on C is additive for $f+g$. We will show (Theorem 1) that this will be true exactly when f and g are positive affine transformations of each other or at least one is constant (as we say, they are 'affinely-related'). Technically, the 'min' functional is additive for every C when and only when we are integrating the sum of two affinely-related functions. Since affine-relatedness is a much stronger condition than comonotonicity, this implies that the 'min' functional will be additive for a much smaller class of functions than the Choquet integral. For this reason, we inquire whether it is possible to obtain broader additivity by imposing symmetric structural restrictions on the C's we want to consider. The surprising answer is that, unless we want to consider *only* DM's whose C is the convex hull of *degenerate* probabilities (i.e., those that assign probability one to a specific state ω), we again have that additivity holds only for affinely-related functions. This tighter version of the additivity result (Theorem 2) shows that the narrow additivity of the 'min' functional cannot be avoided in general. This phenomenon is not limited to the 'min' functional only: in Section 5 we show that all results generalize, with a very peculiar exception, to all functionals which can be expressed as a linear function of the 'min' and the 'max' functionals (the latter associates with a function the *largest* expected value with respect to probabilities in C).

Our exploration of the additivity of the 'min' functional is related to recent work by Klibanoff (1996) that characterizes when a DM described by the CEU or multiple priors models can be made better off by mixing. [5] A necessary property for this (and, more technically, for the DM's preferences to display strict concavity) is that the two acts being mixed are not affinely related (multiple priors) or comonotonic (CEU). It is clear that for both the Choquet and the 'min' functionals, *strict* concavity can arise only when they fail to be additive. For instance, as the Choquet integral is additive when integrating the sum of comonotonic functions, 'mixing' between comonotonic acts does *not* make a CEU DM better off.

The characterization of the additivity of the 'min' functional also serves to delimit the extension of mathematical results which work for the additive integral. For instance, as Ghirardato (1997) uses comonotonic additivity of the Choquet integral to study the validity of Fubini's theorem for Choquet integrals, the results in this paper can be used to characterize the class of functions for which one can safely exchange the order of integration in a multiple priors model. This is quite helpful, for example, in any application where the notion of independent sources of uncertainty is relevant.

[5] A common feature of the two models is that they were originally developed in a set-up by Anscombe and Aumann (1963), in which the existence of an independent randomizing device (a 'roulette wheel') is assumed. This gives an interpretation to 'objective' mixtures (technically: pointwise convex combinations) of acts. See Eichberger and Kelsey (1995) for an argument that CEU with convex capacities need not allow preference for randomization in a Savage framework.

The structure of the rest of the paper is as follows: After introducing some preliminary material in Section 2, we present the basic additivity result in Section 3. In Section 4, we discuss the tightness of Theorem 1 and present the more general Theorem 2, while underlining the connections with comonotonicity. Section 5 closes by discussing the extension of our results to more general functionals and the peculiar exception to the extension.

2. Preliminaries

Let $\mathscr{F}$ be an algebra of subsets of a space Ω which contains all singletons, $\mathscr{P}$ be the set of all finitely additive probabilities defined on $\mathscr{F}$ and $\mathscr{P}^{s}$ be the subset of $\mathscr{P}$ containing all the simple probabilities, i.e., $P \in \mathscr{P}^{s}$ iff there exists a finite set $A \subseteq \Omega$ such that $P(A) = 1$. Notice that each finite set $A \subseteq \Omega$ belongs to $\mathscr{F}$ because this algebra contains all singletons. δ_ω denotes the probability measure concentrated on $\omega \in \Omega$. Let $\mathscr{C}$ be the collection of all non-empty, convex and closed sets in $\mathscr{P}$. As is well-known, every element of $\mathscr{C}$ is weak*-compact. Let $\mathscr{B}(\Omega, \mathscr{F})$ be the uniform closure of the set of all simple (real-valued and finite-ranged) functions defined on $\mathscr{F}$, which will henceforth be denoted just $\mathscr{B}$.

For a given $C \in \mathscr{C}$ we define the 'min' functional $I_C : \mathscr{B} \to \mathbb{R}$ as follows: [6]

$$I_C(f) = \min_{P \in C} \int_\Omega f(\omega) \mathrm{d} P. \tag{1}$$

So I_C associates with every function $f \in \mathscr{B}$ the smallest possible integral with respect to probabilities in C. Since C is weak*-compact and $\mathscr{B}$ was chosen as above, the functional is well-defined.

Finally, we need to recall a relation between pairs of functions which is well-known in the literature on Choquet integration.

Definition 1: We say that $f, g \in \mathscr{B}$ are *comonotonic* if for every $\omega, \omega' \in \Omega$,

$$(f(\omega) - f(\omega'))(g(\omega) - g(\omega')) \geq 0$$

In other words two functions are comonotonic (short for 'commonly monotonic') if they have the same 'type' of monotonicity.

3. Affine-relatedness and additivity

In this section, we ask when the I_C functional is additive. A natural candidate condition is comonotonicity. The following example (from Klibanoff, 1996) shows, however, that comonotonicity does not guarantee additivity.

[6] To integrate with respect to a finitely additive probability we use the Stiltjes integral introduced by Hildebrandt (1934). For the full definition and related results see, e.g., Marinacci (1993).

Example 1: Suppose that $\Omega = \{\omega_1, \omega_2, \omega_3\}$ and $\mathscr{F} = 2^{\Omega}$. Consider comonotonic f and g defined as follows:

$$f(\omega_1) = 1.5 \quad f(\omega_2) = 2 \quad f(\omega_3) = 3.5 \quad \text{and}$$
$$g(\omega_1) = 0 \quad g(\omega_2) = 2.1 \quad g(\omega_3) = 4$$

Consider the set $C = \{P: P = \alpha\delta_{\omega_2} + (1-\alpha)(1/2\,\delta_{\omega_1} + 1/2\,\delta_{\omega_3}),\ \alpha \in [0,1]$. Observe that:

$$I_C(f) + I_C(g) = 2 + 2 < 4.1 = I_C(f+g).$$

We now introduce formally the appropriate condition.

Definition 2: Two functions f and g in $\mathscr{B}$ are *affinely related* if there exist $\alpha \geq 0$ *and* $\beta \in \mathbb{R}$ such that either $f(\omega) = \alpha g(\omega) + \beta$ for all $\omega \in \Omega$ or $g(\omega) = \alpha f(\omega) + \beta$ for all $\omega \in \Omega$ or both.

In other words, f and g are affinely related if either f is constant or g is constant or there exists $\alpha > 0$ and $\beta \in \mathbb{R}$ such that $f(\omega) = \alpha g(\omega) + \beta$. It is immediate to see that affinely related functions are comonotonic. The converse is, in general, not true. However, the converse does hold when both functions are defined on a set with only two points. [7]

Theorem 1: Let $f, g \in \mathscr{B}$. The following two statements are equivalent:
(i) f and g are affinely related;
(ii) $I_C(f+g) = I_C(f) + I_C(g)$ for all $C \in \mathscr{C}$.

The proof of this result makes use of the following lemma, which says that the operator I_C is additive for two functions f and g if and only if the integral of f and g is minimized (over C) by the same probability.

Lemma 1: For a given set $C \in \mathscr{C}$ and $f, g \in \mathscr{B}$, the following two statements are equivalent:
(i) $I_C(f+g) = I_C(f) + I_C(g)$;
(ii) $(\arg\min_{P \in C} \int f \mathrm{d}P) \cap (\arg\min_{P \in C} \int g \mathrm{d}P) \neq \emptyset$.

[7] Note that this observation implies that the ‘trade-off’ method for eliciting the von Neumann–Morgenstern utility index, as developed in Wakker and Deneffe (1996), can be used for decision makers whose preferences are represented by the ‘min’ functional. They show that the method can be applied to a class of preference functionals that include the CEU functional. Since they essentially use a state space with only two elements, the observation shows that the method applies to preferences represented by the ‘min’ functional as well.

Proof: The implication (ii) $\Rightarrow$ (i) is obvious. To see the converse, by (i), for $P_+ \in \arg\min_{P \in C} \int (f+g) \mathrm{d}P$:

$$\min_{P \in C} \int (f+g) \mathrm{d}P = \int f \mathrm{d}P_+ + \int g \mathrm{d}P_+ = \min_{P \in C} \int f \mathrm{d}P + \min_{P \in C} \int g \mathrm{d}P.$$

Suppose $\int f \mathrm{d}P_+ > \min_{P \in C} \int f \mathrm{d}P$. The above equalities imply $\int g \mathrm{d}P_+ < \min_{P \in C} \int g \mathrm{d}P$, a contradiction. Consequently, $\int f \mathrm{d}P_+ \leq \min_{P \in C} \int f \mathrm{d}P$. Similarly, one can show $\int g \mathrm{d}P_+ \leq \min_{P \in C} \int g \mathrm{d}P$. We conclude that:

$$P_+ \in \left(\arg\min_{P \in C} \int f \mathrm{d}P \right) \cap \left(\arg\min_{P \in C} \int g \mathrm{d}P \right),$$

which is what we wanted to prove.

To prove Theorem 1, we start by observing that each function induces, through expectation, an ordering on probabilities. We then use Lemma 1 to show that if additivity holds for functions f and g *for all* $C \in \mathscr{C}$, then the induced orderings must be identical. As they satisfy the von Neumann–Morgenstern axioms, the result then follows from the uniqueness part of their classical result.

Proof of Theorem 1: The implication (i) $\Rightarrow$ (ii) is obvious. As to the converse, if either f or g is constant we are done. Assume neither of them is constant. Define the ordering $\succeq_f$ on $\mathscr{P}$ as follows:

$$P \succeq_f P' \Leftrightarrow \int f \mathrm{d}P \geq \int f \mathrm{d}P'.$$

This ordering is transitive and complete. The ordering $\succeq_g$ is defined similarly. We show that they are equivalent, i.e.,

$$P \succeq_f P' \Leftrightarrow P \succeq_g P'.$$

Suppose, to the contrary, that there existed $P,\ P' \in \mathscr{P}$ such that $P \succ_f P'$ and $P \preceq_g P'$ (the other case is handled similarly). There are two cases to consider: (1) Suppose $P \prec_g P'$. An application of Lemma 1 with $C = \{\alpha P + (1-\alpha)P' : \alpha \in [0,1]\}$ yields a contradiction. (2) Assume $P \sim_g P'$. As g is not a constant, there exists P'' such that $\int g \mathrm{d}P \neq \int g \mathrm{d}P''$. Without loss of generality, suppose $\int g \mathrm{d}P < \int g \mathrm{d}P''$. Then $P \prec_g P''$ and, by case 1, $P'' \succeq_f P$. Hence,

$$\int g \mathrm{d}P' = \int g \mathrm{d}P < \int g \mathrm{d}P'' \text{ and}$$
$$\int f \mathrm{d}P' < \int f \mathrm{d}P \leq \int f \mathrm{d}P''.$$

For each $\alpha \in (0,1)$, $\int g \mathrm{d}(\alpha P'' + (1-\alpha)P') > \int g \mathrm{d}P$. By case 1:

$$\int f \mathrm{d}(\alpha P'' + (1-\alpha)P') \geq \int f \mathrm{d}P.$$

By continuity, $\int f \mathrm{d}P' \geq \int f \mathrm{d}P$, which contradicts $P \succ_f P'$.

We have thus proved that $P \succeq_f P' \Leftrightarrow P \succeq_g P'$. Consider now the orderings on the simple probabilities: Let $\succeq_f^s$ and $\succeq_g^s$ respectively be the restrictions to $\mathscr{P}^s$ of $\succeq_f$ and $\succeq_g$. It is easy to check that $\succeq_f^s$ and $\succeq_g^s$ are identical orderings satisfying all the axioms of the classic representation theorem of von Neumann and Morgenstern (see, e.g., Theorem 8.2 of Fishburn, 1970) on $\mathscr{P}^s$. As the representation is unique up to a positive affine transformation, for some $\alpha > 0$ and $\beta \in \mathbb{R}$ we have $f(\omega) = \alpha g(\omega) + \beta$ for all $\omega \in \Omega$.

To summarize, either one of f and g is constant, or $f = \alpha g + \beta$ for some $\alpha > 0$ and $\beta \in \mathbb{R}$, and the implication (ii) $\Rightarrow$ (i) is proved.

Remark 1: As is immediate from the proof of Theorem 1, in order to obtain this result, we do not need to use the fact that the sets $C \in \mathscr{C}$ may contain non-simple probabilities. The theorem could be restated using the set $\mathscr{C}^s$ of closed and convex sets of *simple* probabilities. In fact we shall see in Section 4 that we can restrict $\mathscr{C}$ much more without forfeiting the result.

4. From comonotonicity to affine-relatedness

An interesting question to ask is whether the result in Theorem 1 is tight. That is, do we really need the assumption that additivity (condition (ii)) holds for every $C \in \mathscr{C}$? More generally: how does the relation between f and g depend on the size of the class of Cs for which we ask additivity to hold? Lemma 1, for instance, says that for the case of the class made of a single C, we can only prove that there must be a $P \in C$ minimizing the integral of both f and g, which is much weaker than saying that f and g are affinely related (except, once again, when the Ω consists of only two points). And one might then wonder whether there can be subclasses of $\mathscr{C}$, which we can interpret decision-theoretically as describing certain 'types' of preferences, such that restricting attention to sets of priors of that type yields a less demanding additivity result. The following proposition presents one well-known example of such a subclass.

Notation We denote by $\mathscr{C}^{11}$ the set of all the Cs which have the following form: There are $\omega,\ \omega' \in \Omega$ such that:

$$C = \{\alpha \delta_\omega + (1-\alpha)\delta_{\omega'} : \alpha \in [0,1]\}.$$

We can now state (the simple proof is omitted).

Proposition 1: Let $f,\ g \in \mathscr{B}$. The following two statements are equivalent:

(i) f and g are comonotonic;
(ii) $I_C(f+g) = I_C(f) + I_C(g)$ for all $C \in \mathscr{C}^{11}$.

This result is closely connected to others in the literature. Notice that the class $\mathscr{C}^{11}$ is contained in the class of all the cores of supermodular capacities on $(\Omega, \mathscr{F})$, i.e., every $C \in \mathscr{C}^{11}$ is the core of a supermodular capacity. Since the functional I_C for such a C is a Choquet integral, going from (i) to (ii) in Proposition 1 follows immediately from the fact that the Choquet integral is comonotonic additive (see Dellacherie, 1970 and Schmeidler, 1986). The opposite direction is a slight generalization of a result of Bassanezi and Greco (1984), which is the analogue for the Choquet integral of Theorem 1.

It is fairly apparent how $\mathscr{C}^{11}$ describes a 'type' of DM: One who considers only two states possible, but is agnostic beyond that (i.e., considers all the priors which have those two points as a support in her C). Clearly, the power of the result lies in the fact that we can construct a DM of this type for *every* possible pair of points $\omega, \omega' \in \Omega$. Limiting consideration to only *some* pairs would amount to imposing, beyond the structural restriction of degeneracy, conditions on the DM's beliefs which are difficult to justify a priori, as they would favor some states or weightings over others. So we can read Proposition 1 as follows: Suppose that we want to prove additivity of the 'min' operator for all Cs in some family of sets of priors $\tilde{\mathscr{C}} \subset \mathscr{C}$. It seems quite natural that $\tilde{\mathscr{C}}$ could contain some (finite) polytopes which are convex hulls of degenerate probabilities on states. To avoid imposing arbitrary restrictions, we then have to include in $\tilde{\mathscr{C}}$ all finite polytopes, since none can be excluded a priori. In particular, $\tilde{\mathscr{C}}$ should include the subset $\mathscr{C}^{11}$ of all the Cs which are convex hulls of *two* degenerate probabilities. But then Proposition 1 says that f and g must be comonotonic for additivity to hold whatever our choice of C in $\mathscr{C}^{11}$ (hence in $\tilde{\mathscr{C}}$).

It is important to stress that, as we did above, in what follows we shall avoid arbitrariness by focusing our attention on families $\tilde{\mathscr{C}}$ that are only defined by *symmetric structural* restrictions, in the sense that if, say, $C \in \tilde{\mathscr{C}}$ is the convex hull of a two-point distribution and a degenerate distribution, then $\tilde{\mathscr{C}}$ must contain all such Cs.

Now, comonotonicity is significantly weaker than affine-relatedness. So, one might wonder whether by choosing some $\tilde{\mathscr{C}}$ which is strictly larger than $\mathscr{C}^{11}$, but still not too large, one can obtain a result which gives additivity for some relation between f and g which is weaker than affine-relatedness and stronger than comonotonicity (an obvious example: f is an increasing transformation of g). The surprising answer to this is 'no'. That is, as soon as we enlarge $\mathscr{C}^{11}$ in a fashion which, as suggested above, only imposes structural restrictions on the $C \in \tilde{\mathscr{C}}$, we can have additivity for all Cs only if f and g are affinely related. This is the upshot of the next theorem.

Before moving to that, however, let us discuss how to enlarge $\mathscr{C}^{11}$ in a symmetric way. There are two different ways to proceed: (1) to increase the number of extreme points of the Cs; (2) to allow the extreme points of the Cs to be non-degenerate. Interestingly, increasing the number of extreme points while maintaining their elementary structure does not change the result of Proposition 1.

In fact, any C which is the convex hull of (finitely many) degenerate probabilities is the core of a supermodular capacity.[8] In such a case, I_C is a Choquet integral, which is additive for all comonotonic f and g. Thus I_C is comonotonic additive for all Cs which are the convex hull of any (finite) number of degenerate probabilities. The other possibility, making the structure of extreme points richer, is the one we follow in the next definition. We enlarge $\mathscr{C}^{11}$ by considering Cs generated by two probabilities, one of which can have a support of two points, rather than only one. As we want this enlargement to be symmetric, so we shall consider all Cs with this structure.

Notation: We denote by $\mathscr{C}^{12}$ the subset of $\mathscr{C}$ consisting of all $C \in \mathscr{C}$ of the form:

$$C = \{ \alpha P + (1-\alpha) Q: \alpha \in [0,1] \}$$

where there exist $\beta \in [0,1]$ and $\omega,\ \omega',\ \omega'' \in \Omega$ such that:

$$P = \delta_\omega \text{ and } Q = \beta \delta_{\omega'} + (1-\beta)\, \delta_{\omega''}.$$

The family $\mathscr{C}^{12}$ is clearly the *smallest* symmetric enlargement of $\mathscr{C}^{11}$ in the direction outlined above.[9]

Theorem 2: Let $f,\ g \in \mathscr{B}$. The following two statements are equivalent:
(i) f and g are affinely related;
(ii) $I_C(f+g) = I_C(f) + I_C(g)$ for all $C \in \mathscr{C}^{12}$.

To prove this result, we first show that if additivity holds for all $C \in \mathscr{C}^{12}$, then f must be a strictly increasing transformation of g. Then we show that the transformation must be concave as well as convex, thus linear.

Proof: The implication (i) $\Rightarrow$ (ii) is obvious. As to the converse, by Proposition 1 the two functions f and g are comonotonic. We first show that $f(\omega) \geq f(\omega')$ if and only if $g(\omega) \geq g(\omega')$. Suppose, to the contrary, that there existed $\omega,\ \omega' \in \Omega$ such that $f(\omega) > f(\omega')$ and $g(\omega) = g(\omega')$. Suppose g is not constant (otherwise,

[8] If C is the convex hull generated by, say, $\{\delta_{\omega_i}\}_{i=1}^n$, then it is the core of the supermodular capacity (called *unanimity game*) $u_{\{\omega_1, \ldots, \omega_n\}}$, which assigns weight 1 to the set $\{\omega_1,\ \ldots,\ \omega_n\}$ and all its supersets.

[9] A decision-theoretic interpretation of $\mathscr{C}^{12}$ is that such a DM acts as if she knows exactly the relative likelihood of two states, but is uncertain about the relative likelihood of the union of the two states compared to a third.

we are done). Take $\omega'' \in \Omega$ such that $g(\omega'') \neq g(\omega)$. Without loss of generality, suppose $g(\omega'') < g(\omega')$. By comonotonicity:

$$f(\omega) > f(\omega') \geq f(\omega'') \text{ and}$$
$$g(\omega) = g(\omega') > g(\omega'').$$

Consider $C^* = \{\alpha P + (1-\alpha)Q\colon\ \alpha \in [0,1]\}$ where, for $\beta \in [0,1]$, $Q = \beta\delta_\omega + (1-\beta)\delta_{\omega''}$ and $P = \delta_{\omega'}$. For $\beta < 1$ large enough, we have:

$$\beta f(\omega) + (1-\beta) f(\omega'') > f(\omega') \text{ and}$$
$$\beta g(\omega) + (1-\beta) g(\omega'') < g(\omega').$$

This implies $(\arg\min_{R \in C^*} \int f \mathrm{d} R) \cap (\arg\min_{R \in C^*} \int g \mathrm{d} R) = \emptyset$, which is impossible by Lemma 1. Since $f(\omega) \geq f(\omega')$ if and only if $g(\omega) \geq g(\omega')$, there exists a strictly increasing transformation ϕ: Range$(g) \to \mathbb{R}$ such that $f(\omega) = \phi(g(\omega))$ for all $\omega \in \Omega$. We show that ϕ is concave. Choose any ω, ω', $\omega'' \in \Omega$ such that $g(\omega) \leq g(\omega') \leq g(\omega'')$. For $\alpha \in [0,1]$, $g(\omega) \leq \alpha g(\omega) + (1-\alpha)\ g(\omega'') \leq g(\omega'')$. By Lemma 1 on C^*,

$$(\alpha f(\omega) + (1-\alpha) f(\omega'') - f(\omega'))(\alpha g(\omega) + (1-\alpha) g(\omega'') - g(\omega')) \geq 0.$$

Since $f = \phi \circ g$, this implies that it can never be the case that:

$$\alpha\phi(g(\omega)) + (1-\alpha)\phi(g(\omega'')) > \phi(g(\omega')) \text{ and}$$
$$\alpha g(\omega) + (1-\alpha) g(\omega'') < g(\omega').$$

We claim that this implies that ϕ is concave on the range of g. For, suppose not. Then, there exist $g(\omega)$, $g(\omega')$, $g(\omega'')$ (without loss of generality assume $g(\omega) < g(\omega'')$) and $\bar{\alpha} \in (0,1)$ such that $g(\omega') = \bar{\alpha} g(\omega) + (1-\bar{\alpha}) g(\omega'')$ and:

$$\phi(\bar{\alpha} g(\omega) + (1-\bar{\alpha}) g(\omega'')) < \bar{\alpha}\phi(g(\omega)) + (1-\bar{\alpha})\phi(g(\omega'')).$$

Since ϕ is strictly increasing, $\phi(g(\omega)) < \phi(g(\omega')) < \phi(g(\omega''))$. Therefore, for $\alpha > \bar{\alpha}$ but close to $\bar{\alpha}$ we have $\alpha\phi(g(\omega)) + (1-\alpha)\phi(g(\omega'')) > \phi(g(\omega'))$ and $\alpha g(\omega) + (1-\alpha) g(\omega'') < g(\omega')$, a contradiction. Similarly, as it cannot be that:

$$\alpha\phi(g(\omega)) + (1-\alpha)\phi(g(\omega'')) < \phi(g(\omega')) \text{ and}$$
$$\alpha g(\omega) + (1-\alpha) g(\omega'') > g(\omega').$$

one can prove that ϕ is convex on the range of g. We conclude that ϕ is linear on the range of g, as wanted.

Theorem 2 shows that the result of Theorem 1 does not depend on the fact that we require additivity for all $C \in \mathscr{C}$. In fact the result is true if we only require additivity to hold on $\mathscr{C}^{12}$. The theorem also proves that as soon as we allow the possibility that some of the extreme points of the Cs are nondegenerate then additivity can hold in general only if f and g are affinely related. Quite surprisingly then, there are no 'intermediate' relations between comonotonicity and affine-relatedness.

5. Max and min

For obvious reasons of symmetry, all the results of this paper can be proved for the 'max' operator:

$$J_C(f) = \max_{P \in C} \int_\Omega f(\omega) \mathrm{d}P.$$

For instance, we can rewrite Lemma 1 as follows: Given $C \in \mathscr{C}$, the J_C functional is additive if and only if:

$$\left(\arg\max_{P \in C} \int f \mathrm{d}P\right) \cap \left(\arg\max_{P \in C} \int g \mathrm{d}P\right) \neq \emptyset.$$

An interesting question is whether these results could also be generalized to a larger class of functionals. For instance, a fairly general class is the one containing functionals of the form, for $\lambda \in [0,1]$,

$$K_C(f) = \lambda I_C(f) + (1-\lambda) J_C(f). \tag{2}$$

Remark 2: Some thought reveals that the class discussed above is equivalent to the class of all functionals with the following structure: For a given $f \in \mathscr{B}$ and a given $C \in \mathscr{C}$, consider the set $\mathrm{Range}_C(f) \equiv \{x \in \mathbb{R}\colon \exists P \in C \text{ s.t. } x = \int f \mathrm{d}P\}$. It can be seen to be a closed and bounded interval in $\mathbb{R}$, say $[a,b]$. Consider now [0,1] with the usual Borel σ-algebra and a measure μ on it. Let:

$$K_C(f) = (b-a)\left(\int_{[0,1]} x \mathrm{d}\mu\right) + a.$$

This corresponds to the following idea: For every function f find its range of possible values $[a,b]$ (which could of course be degenerate), and then obtain a 'summary statistic' of the interval by (transforming it into [0,1] and) integrating with respect to some measure μ. The key aspect is that the procedure does not depend on the identity of the interval $[a,b]$, that is, all intervals are treated in the same way. It is clear that the 'min' and 'max' operator fall in this class. Another obvious example is the mean of the values of the interval $(\lambda = 1/2)$, which corresponds to the case in which has a uniform density over [0,1]. Clearly every such functional can be written as in Eq. (2) for some $\lambda \in [0,1]$.

In Sections 5.1 and 5.2 we show that, while the results in the previous sections can all be generalized to this class of functionals when $\lambda \neq 1/2$, the 'uniform' case of $\lambda = 1/2$ is a bit delicate.

5.1. The case of $\lambda \neq 1/2$

The following proposition contains the required extensions.

Proposition 2: Suppose that $\lambda \neq 1/2$ and $f,\ g \in \mathscr{B}$. Then the K_C functional is additive for every $C \in \mathscr{C}^{11}$ if and only if f and g are comonotonic. It is additive for every $C \in \mathscr{C}^{12}$ if and only if f and g are affinely related.

The key result for proving the proposition is Lemma 2, which requires the following.

Notation: We denote by $\mathscr{C}^{PQ}$ the set of all $C \in \mathscr{C}$ which have the following form:

$$C = \{\alpha P + (1-\alpha)Q \colon \alpha \in [0,1]\}$$

for some $P,\ Q \in \mathscr{P}$.

Lemma 2: For every $f,\ g \in \mathscr{B}$ and for every $C \in \mathscr{C}^{PQ}$,

$$I_C(f+g) - I_C(f) - I_C(g) = J_C(f) + J_C(g) - J_C(f+g). \tag{3}$$

Proof: Let P and Q be the extreme points of $C \in \mathscr{C}^{PQ}$. It is immediate to notice that in calculating smallest and largest integrals, we can restrict our attention to P and Q, rather than their convex combinations. Now observe the following:

$$\begin{aligned} J_C(f) + J_C(g) - J_C(f+g) &= -I_C(-f) - I_C(-g) - (-I_C(-f-g)) \\ &= I_C(-f-g) - I_C(-f) - I_C(-g). \end{aligned}$$

In general, the integrals of the different functions will be minimized by different probabilities. But it is clear that if, say, $I_C(f) = \int f \mathrm{d}P$ then $I_C(-f) = \int (-f)\mathrm{d}Q$. For instance, suppose that P minimizes g and $f+g$ and Q minimizes f. Then:

$$\begin{aligned} I_C(f+g) - I_C(f) - I_C(g) &= \int f \mathrm{d}P - \int f \mathrm{d}Q \\ &= I_C(-f-g) - I_C(-f) - I_C(-g). \end{aligned}$$

Checking all the different cases (and remembering Lemma 1), we can conclude that Eq. (3) holds for all $C \in \mathscr{C}^{PQ}$.

Proof of Proposition 2: We notice that for every $\lambda \in [0,1]$:

$$K_C(f+g) = K_C(f) + K_C(g) \tag{4}$$

can be rewritten as follows:

$$\lambda\big(I_C(f+g) - I_C(f) - I_C(g)\big) = (1-\lambda)\big(J_C(f) + J_C(g) - J_C(f+g)\big). \tag{5}$$

In turn, by Lemma 2, for $C \in \mathscr{C}^{PQ}$ this equality becomes:

$$\lambda\big(I_C(f+g) - I_C(f) - I_C(g)\big) = (1-\lambda)\big(I_C(f+g) - I_C(f) - I_C(g)\big). \tag{6}$$

This shows that the results presented so far hold for the K_C functional as long as $\lambda \neq 1/2$. In fact, in such case the equality (Eq. (6)) holds if and only if $I_C(f+g) - I_C(f) - I_C(g) = 0$, i.e.,

$$I_C(f+g) = I_C(f) + I_C(g).$$

Consequently, as long as $\lambda \neq 1/2$, the analysis of the additivity of K_C reduces to that of the 'min' functional I_C. Therefore, as both $\mathscr{C}^{11}$ and $\mathscr{C}^{12}$ are contained in $\mathscr{C}^{PQ}$, a straightforward application of the results of the previous sections completes the proof.

5.2. *The case of $\lambda = 1/2$*

What happens in this case? At first, this balance of 'min' and 'max' might appear to lead back to additivity, (decision theoretically: to SEU). The following proposition shows that this is true for $C \in \mathscr{C}^{PQ}$.

Proposition 3: Suppose $\lambda = 1/2$. Then the K_C functional is linear on $\mathscr{B}$ if $C \in \mathscr{C}^{PQ}$. In particular:

$$K_C(f) = \int f \mathrm{d}\rho$$

where ρ: $\mathscr{F} \to [0,1]$ is given by:

$$\rho = (1/2)P + (1/2)Q.$$

Proof: The functional K_C is positive homogeneous. By Lemma 2, it is also additive. For any $A \in \mathscr{F}$, set $\rho(A) = K_C(1_A)$, where 1_A is the indicator function of A. Let f be a simple function in $\mathscr{B}$, and let $c_f \in \mathbb{R}$ be such that $f(\omega) + c_f \geq 0$ for all $\omega \in \Omega$. It is easy to check that $K_C(f + c_f) = \int (f + c_f) \mathrm{d}\rho$, so that:

$$K_C(f) + c_f = K_C\big(f + c_f\big) = \int\big(f + c_f\big)\mathrm{d}\rho = \int f \mathrm{d}\rho + c_f.$$

Therefore, $K_C(f) = \int f \mathrm{d}\rho$. Using uniform convergence, it is easy to check that this holds for all $f \in \mathscr{B}$. Since for all $A \in \mathscr{F}$:

$$K_C(1_A) = (1/2) P(A) + (1/2) Q(A),$$

if P and Q are the extreme points of C, the result follows.

The following example demonstrates, however, that additivity does not hold for every $C \in \mathscr{C}$.

Example 2: Suppose that $\Omega = \{\omega_1, \omega_2, \omega_3\}$ and $\mathscr{F} = 2^{\Omega}$. Consider f and g as in example 1. Set C = conv$(1/2\,\delta_{\omega_1} + 1/2\,\delta_{\omega_3}, \delta_{\omega_2}, \delta_{\omega_3})$, where 'conv' denotes the convex hull. It is immediate to calculate that:

$$I_C(f+g) - I_C(f) - I_C(g) = 0.1 > 0 = J_C(f) + J_C(g) - J_C(f+g),$$

thus violating additivity by Eq. (5).

Remark 3: In fact, one can show the following: For $\lambda = 1/2$, if Ω is finite then $C \notin \mathscr{C}^{PQ}$ implies that K_C is not additive. That is, the converse of Proposition 3 holds for finite Ω. [10]

Having thus concluded that K_C is not in general additive when $\lambda = 1/2$, we now ask whether we can prove an additivity result analogous to that in Section 5.1. The 'if' statements in Proposition 2 obviously extend. The following example shows that the 'only if' statements, however, do not extend. Specifically, we present a pair of non-comonotonic functions for which additivity when $\lambda = 1/2$ holds for all C.

Example 3: Suppose that $\Omega = \{\omega_1, \omega_2, \omega_3\}$ and $\mathscr{F} = 2^{\Omega}$. Consider f and g defined as follows:

$$\begin{array}{lll} f(\omega_1) = a & f(\omega_2) = 0 & f(\omega_3) = -a \text{ and} \\ g(\omega_1) = -b & g(\omega_2) = 0 & g(\omega_3) = b \end{array}$$

where $a \geq b > 0$. It is easy to convince oneself that for every set $C \in \mathscr{C}$, only two extreme points will play a role in the evaluation of f, g and $f+g$. But then, if $\lambda = 1/2$, Lemma 2 applies and shows that additivity holds.

[10] A proof is available from the authors upon request. We conjecture that, at least when C has finitely many extreme points, the converse also holds for infinite Ω.

One should notice that the example is quite general, and that it does not depend on our choice of a set Ω with three points. That is, for every set Ω, it is possible to construct a pair of functions like f and g in the example, such that: (1) they are not comonotonic (hence, not affinely related), (2) whatever the set C is, f and g and $f+g$ are integrated only with respect to two probabilities, so that additivity of K_C follows from Lemma 2. So, quite surprisingly, we conclude by observing that the powerful conclusions obtained for the other cases do not hold for the case of $\lambda = 1/2$.

Acknowledgements

We thank Michèle Cohen, Truman Bewley and three anonymous referees for helpful comments. Ghirardato is very grateful to CORE, Université Catholique de Louvain, for its hospitality during the period in which a draft of this paper was written. We also thank participants at the Midwest Mathematical Economics Meetings at Washington University, St. Louis (October 1996). Marinacci is grateful to the Social Sciences and Humanities Research Council of Canada for financial support. A previous version was circulated as 'Linearity with Multiple Priors.'

References

Anscombe, F.J., Aumann, R.J., 1963. A definition of subjective probability. Ann. Math. Stat. 34, 199–205.

Bassanezi, R.C., Greco, G.H., 1984. Sull'additività dell'integrale. Rendiconti del Seminario di Matematica dell'Università di Padova 72, 249–275.

Chateauneuf, A., 1991. On the use of capacities in modelling uncertainty aversion and risk aversion. J. Math. Econ. 20, 343–369.

Dellacherie, Claude, 1970. Quelques commentaires sur les prolongements de capacités. Séminaire Probabilités V, Strasbourg, number 191. Lecture Notes in Mathematics. Springer-Verlag, Berlin.

Eichberger, J., Kelsey, D., 1995. Uncertainty aversion and preference for randomization. J. Econ. Theory 71, 31–43.

Ellsberg, D., 1961. Risk, ambiguity, and the Savage axioms. Q. J. Econ. 75, 643–669.

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

Ghirardato, P., 1997. On independence for non-additive measures, with a Fubini theorem. J. Econ. Theory 73, 261–291.

Gilboa, I., Schmeidler, D., 1988. Maxmin expected utility with a non-unique prior. J. Math. Econ. 18, 141–153.

Hildebrandt, T.H., 1934. On bounded linear functional operators. Trans. Am. Math. Soc. 36, 868–875.

Klibanoff, Peter, July 1996. Characterizing uncertainty aversion through preference for mixtures. CMS-EMS Discussion Paper 1159, Northwestern University.

Marinacci, Massimo, September, 1993. Finitely additive and epsilon Nash equilibria. Mimeo, Northwestern University (Forthcoming in International Journal of Game Theory).

Savage, Leonard J., 1954. The Foundations of Statistics. Wiley, New York.
Schmeidler, D., 1986. Integral representation without additivity. Proc. Am. Math. Soc. 97, 255–261.
Schmeidler, D., 1989. Subjective probability and expected utility without additivity. Econometrica 57, 571–587.
Wakker, P.P., Deneffe, D., 1996. Eliciting von Neumann–Morgenstern utilitites when probabilities are distorted or unknown. Manage. Sci. 42, 1131–1150.