DIVISION OF THE HUMANITIES AND SOCIAL SCIENCES
CALIFORNIA INSTITUTE OF TECHNOLOGY
PASADENA, CALIFORNIA 91125

LINEARITY WITH MULTIPLE PRIORS

Paolo Ghirardato
California Institute of Technology

Peter Klibanoff
Northwestern University

Massimo Marinacci
University of Toronto

SOCIAL SCIENCE WORKING PAPER 980

July 1996

Linearity with Multiple Priors*

Paolo Ghirardato Peter Klibanoff Massimo Marinacci

Abstract

We characterize the types of functions over which the functional defined as the "min" of integrals with respect to probabilities in a given non-empty closed and convex class is linear. This happens exactly when "integrating" functions which are positive affine transformations of each other (or when one is constant). We show that the 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. JEL Classification Number: D81.

Introduction and Motivation

In the wake of Daniel Ellsberg's intriguing experiments [5], decision theorists have produced a number of interesting axiomatic models of behavior aimed at explaining his results and, more generally, capturing ambiguous beliefs. Schmeidler [14] suggested extending the classical Subjective Expected Utility (SEU) model of Anscombe and Aumann [1] and Savage [12] by allowing the decision maker (DM)'s preferences 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 [8] and Chateauneuf [3]. As is well-known (see Schmeidler [14]) 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*.

* Ghirardato is very grateful to CORE, Université Catholique de Louvain, for its hospitality during the period in which this paper was written.

While the linearity of the Choquet integral has been carefully studied by mathematicians and decision theorists, the linearity of the functional defined as the "min" of integrals over a (closed and convex) set of probabilities has not. This paper addresses exactly that issue. More specifically we ask the following question: Suppose that we are modeling 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 C. Given two acts f and g,[1] we want to know whether there is a C such that the "min" of the expected utilities on C is not additive for $f+g$. We will show (Theorem 1) that this will always be possible as long as f and g are not positive affine transformations of each other and neither is constant (as we say: they are not "affinely related"). Technically, the "min" functional is linear for every C when and only when we are integrating the sum of two affinely related functions. Since affine-relatedness is stronger than comonotonicity, this implies that the "min" functional will be linear for a much smaller class of functions than the Choquet integral.

A common feature of the CEU and multiple prior models is that they were originally developed in an Anscombe-Aumann set-up [1], in which the existence of an independent randomizing device (a "roulette wheel") is assumed. This gives meaning to "objective" mixtures (technically: pointwise convex combinations) of acts. In recent work, Klibanoff [10] characterizes when a DM described by one of these two models can be made better off by mixing. 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). Our exploration of the linearity of the "min" functional is related to this since 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 (see below), "mixing" between comonotonic acts does *not* make a CEU DM better off.

Besides its import on the issue of preference for mixtures, and its purely mathematical interest, the question of the linearity of the "min" functional also bears on the possibility of extending results which work for the additive integral. For instance, as Ghirardato [7] 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 obtain a similar result for the "min" functional. The very limited linearity we get, though, implies that the scope of such a result would be very narrow. For this reason, we inquire whether it is possible to obtain broader linearity 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 (which assign probability one to a specific state ω), we again have that linearity holds only for affinely related functions. This tighter version of the linearity result (Theorem 2) shows that the narrow linearity of the "min" functional cannot be avoided in general.

[1] From here on we fix the utility function u and write simply f and g for $u \cdot f$ and $u \cdot g$. Hence all acts are random variables.

Finally we consider the validity of the results discussed so far for other functionals. They immediately extend to the "max" functional, which associates with a function the largest expected value with respect to probabilities in C. Also they extend to any linear function of the "min" and "max" (Proposition 2), with a very peculiar exception: Those functions which treat the "min" and "max" identically.

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

1 Preliminaries

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

For a *given* $C \in \mathcal{C}$ we define the "min" functional $I_C : \mathcal{B} \to \mathbf{R}$ as follows:[2]

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

So I_C associates with every function $f \in \mathcal{B}$ the smallest possible integral with respect to probabilities in C. Since C is weak*-compact and $\mathcal{B}$ 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 $f, g \in \mathcal{B}$ 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.

[2] To integrate with respect to a finitely additive probability we use the Stiltjes integral introduced by Hildebrandt [9]. For the full definition and related results see, e.g., Marinacci [11].

2 Affine-Relatedness and Linearity

In this section we prove the basic result that the I_C functional can be additive for a pair of functions f and g in $\mathcal{B}$ for *every* class of probabilities $C \in \mathcal{C}$ if and only if f and g are positive affine transformations of each other or one is constant. First we introduce formally the relation between functions that we will use.

Definition 2 *Two functions f and g in $\mathcal{B}$ are affinely related if there exist $\alpha \geq 0$ and $\beta \in \mathbf{R}$ 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 $\beta \in \mathbf{R}$ 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.

Theorem 1 *Let $f, g \in \mathcal{B}$. 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 $C \in \mathcal{C}$.

To prove theorem 1 we shall use the following lemma, which says that the operator I_C is linear for two functions f and g if the integral of f and g is minimized (over C) by the same probability.

Lemma 1 *For a given set $C \in \mathcal{C}$ and $f, g \in \mathcal{B}$, the following two statements are equivalent*

(i) $I_C(f+g) = I_C(f) + I_C(g)$;

(ii) $\left(\arg\min_{P \in C} \int f dP\right) \cap \left(\arg\min_{P \in C} \int g dP\right) \neq \emptyset$.

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

$$\begin{aligned} \min_{P \in C} \int (f+g)\, dP &= \int f\, dP_+ + \int g\, dP_+ \\ &= \min_{P \in C} \int f\, dP + \min_{P \in C} \int g\, dP. \end{aligned}$$

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

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

which is what we wanted to prove. ∎

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

$$P \succeq_f P' \Longleftrightarrow \int f dP \geq \int f dP'.$$

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

$$P \succeq_f P' \Longleftrightarrow P \succeq_g P'.$$

Suppose, to the contrary, that there existed $P, P' \in \mathcal{P}$ 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 $\int g\, dP \neq \int g\, dP''$. Without loss of generality, suppose $\int g\, dP < \int g\, dP''$. Then $P \prec_g P''$ and, by case 1, $P'' \succeq_f P$. Hence

$$\begin{array}{rcccl} \int g\, dP' & = & \int g\, dP & < & \int g\, dP'' \text{ and} \\ \int f\, dP' & < & \int f\, dP & \leq & \int f\, dP''. \end{array}$$

   For each $\alpha \in (0,1)$, $\int g\, d\left(\alpha P'' + (1-\alpha) P'\right) > \int g\, dP$. By case 1

$$\int f\, d\left(\alpha P'' + (1-\alpha) P'\right) \geq \int f\, dP.$$

   By continuity, $\int f\, dP' \geq \int f\, dP$, 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 on $\mathcal{P}^s$ of $\succeq_f$ and $\succeq_g$. It is easy to check that $\succeq_f$ and $\succeq_g$ are identical orderings satisfying all the axioms of the classic representation theorem of von Neumann and Morgenstern (see e.g. Fishburn [6, Thm. 8.2]) on $\mathcal{P}^s$. As the representation is unique up to a positive affine transformation, for some $\alpha > 0$ and $\beta \in \mathbf{R}$ we have $f(\omega) = \alpha g(\omega) + \beta$ for all $\omega \in \Omega$.

To summarize, either one of f and g are constant, or $f = \alpha g + \beta$ for some $\alpha > 0$ and $\beta \in \mathbf{R}$, and the implication (ii) $\Longrightarrow$ (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 $C \in \mathcal{C}$ may contain non-simple probabilities. The theorem could be restated using the set $\mathcal{C}^s$ of closed and convex sets of *simple* probabilities. In fact we shall see in the next section that we can restrict $\mathcal{C}$ much more without forfeiting the result.

3 From Comonotonicity to Affine-Relatedness

An interesting question to ask is whether the result in Theorem 2 is tight. That is, do we really need the assumption that linearity (condition (ii)) holds for *every* $C \in \mathcal{C}$? More generally: how does the relation between f and g depend on the size of the class of Cs for which we ask linearity 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 $\mathcal{C}$, 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 linearity result. The following proposition presents one well-known example of such a sub-class. Let $\mathcal{C}'$ be 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 $f, g \in \mathcal{B}$. They are comonotonic if and only if*

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

for all $C \in \mathcal{C}'$.

This result is closely connected to others in the literature. Notice that the class $\mathcal{C}'$ is contained in the class of all the cores of supermodular capacities on $(\Omega, \mathcal{F})$, i.e., every $C \in \mathcal{C}'$ is the core of a supermodular capacity. Since the functional I_C for such a C is a Choquet integral, the "only if" in Proposition 1 follows immediately from the fact that the Choquet integral is comonotonic additive (see Dellacherie [4] and Schmeidler [13]). The "if" is a slight generalization of a result of Bassanezi and Greco [2], which is analogous to our Theorem 1.

It is fairly apparent how $\mathcal{C}'$ 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 linearity of the "min" operator for all Cs in some family of sets of priors $\tilde{\mathcal{C}} \subset \mathcal{C}$. It seems quite natural that $\tilde{\mathcal{C}}$ 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 $\tilde{\mathcal{C}}$ *all* finite polytopes, since none can be excluded *a priori*. In particular, $\tilde{\mathcal{C}}$ should include the subset $\mathcal{C}'$ 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 linearity to hold whatever our choice of C in $\mathcal{C}'$ (hence in $\tilde{\mathcal{C}}$).

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

Now, comonotonicity is significantly weaker than affine-relatedness. So one might wonder whether by choosing some $\tilde{\mathcal{C}}$ which is strictly larger than $\mathcal{C}'$, but still not too large, one can obtain a result which gives linearity 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 $\mathcal{C}'$ in a fashion which, as suggested above, only puts structural restrictions on the $C \in \tilde{\mathcal{C}}$, we can have linearity 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 $\mathcal{C}'$ 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.[3] 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 $\mathcal{C}'$ 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. Let $\mathcal{C}''$ be the subset of $\mathcal{C}$ consisting of all $C \in \mathcal{C}$ of the form

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

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

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

The family $\mathcal{C}''$ is clearly the *smallest* symmetric enlargement of $\mathcal{C}'$ in the direction outlined above.

Theorem 2 *Let $f, g \in \mathcal{B}$. They are affinely related if and only if*

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

for all $C \in \mathcal{C}''$.

[3] 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.

Proof: The "only if" part 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 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

$$\begin{array}{rcccl} f(\omega) & > & f(\omega') & \geq & f(\omega'') \text{ and} \\ g(\omega) & = & g(\omega') & > & g(\omega''). \end{array}$$

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

$$\begin{array}{rcl} \beta f(\omega) + (1-\beta) f(\omega'') & > & f(\omega') \text{ and} \\ \beta g(\omega) + (1-\beta) g(\omega'') & < & g(\omega'). \end{array}$$

This implies $\left(\arg\min_{P \in C^*} \int f\, dP\right) \cap \left(\arg\min_{P \in C^*} \int g\, dP\right) = \emptyset$, 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 $\phi : \text{Range}(g) \to \mathbf{R}$ such that $f(\omega) = \phi(g(\omega))$ for all $\omega \in \Omega$. We show that ϕ is concave. Choose any $\omega, \omega', \omega'' \in \Omega$ 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

$$\begin{array}{rcl} \alpha\phi\left(g(\omega)\right) + (1-\alpha)\,\phi\left(g(\omega'')\right) & > & \phi\left(g(\omega')\right) \text{ and} \\ \alpha g(\omega) + (1-\alpha)\, g(\omega'') & < & g(\omega'). \end{array}$$

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 $\alpha\phi(g(\omega)) + (1-\alpha)\,\phi\left(g(\omega'')\right) > \phi\left(g(\omega')\right)$ and $\alpha g(\omega) + (1-\alpha)\, g(\omega'') < g(\omega')$, a contradiction. Similarly, as it cannot be that

$$\begin{array}{rcl} \alpha\phi\left(g(\omega)\right) + (1-\alpha)\,\phi\left(g(\omega'')\right) & < & \phi\left(g(\omega')\right) \text{ and} \\ \alpha g(\omega) + (1-\alpha)\, g(\omega'') & > & g(\omega'). \end{array}$$

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 linearity for all $C \in \mathcal{C}$. In fact the result is true if we only require linearity to hold on $\mathcal{C}''$. The theorem also proves that as soon as we allow the possibility that some of the extreme points of the Cs are nondegenerate[4] then linearity 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.

4 Extensions

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

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

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

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

An interesting question is whether these results could also be generalized to a larger class of functionals. For instance, a fairly general (see Remark 2 below) class is the one containing functionals of the form, for $\lambda \in [0,1]$,

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

It turns out that the results in the previous sections can be all generalized to this class of functionals, with one peculiar exception. The following lemma provides the key result for the extension. Let $\bar{\mathcal{C}}$ be the class of all the sets of probabilities with two extreme points. That is, $C \in \bar{\mathcal{C}}$ if there are $P, Q \in \mathcal{P}$ such that $C = \{\alpha P + (1-\alpha)Q : \alpha \in [0,1]\}$.

Lemma 2 *For every $f, g \in \mathcal{B}$ and for every $C \in \bar{\mathcal{C}}$,*

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

Proof: Let P and Q be the extreme points of $C \in \bar{\mathcal{C}}$. 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:

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

[4] A decision-theoretic interpretation 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.

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

$$\begin{aligned} I_C(f+g) - I_C(f) - I_C(g) &= \int f\,dP - \int f\,dQ \\ &= I_C(-f-g) - I_C(-f) - I_C(-g). \end{aligned}$$

Checking all the different cases (and remembering lemma 1), we can conclude that (3) holds for all $C \in \bar{\mathcal{C}}$. ∎

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:

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

In turn, by Lemma 2, for $C \in \bar{\mathcal{C}}$ this equality becomes

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

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 (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 linearity of K_C reduces to that of the "min" functional I_C. Therefore, as both $\mathcal{C}'$ and $\mathcal{C}''$ are contained in $\bar{\mathcal{C}}$, the results of the previous sections immediately imply the following:

Proposition 2 *Suppose that $\lambda \neq 1/2$ and $f, g \in \mathcal{B}$. Then (4) holds for every $C \in \mathcal{C}'$ if and only if f and g are comonotonic. It holds for every $C \in \mathcal{C}''$ if and only if f and g are affinely related.*

So our results extend to the case in which $\lambda \neq 1/2$. But what about $\lambda = 1/2$? Then by Lemma 2 we can have (4) holding for every $C \in \bar{\mathcal{C}}$ (and hence in all $C \in \mathcal{C}''$) for *every* f and g, regardless of their form. Is it possible to do more by going beyond $\bar{\mathcal{C}}$? The following example shows that it is even impossible to obtain a general result yielding comonotonicity.

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

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

where $a \geq b > 0$. Clearly f and g are not comonotonic. However it is easy to convince oneself that for every set $C \in \mathcal{C}$, only two extreme points will play a role in the evaluation of f, g and $f+g$. But then lemma 2 applies and it shows that (4) holds. △

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 linearity of K_C follows from lemma 2. So, quite surprisingly, the powerful conclusions obtained for the other cases do not hold for the case of $\lambda = 1/2$.

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

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

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, which corresponds to the case in which has a uniform density over $[0, 1]$. Now clearly every such functional can be written as in Eq. (2) for some $\lambda \in [0, 1]$. The case of the mean (μ uniform) corresponds to $\lambda = 1/2$, which is the black sheep in the results above.

Remark 3 Needless to say, nothing of the above strictly relies on the fact that the functional K_C is a convex combination of I_C and J_C. If K_C has the structure

$$K_C(f) = \gamma I_C(f) + \lambda J_C(f),$$

for $\gamma, \lambda \in \mathbf{R}$, then Proposition 2 holds with the assumption $\gamma \neq \lambda$. If $\gamma = \lambda$ then the problems outlined above for the case of $1/2$ arise in the same fashion. So the results in the paper extend to any linear function of I_C and J_C, as long as the coefficients are different.

References

[1] Frank J. Anscombe and Robert J. Aumann. A definition of subjective probability. *Annals of Mathematical Statistics*, 34:199–205, 1963.

[2] Rodney C. Bassanezi and Gabriele H. Greco. Sull'additività dell'integrale. *Rendiconti del Seminario di Matematica dell'Università di Padova*, 72:249–275, 1984.

[3] Alain Chateauneuf. On the use of capacities in modeling uncertainty aversion and risk aversion. *Journal of Mathematical Economics*, 20:343–369, 1991.

[4] Claude Dellacherie. Quelques commentaires sur les prolongements de capacités. In *Séminaire Probabilités V, Strasbourg*, number 191 in Lecture Notes in Mathematics, Berlin and New York, 1970. Springer-Verlag.

[5] Daniel Ellsberg. Risk, ambiguity, and the Savage axioms. *Quarterly Journal of Economics*, 75:643–669, 1961.

[6] Peter C. Fishburn. *Utility Theory for Decision Making.* J.Wiley and Sons, New York and London, 1970.

[7] Paolo Ghirardato. On independence for non-additive measures, with a Fubini theorem. Social Science Working Paper 940, Caltech, September 1995. (Forthcoming in *Journal of Economic Theory*).

[8] Itzhak Gilboa and David Schmeidler. Maxmin expected utility with a non-unique prior. *Journal of Mathematical Economics*, 18:141–153, 1988.

[9] T.H. Hildebrandt. On bounded linear functional operators. *Transactions of the American Mathematical Society*, 36:868–875, 1934.

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

[11] Massimo Marinacci. Finitely additive and epsilon Nash equilibria. Mimeo, Northwestern University (forthcoming in *International Journal of Game Theory*), September 1993.

[12] Leonard J. Savage. *The Foundations of Statistics.* J.Wiley and Sons, New York and London, 1954.

[13] David Schmeidler. Integral representation without additivity. *Proceedings of the American Mathematical Society*, 97:255–261, 1986.

[14] David Schmeidler. Subjective probability and expected utility without additivity. *Econometrica*, 57:571–587, 1989.