Aggregation of Opinions and Risk Measures

Massimiliano Amarante

Université de Montréal et CIREQ

Abstract. A ranking $\succsim$ over a set of alternatives is an *aggregation of experts' opinions* (AEO) if it depends on the experts' assessments only. We study both those rankings that result from pooling Bayesian experts and those that result from pooling possibly non-Bayesian experts. In the non-Bayesian case, we allow for the simultaneous presence of experts that may display very different attitudes toward uncertainty. We show that a unique axiom along with a mild regularity condition fully characterize those AEO rankings which are "generalized averages" of experts' opinions, in the sense that the average is obtained by using a capacity rather than a probability measure. We call these rankings *non-linear pools*. We consider a number of special cases such as linear pools (Stone 1961), concave/convex pools (Crès, Gilboa and Vieille 2011), quantiles and pools of equally reliable experts. We, then, apply our results to the theory of risk measures. We show that a wide class of risk measures can be regarded as non-linear pools. This not only includes the coherent risk measures of Artzner et al. 1999, but also measures like the Value at Risk, which fail sub-additivity. Our results cast a different light on the dispute on the use of risk measures as well as on the debate on the properties that are desirable in a risk measure. We also briefly discuss the possibility of extending our findings to include the convex measures of Föllmer and Schied (2002) as well as their non-subadditive extensions.

1. Introduction

More often than not, decisions are made with the aid of expert advise. This is typical in the case of Governments, Banks, CEOs, who all use experts of various sorts, but it is not less common in the case of everyday life decisions: we search the internet, watch tv, read newspapers, consult doctors and financial experts, and so on. In this paper, we study how the opinions that come from these different sources are to aggregated in order to make decisions.

As the topic is certainly not new, it is appropriate to begin by highlighting the added value of the present work. For one, the analysis is conducted here at a high level of generality: the paper's main result obtains under one minor regularity condition and only one (very mild) axiom. This is not, however, generality for the sake of itself as it immediately leads to applications. As we shall see, our general class of aggregation procedures includes, and thus provides a rationale for, many

Key words and phrases. aggregation of opinions, non-linear opinion pools, non-Bayesian experts, concave pools, equal reliability, risk measures, value at risk, choquet integral

JEL classification: D7; D8

email: massimiliano.amarante1@gmail.com.

aggregation procedures that are commonly used in practice but that do not fit existing models. What is more, the generality of our result allows one to interpret many economic indicators as being the outcome of an aggregation procedure satisfying our conditions. In many cases, this possibility casts an entirely new light on the issue of which properties are desirable for these indicators.

The paper unfolds as follows. We begin our study in Section 2. The set-up portrays an individual who has to choose among a set of alternatives. He has a utility function on consequences, but the consequences of his choices are not known in advance because of the presence of uncertainty. To deal with the uncertainty, the individual resorts to the opinions of a collection of experts. As different experts have, generally speaking, different opinions, he needs to aggregate these opinions in order to rank the alternatives. We focus first on the case where each expert is Bayesian, that is she describes the uncertainty by means of a single probability measure over a set of states (Section 2). There are at least three reasons which suggest preceding the study of the general case with that of Bayesian experts. Firstly, with the notable exception of the paper by Crès, Gilboa and Vieille [**9**], the literature has focused mostly on this case. Adopting the same setting will make it easy to identify the nature and the entity of our contribution. Secondly, by focusing on *linear* (=Bayesian) experts, we will be able to ascribe emerging non-linearities only to the aggregation procedures. This will result in a better understanding of these procedures and will lead to more transparent interpretations. Thirdly, and perhaps most importantly, Proposition 2 in Section 3 shows that the problem of aggregating a large variety of non-Bayesian experts (not even necessarily of the same type) can be represented as one of aggregating a set of fictitious Bayesian experts. Equivalently, Proposition 2 shows that the individual's rankings (or the corresponding economic indicators, in the intepretation mentioned above) that may result from the aggregation of non-Bayesian experts are exactly the same as those that obtain from the aggregation of Bayesian experts.

The main result on the aggregation of Bayesian experts, Theorem 1 in Section 2, states that any such aggregation can be regarded as a non-linear weighing of experts' opinions, in the sense that the operation of weighing is represented by a capacity rather than a probability measure. In Subsection 2.1, we give some examples of these non-linear aggregation protocols, and show that they have distinctive advantages over the classical linear ones (that is, those that use a probability measure to weigh experts). For instance, in the case of a decision maker who considers investing in a financial portfolio, a classical linear protocol demands that the decision maker would assign weights on the experts irrespective of the portfolio to be evaluated. In contrast, in our non-linear protocols, the decision maker can account for the fact that the experts' competence may vary across different portfolios by varying the weights as a function of the portfolio being evaluated. This might be necessary, for instance, because the decision maker knows that some experts have better information about, say, market X while others have better information about market Y. In Subsection 2.2, still within the framework of Bayesian experts, we characterize the subclass of aggregation procedure that are "cautious" in the sense that they are biased toward the most pessimistic among the experts' evaluations. Needless to say, the characterization of the twin "optimistic" aggregation procedures follows at once. In Section 3, we move to the study of non-necessarily Bayesian experts. As stated above, the basic result is that of Proposition 2, which implies that any outcome achievable from the aggregation of non-Bayesian experts can be replicated by an aggregation of Bayesian experts. It is worth noting that we allow for a large variety of experts' pools as these may contain experts that are ambiguity averse alongside with experts that are ambiguity loving as well as experts that display a variable attitude toward ambiguity. In Subsection 3.1, we consider the special case where each expert is of the maxmin expected utility type, and derive the representation of Crès, Gilboa and Vieille [**9**]. In Section 4, we consider another subclass of the aggregation procedures of Proposition 2. This consists of the aggregations that deem all experts equally reliable. Most of the results of this section are already known, mainly from other strands of literature, and are included here only as examples. They serve an important purpose, nonetheless. For one thing, they correspond to aggregation procedures that are widely used in practice and are at the foundations of many economic indicators. At the same time, since they are as general as the procedures of Proposition 2 (aside from the hypothesis of equal reliability which, of course, renders the characterization sharper), they show that the generality of Proposition 2 is necessary to accommodate these procedures.

Throughout the paper, we give several applications and examples; some are generic, so as to highlight the possibility of a certain aggregation procedure (see Subsection 2.1), some are specific (see Section 4). There is one application that we develop quite extensively: that to the theory of Risk Measures (Section 5). Risk Measures are indicators of the variability associated with a portfolio or a company and are commonly used in practice, often for regulatory purposes (see, for instance, the Basel Accords of 1988, 1999 and 2010). They are constantly at the center of heated debates concerning their empirical performances, the legitimacy of their use for certain specific purposes as well as debates regarding the properties that are desirable in a risk measure. For instance, the *Value at Risk* (VaR; see [**11**] and [**18**]), the most popular risk measure, whose utilization is recommended in all Basel Accords, is constantly criticized for not being sub-additive since, by lacking this property, it seems not to recognize the value of diversification. Some of these debates have even reached the public eye, like those following the financial crisis of 2007-08, and that culminated with N. Taleb testifying before the US Congress and asking for the banning of the VaR. Taleb has gone as far as questioning the very possibility of a risk measure on the grounds that the risk of rare events is inherently non-measurable. To a certain extent, arguments of this sort can be traced back to Markowitz, one of the forefathers of the theory of risk measures. Markowitz, in complete accordance with Savage's view, took a fully subjective approach to the theory of risk measures, which he viewed as representations of the analysist's subjective perception of risk (see [**21**] and [**22**]). In Section 5, we notice that risk measures can be viewed as the result of an aggregation of experts' assessments of risk (Bayesian or not). This result is fully compatible with the subjective viewpoint. In fact, one can possibly argue that it is precisely the recognition of the subjectivity involved in the risk's assessments that justifies and motivates their aggregation. Viewing risk measures as the result of an aggregation has important consequences with regard to the properties that might be deemed desirable in a risk measure. For instance, one can give examples where each expert proposes the adoption of a sub-additive risk measure, but the aggregation results in a non-necessarily sub-additive risk measure. In fact, as we shall see, the sub-additivity of the aggregated measure corresponds to a pessimistic view about the experts' assessments (Corollary 5), which goes to show that sub-additivity, or lack thereof, might not be an issue about the value of diversification but rather one about the confidence in the experts' opinions.

The present paper belongs to the body of literature stemming from the work of Stone [**29**]. In [**29**], Stone proposed the *linear opinion pool* (which Stone attributed to Laplace) as a way of aggregating Bayesian experts. The linear opinion pool simply consists of fixing a system of weights on the set of experts, and of using it to average the experts' opinions. Recently, Crès, Gilboa and Vieille [**9**] (to which we refer for a thorough survey of the literature) have generalized Stone's proposal with regard to both the class of admissible experts and the weighing system: they allow for experts of the Maxmin Expected Utility type and for weights that are also of the maxmin type. Here, we provide a direct generalization of the linear opinion pool whereby capacities replace probability measures. Our class also extends that of Crès et al. as neither our generalized averages need be of the "cautious" type nor our experts need be of the maxmin type. In fact, we allow for the simultaneous presence of experts who may display very different attitudes toward uncertainty.

There are two other strands in literature that relate to the content of the present paper. The first is the literature on the social aggregation of individual preferences. At the structural level, the main difference between the literature on social aggregation and that on aggregation of experts resides in the fact that the former has to deal with individuals who possibly display different preferences on outcomes, a problem that is altogether absent in the case of aggregation of experts. Other differences include the presence of axioms specific to the fact that the entity aggregating the individual preferences is a society. At the technical level, the literature on social aggregation has focused mostly on the linear aggregation of "linear" individuals (that is, individuals whose preferences are represented by linear/affine functionals such as vN-M utilities or Expected utilities); this is the case, for instance, for Harsanyi [**17**], Coulhon-Mongin [**8**], Blackorby et al. [**6**] and Dhillon-Mertens [**10**]. Only recently has this literature moved to consider possibly non-linear aggregations of linear individuals with the works of Sprumont [**28**] and Alon and Gayer [**1**] or of possibly non-linear individuals with the work of Nascimento [**25**]. All these papers, however, always restrict to aggregations of the "pessimistic" type. Within the limits of the structural difference mentioned above, our paper contributes to this literature a number of methods concerning the non-linear aggregation of non-linear individuals. The other literature, whose results we use in Section 4, is that on Ordered Weighted Averaging (OWA) operators, which has mainly developed in theoretical computer science. OWA operators were introduced by Yager [**30**] and have been generalized in various ways (see [**23**] and [**4**]). When interpreted as aggregators of experts' opinions, they encode the additional requirement of equal reliability of the experts, a requirement that takes the form of an "anonimity" axiom in the literature on social aggregation.

2. Aggregating Bayesian Experts

We consider an individual that has to choose among the alternatives in a set $\mathcal{A}$. The outcome associated to each choice depends on the realization of a state of the world $s \in S$. In the process of reaching his decision, the individual is assisted by a set $\mathcal{E}$ of experts. We begin with the case where each expert $e \in \mathcal{E}$ is Bayesian, in the sense that she reduces all the uncertainty to risk: e's view of uncertainty is represented by a probability measure P_e on S. Let $\varphi \in \mathcal{A}$ and let u denote the individual's utility on outcomes. Then, according to expert e, the correct evaluation of alternative

φ is given by the value $\int_S u(\varphi)dP_e$. Different experts have different probabilities on S, which can be interpreted as the result of differences of information across experts or simply as differences of opinions. The problem of our individual is that of aggregating these different views in order to obtain a ranking of the alternatives.

This description gives rise to the following formal model. An alternative $\varphi \in \mathcal{A}$ is a mapping $\varphi : S \longrightarrow X$, where S is a set of states and X is a set of outcomes. We assume that X is a mixture space and that our individual has a preference over outcomes represented by a von Neumann-Morgenstern utility u on X. The set of states is endowed with the coarsest σ-algebra, Σ, which makes all the mappings of the form $u \circ \varphi$ measurable, where φ ranges over the set of alternatives. The set of all mappings of the form $u \circ \varphi$ is thus a subset of $B(\Sigma)$, the Banach space (sup-norm) of bounded Σ-measurable mappings. The view of each expert $e \in \mathcal{E}$ is represented by a (finitely additive) probability measure P_e on Σ. We identify each expert $e \in \mathcal{E}$ with his own view P_e. In this way, the set of experts $\mathcal{E}$ is identified to a subset of the norm dual, $ba(\Sigma)$, of $B(\Sigma)$ and this subset is endowed with the weak*-topology produced by the duality $(ba(\Sigma), B(\Sigma))$. Due to the presence of experts, each alternative $\varphi \in \mathcal{A}$ is naturally associated to a collection of evaluations $\left\{\int u(\varphi)dP_e\right\}_{e\in\mathcal{E}}$, one for each expert. This collection can be viewed as a mapping $E_\varphi : \{P_e\}_{e\in\mathcal{E}} \longrightarrow \mathbb{R}$ where E_φ is defined by $E_\varphi(P_e) = \int u(\varphi)dP_e$, and we will denote by κ the association

$$\kappa : \varphi \longmapsto E_\varphi$$

that is, κ associates an alternative $\varphi \in \mathcal{A}$ with the corresponding set of experts' evaluations E_φ. We say that a preference relation $\succsim$ over $\mathcal{A}$ is an *Aggregation of Bayesian Experts' Opinions* (ABEO) if it depends only on the experts' evaluations. The next definition expresses this formally.

Definition 1. *A preference relation $\succsim$ over $\mathcal{A}$ is an Aggregation of Bayesian Experts' Opinions (ABEO) if the functional that represents it factors as $I = V \circ \kappa$.*

We are going to focus exclusively on functionals I (equivalently, on functionals V) which are $\mathbb{R}$-valued, hence on preferences that are Archimedean. Axioms guaranteeing the existence of such representations are well-known, and we refer the reader to the relevant literature.

Before proceeding, it is worth noting that this set-up can accommodate two protocols that are, in principle, different. In one, for each alternative $\varphi \in \mathcal{A}$, each expert informs the decision maker of what is, in her opinion, the correct evaluation of φ, and then the decision maker aggregates these evaluations. In this protocol, the decision maker need not be explicitly informed about the expert's probability on S. In the other, at the outset each expert informs the decision maker of her probability on S, and then for each $\varphi \in \mathcal{A}$ the decision maker uses these probabilities to compute the evaluation that each expert would have given for that alternative. Under the assumption, customary in the literature, that the set of alternatives is rich enough so that the space of mappings of the form $u \circ \varphi$, $\varphi \in \mathcal{A}$, generates the whole $B(\Sigma)$, these two protocols are clearly equivalent. While we too make this assumption, thus rendering the distinction moot, it is worth keeping in mind that the underlying protocol can be of either type.

We shall see momentarily that an ABEO preference relation $\succsim$ over $\mathcal{A}$ is a sort of generalized weighing of experts' evaluations if and only if two mild conditions are satisfied. The first is a regularity condition:

Assumption 0: *The functional V is equivariant with respect to positive affine transformations.*

In the Bayesian case, this means that for all $\lambda \geq 0$, $\beta \in \mathbb{R}$ and for all $E_\varphi \in \kappa(B(\Sigma))$ we must have $V(\lambda E_\varphi + \beta \mathbf{1}) = \lambda V(E_\varphi) + \beta V(\mathbf{1})$, where $\mathbf{1}$ denotes the function identically equal to 1 on its domain.[1] Assumption 0 is very natural, probably almost mandatory, in our setting. It is motivated by the assumption that the decision maker's utility on the outcome space X is of the von Neumann-Morgenstern type. Thus, Assumption 0 is simply the condition that the preference does not depend on the choice of the representing utility on outcomes.

The second is a unanimity condition that is customarily imposed in settings like ours (see for instance [**9**]), and is probably the minimal necessary requirement that enables us to talk about aggregating experts' opinion. It states that if all experts (when using the decision maker's utility on consequences) agree on the ranking of two alternatives, then the decision maker adopts this ranking. In order not to duplicate our assumptions, we state the condition by using a terminology more general than what is necessary in the Bayesian case. For $\varphi, \psi \in \mathcal{A}$, let us stipulate that $\varphi \succsim_e \psi$ means that "by using the decision maker's utility u on consequences, expert $e \in \mathcal{E}$ evaluates that φ should be preferred to ψ". Of course, in the Bayesian case, $\varphi \succsim_e \psi$ means precisely that $\int u(\varphi) dP_e \geq \int u(\psi) dP_e$.

Axiom 1 (Unanimity). *Let $\varphi, \psi \in \mathcal{A}$. If $\varphi \succsim_e \psi$ for all $e \in \mathcal{E}$, then $\varphi \succsim \psi$.*

Theorem 1 below states that an (archimedean) ABEO preference relation satisfies Assumption 0 and Unanimity if and only if it is a "weighted average" of experts' opinions, where the weighing is represented by a capacity. Let us denote by $\tilde{\mathcal{C}}$ the closed convex hull of the set $\{P_e\}_{e\in\mathcal{E}} \subset ba(\Sigma)$, $\tilde{\mathcal{C}} = \overline{co}\{P_e\}_{e\in\mathcal{E}}$, and let $\tilde{\kappa} : \varphi \longmapsto \tilde{E}_\varphi$, where $\tilde{E}_\varphi : \tilde{\mathcal{C}} \longrightarrow \mathbb{R}$ is defined by $\tilde{E}_\varphi(P) = \int u(\varphi) dP$; that is, $\tilde{E}_\varphi$ is the extension of E_φ to $\tilde{\mathcal{C}}$ and is formally defined in exactly the same way.

Theorem 1. *An ABEO preference relation $\succsim$ over $\mathcal{A}$ satisfies Assumption 0 and AXIOM 1 iff*

$$\varphi \succsim \psi \qquad \Longleftrightarrow \qquad \int_{\tilde{\mathcal{C}}} \tilde{E}_\varphi d\Gamma \geq \int_{\tilde{\mathcal{C}}} \tilde{E}_\psi d\Gamma$$

where Γ is a capacity on the Borel sets of $\tilde{\mathcal{C}} = \overline{co}\{P_e\}_{e\in\mathcal{E}}$ and the integral is taken in the sense of Choquet.

Thus, Theorem 1 (whose proof, like all proofs in this paper, is in the Appendix) shows that a preference $\succsim$ satisfying Assumption 0 and AXIOM 1 admits a representation by means of the functional

$$I(\varphi) = \int_{\tilde{\mathcal{C}}} \tilde{E}_\varphi d\Gamma = \int_{\tilde{\mathcal{C}}} \int_S u(\varphi) dP d\Gamma$$

[1] Because of the richness assumption in the text, $E_\varphi \in \kappa(B(\Sigma))$ implies $\lambda E_\varphi + \beta \mathbf{1} \in \kappa(B(\Sigma))$.

As observed in the next corollary, the preferences of Theorem 1 are direct extensions of the classical *linear opinion pools* of Stone [**29**], which obtain precisely when the capacity Γ in Theorem 1 is a probability measure. In this case, Γ can be replaced by its barycenter, which corresponds to the "average" expert or, equivalently, to a fixed systems of weights on the experts.

COROLLARY 1. *An ABEO preference relation $\succsim$ over $\mathcal{A}$ is a linear opinion pool iff the capacity Γ in Theorem 1 is a probability measure. In particular, if the set $\mathcal{E}$ is finite, $\mathcal{E} = \{e_1, e_2, ..., e_n\}$, there exists a system of weights $\lambda = (\lambda_e)_{e\in\mathcal{E}}$ such that*

$$I(\varphi) = \sum_{e=1}^{n} \lambda_e \int_S u(\varphi) dP_e$$

In the general case, preferences satisfying the assumptions of Theorem 1 can be seen as non-linear pools as the experts' opinions are weighed in a non-additive way. We shall see momentarily that this non-additivity can be interpreted as the result of varying the weights on the experts depending on the alternative to be evaluated, possibly to account for the fact that an expert's reliability may not be the same across all alternatives.

2.1. Interpreting non-linear pools; quantile functions. For the preference $\succsim$ to be an ABEO, the representing functional must factor as $I = V \circ \kappa$. In the Appendix, we show that one can always represent such an I in the form $I = V \circ \kappa = \tilde{V} \circ \tilde{\kappa}$, where $\tilde{\kappa}$ is the extension of κ introduced above. Given this, it is pretty clear that an ABEO preference satisfies the axiom of unanimity if and only if the functional $\tilde{V}$ (equivalently, V) is monotone on its domain, that is $\tilde{E}_\varphi \geq \tilde{E}_\psi$ (pointwise) implies $\tilde{V}(\tilde{E}_\varphi) \geq \tilde{V}(\tilde{E}_\psi)$. In order to gain more insights into the non-linear pools of Theorem 1, let us begin by observing that any monotone $\tilde{V} : \tilde{\kappa}(B(\Sigma)) \longrightarrow \mathbb{R}$ can be written in the form

$$\tilde{V}(\tilde{E}_\varphi) = \tilde{\alpha}(\tilde{E}_\varphi) \inf_{\tilde{\mathcal{C}}} \tilde{E}_\varphi + (1 - \tilde{\alpha}(\tilde{E}_\varphi)) \sup_{\tilde{\mathcal{C}}} \tilde{E}_\varphi \tag{2.1}$$

where $\tilde{\alpha}(\tilde{E}_\varphi) \in [0,1]$. The values of the coefficients $\tilde{\alpha}(\tilde{E}_\varphi)$ can be interpreted as indicators of the decision maker's confidence in the experts' opinions: higher values indicate that the decision maker leans toward the more conservative evaluations possibly because he is not very confident in the experts' assessments, and an analogous interpretation can be given for lower values of $\tilde{\alpha}(\tilde{E}_\varphi)$. When $\tilde{E}_\varphi$ is a constant function (that is, all the experts agree on the evaluation of φ), this coefficient is evidently irrelevant; when $\tilde{E}_\varphi$ is non-constant, one can solve (2.1) for $\tilde{\alpha}(\tilde{E}_\varphi)$ and obtain

$$\tilde{\alpha}(\tilde{E}_\varphi) = \frac{\tilde{V}(\tilde{E}_\varphi) - \sup \tilde{E}_\varphi}{\inf \tilde{E}_\varphi - \sup \tilde{E}_\varphi}$$

When the opinion pool is linear, the capacity Γ in Theorem 1 is a measure and for any $\varphi \in \mathcal{A}$

$$\tilde{V}(\tilde{E}_\varphi) = \tilde{E}_\varphi(P_\Gamma)$$

where P_Γ denotes the barycenter of Γ. This shows that a linear opinion pool consists of choosing, once and for all, a point in $\tilde{\mathcal{C}}$, and of using this point to evaluate the alternatives. This point, that can be interpreted as a representative opinion of the pool, is chosen independently of the act being evaluated, and thus corresponds to the operation of putting fixed weights on the experts. In contrast, a non-linear pool obtains when the representative opinion (the point at which $\tilde{E}_\varphi$ is evaluated) is allowed to vary with φ: the decision maker leans, say, toward conservative evaluations in correspondence to a certain φ while he leans toward optimistic evaluations in correspondence to a certain $\psi \neq \varphi$. This might reflect, for instance, different levels of confidence in the experts' reliability of evaluating certain alternatives.

An important example of non-linear pools is provided by quantile functions. Recall that, for Λ a probability measure on the Borel sets of $\tilde{\mathcal{C}}$, a lower quantile with respect to Λ is a functional $\tilde{Q} : \tilde{\kappa}(B(\Sigma)) \longrightarrow \mathbb{R}$ of the type

$$\tilde{Q}(\tilde{\kappa}(\varphi)) = \inf\{x \mid \Lambda(\{P : \tilde{\kappa}(\varphi)(P) \geq x\}) \leq \beta\}$$

where β is a fixed number in $[0, 1)$. Similarly, an upper quantile is a functional of the type

$$\tilde{Q}(\tilde{\kappa}(\varphi)) = \sup\{x \mid \Lambda(\{P : \tilde{\kappa}(\varphi)(P) \geq x\}) \geq \beta\}$$

A functional $\tilde{Q}$ is simply a quantile function if it is either a lower quantile or an upper quantile. As quantile functions can be represented by means of Choquet integrals (see [**7**]), Theorem 1 guarantees that quantile functions give rise to ABEO preferences that satisfy Assumption 0 and Unanimity. An example of a quantile-type aggregation is a rule that assigns value 100 to alternative φ if 95% of the experts evaluate that φ is worth at least 100. Aggregation rules of this type, that is rules that discard extreme opinions and aggregate the remaining by either a simple average or by taking a max or a min, are common in practice and are clearly incompatible with the linear opinion pool. We will encounter a few examples later in the paper.

2.2. Preference for Compromise. A sub-class of ABEO preferences that is important for applications, especially those that we will discuss later, is that of preferences which are represented by super-additive functionals. These are precisely those preferences that satisfy an additional axiom, Preference for Compromise, originally introduced by Sprumont in [**28**]. Needless to say, the axiom of Preference for Compromise comes implicitly with a dual companion that characterizes those preferences which are representable by means of sub-additive functionals. By using the same convention as in AXIOM 1, the axiom of Preference for Compromise reads as follows:

AXIOM 2 (Preference for Compromise (Sprumont [**28**])). *Let $\mathcal{F}$ be a non-empty, strict subset of $\mathcal{E}$. If $\varphi, \psi, \chi \in \mathcal{A}$ are such that*

$$\varphi \succsim_e \chi \succsim_e \psi \qquad \textit{for all} \qquad e \in \mathcal{F}$$

$$\psi \succsim_e \chi \succsim_e \varphi \qquad \textit{for all} \qquad e \in \mathcal{E} \setminus \mathcal{F}$$

then either $\chi \succsim \varphi$ or $\chi \succsim \psi$.

As its denomination transparently suggests, alternative χ can be viewed as a compromise between φ and ψ, and the axiom demands that the decision maker would always find it as good as the worst between φ and ψ.

PROPOSITION 1. *Let $\succsim$ be an ABEO preference relation which satisfies the conditions of Theorem 1 as well as the axiom Preference for Compromise. Then, there exists a unique, convex,* *weak*-compact set of probability measures* $\mathcal{CE} \subset ba(\Sigma)$ *such that*

$$\varphi \succsim \psi \qquad iff \qquad \min_{P \in \mathcal{CE}} \int_S u(\varphi) dP \geq \min_{P \in \mathcal{CE}} \int_S u(\psi) dP$$

Moreover, the inclusion $\mathcal{CE} \subset \tilde{\mathcal{C}} = \overline{co}\,\{P_e\}_{e \in \mathcal{E}}$ *holds.*

It is probably useful to clarify that Proposition 1 does not say that, for each alternative φ, the decision maker selects the worst among the experts' evaluations. This would be the case only if the equality $\mathcal{CE} = \overline{co}\,\{P_e\}$ obtains. We will characterize this class in Section 4. In general, however, the inclusion $\mathcal{CE} \subset \overline{co}\,\{P_e\}$ can be strict, and an ABEO preference that displays Preference for Compromise can be thought of as the result of a two-stage operation. In the first stage, one uses a collection of linear opinion pools, whereby replacing the set of experts $\mathcal{E}$ with this set of linear opinion pools. This is the set $\mathcal{CE}$ that can be thought of as a set of "experts" whose opinions are closer to each other than those of the original experts. Then, in the second stage, each alternative is evaluated by taking the worst evaluation of these more "moderate" experts. This is further clarified in the next corollary which, for simplicity, is stated for the case of a finite set of experts.

Corollary 2. *Let* $\mathcal{E}$ *be finite,* $\mathcal{E} = \{e_1, e_2, ..., e_n\}$*. An ABEO preference relation* $\succsim$ *satisfies the conditions of Theorem 1 and the axiom Preference for Compromise iff there exists system of weights* $\Lambda \subseteq \Delta(\mathbb{R}^n)$ *(*$\Delta(\mathbb{R}^n)$ *the simplex in* $\mathbb{R}^n$*) such that* $\succsim$ *is represented by the functional*

$$I(\varphi) = \min_{\lambda \in \Lambda} \sum_{e=1}^{n} \lambda_e \int_S u(\varphi) dP_e$$

We conclude this part by stating the axiom dual to Preference for Compromise and the associated representation:

Axiom 3 (Preference for Polarization). *Under the same conditions as in AXIOM 2, both relations* $\varphi \succsim \chi$ *and* $\psi \succsim \chi$ *hold.*

Corollary 3. *Let* $\succsim$ *be an ABEO preference relation which satisfies the conditions of Theorem 1 as well as the axiom Preference for Polarization. Then, there exists a unique, convex, weak*-compact set of probability measures* $\mathcal{CE} \subset ba(\Sigma)$ *such that*

$$\varphi \succsim \psi \qquad iff \qquad \max_{P \in \mathcal{CE}} \int_S u(\varphi) dP \geq \max_{P \in \mathcal{CE}} \int_S u(\psi) dP$$

Moreover, the inclusion $\mathcal{CE} \subset \tilde{\mathcal{C}} = \overline{co}\,\{P_e\}_{e \in \mathcal{E}}$ *holds.*

2.3. Summary on the aggregation of Bayesian experts. When all experts are Bayesian, an aggregation of the experts' opinions satisfies (i) equivariance with respect to positive affine transformations and (ii) Unanimity iff it can be expressed as a non-linear weighing of experts. This weighing is represented by capacity, which takes the place of the probability used in the linear opinion pool of Stone-Laplace. This generalization provides us with enough flexibility to accommodate many aggregation procedures used in practice such as those that discard extreme opinions.

A special class of our aggregation procedures consists of those that result from always finding a compromise (resp. always choosing an extreme) among the experts' opinions. These procedures, which can also be interpreted as revealing a certain lack of confidence (resp. over-confidence) in the experts' assessments, are represented by a maxmin (resp. maxmax) type aggregator.

3. Non-Bayesian experts

We now extend our study by allowing for experts that are not necessarily Bayesian. The first step is to observe what this extension entails in terms of the objects of the previous section. In the Bayesian case, each expert $e \in \mathcal{E}$ incorporated the decision maker's utility function u on X and gave the evaluation $\int u(\varphi)dP_e$ for the alternative $\varphi \in \mathcal{A}$. The structure is similar in the general case but with the difference that the expert's evaluation functional (which is still obtained by incorporating the decision maker's utility function on outcomes) need not be an expected utility functional. By letting $\mathcal{I}$ denote the set of experts, we denote expert i's evaluation functional, $i \in \mathcal{I}$, by $J_i(\cdot)$. Just like in the Bayesian case, each alternative $\varphi \in \mathcal{A}$ is then naturally associated to the collection of experts' evaluations, $\{J_i(\varphi)\}_{i\in\mathcal{I}}$, which we can view as a mapping $F_\varphi : \mathcal{I} \longrightarrow \mathbb{R}$, defined by $i \longmapsto F_\varphi(i) = J_i(\varphi)$. We denote by γ the association $\gamma : \varphi \longmapsto F_\varphi$. The next definition extends Definition 1 of Section 2, and is motivated in exactly the same way. Just like in the Bayesian case, we restrict attention to functionals I (equivalently, V) that are real-valued.

DEFINITION 2. *A preference relation $\succsim$ over $\mathcal{A}$ is an Aggregation of Experts' Opinions (AEO) if the functional that represents it factors as $I = V \circ \gamma$.*

We are going to allow for a large variety of non-Bayesian experts: our experts could be of the maxmin type, the maxmax type or be a variable convex combination of maxmin and maxmax. The exact restrictions on the class of admissible experts are stated in the next definition.

DEFINITION 3. *We say that expert i is of class $\mathcal{NB}$ if the corresponding functional J_i satisfies:*
(1) Monotonicity: $u \circ \varphi(s) \geq u \circ \psi(s)$ for all $s \in S$ implies $J_i(\varphi) \geq J_i(\psi)$;
(2) C-independence: For all $\alpha \in [0, 1]$ and for all $\varphi \in \mathcal{A}$, $J_i(\alpha\varphi + (1-\alpha)\mathbf{1}) = \alpha J_i(\varphi) + (1-\alpha)$.

When an expert i is of class $\mathcal{NB}$, the resulting preference on $\mathcal{A}$ (the one that the expert suggests to the decision maker) is Invariant Bi-separable and the corresponding functional J_i is a functional of the multiple-prior type (see [**14**]), that is J_i is associated to a set of priors $\mathcal{C}_i$. Proposition 2 below states that an aggregation of experts of class $\mathcal{NB}$ can be thought of as the result of putting together all the experts' priors and then weighing them in a non-additive way. Let $\mathcal{K} = \overline{co}\,\{\cup_{i\in\mathcal{I}}\mathcal{C}_i\}$ and, for every $\varphi \in \mathcal{A}$, let $\tilde{E}_\varphi : \mathcal{K} \longrightarrow \mathbb{R}$ be defined by $\tilde{E}_\varphi(P) = \int u \circ \varphi dP$.

PROPOSITION 2. *Let all experts be of class $\mathcal{NB}$ and let $\succsim$ be an AEO that satisfies Assumption 0 and AXIOM 1. Then, for every $\varphi, \psi \in \mathcal{A}$*

$$\varphi \succsim \psi \qquad \Longleftrightarrow \qquad \int_{\mathcal{K}} \tilde{E}_\varphi d\Gamma \geq \int_{\mathcal{K}} \tilde{E}_\psi d\Gamma$$

where Γ is a capacity on the Borel sets of $\mathcal{K}$ and the integral is taken in the sense of Choquet.

We should like to stress that, as long as all experts are of class $\mathcal{NB}$, the experts of Proposition 2 need not belong to the same sub-class: for instance, within the group of experts of Proposition 2, we could have some experts of the maxmin type, some of the maxmax type and some that belong to neither of these classes. This shows the wide range of applicability of Proposition 2: Not only does it allow for differences across the opinions of (possibly) non-Bayesian experts, but it also allows for different experts' attitudes toward the uncertainty.

In general, the representation of Proposition 2 is not unique as there might be other pairs, $(\tilde{\mathcal{C}}, \Gamma')$, of sets of priors and capacities yielding representations of the type

$$I(\varphi) = \int_{\tilde{\mathcal{C}}} \tilde{E}_{\varphi} d\Gamma'$$

Once again, these representations can be interpreted as the result of (1) treating each expert's prior as a Bayesian expert; (2) applying a collection of linear opinion pools over these priors, thus obtaining a set $\tilde{\mathcal{C}}$ of more "moderate" priors; and, finally (3) weighing these priors in a non-additive way. The extreme points of $\tilde{\mathcal{C}}$ can then be interpreted as a set of fictitious Bayesian experts (see the discussion in the Appendix following the proof of Proposition 2).

Possibly, the most important implication of Proposition 2 as well as of the alternative representations just discussed is as follows: Any aggregation of experts of class $\mathcal{NB}$, which satisfies Unanimity and equivariance with respect to positive affine transformations, can be thought of as an aggregation of Bayesian experts which satisfies the same conditions; equivalently, any outcome obtained by aggregating experts of class $\mathcal{NB}$ can be replicated by aggregating Bayesian experts. For a concrete application, let us observe that many economic indicators, e.g. some prices (see, for instance, the LIBOR example of Section 4), are the outcome of an aggregation of experts' opinions. Then, Proposition 2 says that treating those opinions as if they were Bayesian entails no loss of generality.

3.1. MEU experts.

In this subsection, we specialize our setting to that studied in Crès et al. [**9**], where the experts belong to the Maxmin Expected Utility (MEU) class; that is, for each $i \in \mathcal{I}$, the corresponding evaluation functional is of the form

$$J_i(\varphi) = \min_{p \in \mathcal{C}_i} \int_S u(\varphi) dp$$

where $\mathcal{C}_i \subset ba(\Sigma)$ is a set of probability charges. We are going to present two propositions. The first, Proposition 3, completes the parallel with the Bayesian case: under the axiom of Preference for Compromise one obtains a representation of the same form as that encountered in the Bayesian case. Thus, we can offer the same interpretation as above: each experts communicates to the decision maker her set of priors; then, the decision maker collects all the priors for all experts and weighs them in a "conservative" way. The second, Proposition 4, derives the same representation as in Crès et al. [**9**].

PROPOSITION 3. *Assume that all experts are of class MEU, and let $\succsim$ be an AEO satisfying Assumption 0, AXIOM 1 and AXIOM 2. Then,*

(1) There exists a set $\mathcal{E}$ of (fictitious) Bayesian experts such that for every $\varphi, \psi \in \mathcal{A}$

$$\varphi \succsim \psi \qquad \Longleftrightarrow \qquad \min_{P\in\tilde{\mathcal{C}}} \int_S u(\varphi)dP \geq \min_{P\in\tilde{\mathcal{C}}} \int_S u(\psi)dP$$

where $\tilde{\mathcal{C}} = \overline{co}\left\{P_e\right\}_{e\in\mathcal{E}}$. Moreover, $\tilde{\mathcal{C}} \subseteq \mathcal{K} = \overline{co}\left\{\bigcup_{i\in\mathcal{I}} \mathcal{C}_i\right\}$; equivalently,

(2) There exists a set Λ of measures on (the Borel sets of) $\mathcal{K}$ such that

$$\varphi \succsim \psi \qquad \Longleftrightarrow \qquad \min_{\lambda\in\Lambda} \int_{\mathcal{K}} \int_S u(\varphi)dPd\lambda \geq \min_{\lambda\in\Lambda} \int_{\mathcal{K}} \int_S u(\psi)dPd\lambda$$

The next Proposition, both for simplicity and for a direct comparison to Crès et al. [**9**], is stated for the case of a finite set of experts.

PROPOSITION 4. *Let $\mathcal{I} = \{1, 2, ..., n\}$ denote set of experts, and assume that all experts are of class MEU. Let $\succsim$ be an AEO satisfying Assumption 0, AXIOM 1 and AXIOM 2. Then, there is a collection Λ of weighing systems, $\lambda = (\lambda_1, \lambda_2, ..., \lambda_n)$, such that, for every $\varphi \in \mathcal{A}$, $\succsim$ is represented by*

$$I(\varphi) = \min_{\lambda\in\Lambda} \sum_{i=1}^{n} \lambda_i \min_{P\in\mathcal{C}_i} \int_S u(\varphi)dP = \min_{\lambda\in\Lambda} \sum_{i=1}^{n} \lambda_i J_i(\varphi)$$

We conclude this section with an easy, yet useful, observation. An aggregation of maxmin experts need not give rise to a concave representation functional (if it doesn't satisfy Preference for Compromise). That is, even if all experts have a cautious attitude toward uncertainty, the aggregation need not display a cautious attitude. We will return to this point in the section on Risk Measures.

4. Examples; equal reliability

In this section, we consider another subclass of the aggregation procedures of Proposition 2. This is identified by the additional assumption that the decision maker considers all experts equally reliable. For the most part, the results of this section are known, mostly from the literature on OWA (Ordered Weighted Averaging) operators, and are used only as examples. For a finite set of experts, $\mathcal{I} = \{i_1, i_2, ..., i_n\}$, the assumption of equal reliability reads as follows. Let Π denote the group of permutations of $\mathcal{I}$, let $\varphi \in \mathcal{A}$ and let F_φ be the mapping of experts' evaluations defined at the beginning of Section 3.

AXIOM 4 (Equal Reliability). *Let $\succsim$ be an AEO preference on $\mathcal{A}$. We say that the experts are equally reliable if for every $\varphi \in \mathcal{A}$ and every permutation $\pi \in \Pi$: $V(F_\varphi) = I(\varphi) = V(F_\varphi \circ \pi)$.*

Thus, an aggregation protocol satisfies Equal Reliability if no expert is weighed more than another. Clearly, there is only one linear opinion pool that satisfies Equal Reliability, which obtains when each and every expert is assigned the same weight. In contrast, we shall see that there are infinitely many non-linear pools that satisfy Equal Reliability. As AEO preferences are represented by means of capacities, it is not surprising that the key concept to characterize the aggregations of equally reliable experts is that of permutation invariant capacity.

DEFINITION 4. *A capacity ν on $2^{\mathcal{I}}$ is symmetric or permutation invariant if for every permutation π of S and every $E \in 2^{\mathcal{I}}$*

$$\nu(A) = \nu(\pi(E))$$

It is immediate to see that a capacity is permutation invariant iff its value on a set E depends only on the cardinality of the set. Thus, a normalized (i.e., $\nu(\mathcal{I}) = 1$) permutation invariant capacity is completely determined by the $n+1$ values $\nu = (\nu(0), \nu(1), \nu(2), ..., \nu(n))$, where $\nu(c)$ denotes the value of ν on a set of cardinality c. Conversely, any set of $n+1$ numbers such that $\{x_{j+1} \geq x_j \geq 0; x_0 = 0, x_n = 1\}$ determines a permutation invariant capacity on $2^{\mathcal{I}}$. Hence, the set of permutation invariant capacities is isomorphic to the simplex, $\Delta(\mathbb{R}^n)$, in $\mathbb{R}^n$ via the mapping $w = (w_1, w_2, ..., w_n) \longmapsto \nu_w$, where $w_j = \nu_w(j) - \nu_w(j-1)$.

In what follows, we will denote a generic permutation invariant capacity by ν_w, where w denotes the corresponding vector in $\Delta(\mathbb{R}^n)$. For short, we will also say that a Choquet integral is *symmetric* if it is taken with respect to a permutation invariant capacity. A symmetric Choquet integral, $\int F d\nu_w$, has automatically the property that $\int F d\nu_w = \int F \circ \pi d\nu_w$ for every permutation $\pi \in \Pi$. The first example shows that symmetric Choquet integrals give rise to equally reliable AEO preferences (thus, in particular, there are infinitely many such preferences).

EXAMPLE 1 (Symmetric Choquet Integrals). *Let $\mathcal{I} = \{i_1, i_2, ..., i_n\}$ denote the set of experts and let ν_w be a permutation invariant capacity on $2^{\mathcal{I}}$. Then, the preference $\succsim$ on $\mathcal{A}$ defined by*

$$\varphi \succsim \psi \qquad iff \qquad \int_{\mathcal{I}} F_\varphi d\nu_w \geq \int_{\mathcal{I}} F_\psi d\nu_w$$

is an AEO (if we assume that all experts are of class $\mathcal{NB}$) satisfying Assumption 0, AXIOM 1 and AXIOM 4.

Symmetric Choquet integrals are OWA operators (Yager [**30**]). This observation (originally due to Murofushi and Sugeno [**24**], see also Grabisch [**16**]) is especially useful because it tells us what are the aggregation protocols implied by symmetric Choquet integrals. As we shall see momentarily, many of these protocols are commonly used in practice.

EXAMPLE 2 (Ex. 1 continued: OWA operators, [**30**] and [**24**]). *By using the definition of the Choquet integral, the integrals $\int F_\varphi d\nu_w$ can be written in an equivalent way. Since $\mathcal{I}$ has n elements, $F_\varphi \in \mathbb{R}^n$ and has the form $F_\varphi = (f_1^\varphi, f_2^\varphi, ..., f_n^\varphi)$. Let $D^\varphi = (d_1^\varphi, d_2^\varphi, ..., d_n^\varphi)$ denote a non-increasing rearrangment of the vector F_φ. Then, for every $\varphi \in \mathcal{A}$*

$$\int_{\mathcal{I}} F_\varphi d\nu_w = \sum_{i=1}^{n} d_i^\varphi w_i = \langle D^\varphi, w \rangle \tag{4.1}$$

Operators $\mathbb{R}^n \longrightarrow \mathbb{R}$ that are defined by an equation like (4.1) are called OWA operators.

Thus a Choquet integral $\int F d\nu_w$ with respect to a permutation invariant capacity corresponds to the procedure of (i) fixing a system of weights w; (ii) for each vector of experts' evaluations $F \in \mathbb{R}^n$, F is rearranged in a non-increasing order; and (iii) the rearranged values are averaged with the weights given by w. This observation allows us to characterize those protocols which consist of always choosing the lowest among the experts' evaluations, and that we mentioned in passing following Corollary 2.

Example 3 (Always choosing the lowest/highest evaluation). *The protocol of always choosing the lowest among the experts' evaluation is an aggregation procedure that satisfies both Preference for Compromise and equal reliability. It is represented by the OWA operator corresponding to the system of weights* $m = (0, 0, ..., 0, 1)$. *The corresponding capacity,* ν_m*, is defined by* $\nu_m(E) = 0$ *for all* $E \in 2^{\mathcal{I}} \diagdown \mathcal{I}$ *and* $\nu_m(\mathcal{I}) = 1$. *If we replace Preference for Compromise with Preference for Polarization, we obtain the protocol that consists of always choosing the highest among the experts' evaluations. This corresponds to the OWA operator defined by the system of weights* $M = (1, 0, ..., 0, 0)$. *The capacity* ν_M *is the conjugate capacity of* ν_m*, that is* $\nu_M(E) = \nu_m(\mathcal{I}) - \nu_m(E^c)$*, for all* $E \in 2^{\mathcal{I}}$*, where* E^c *denotes the complement of* E.

Example 4 (Discarding the highest and the lowest). *Another protocol that is commonly used in practice consists of discarding both the highest and lowest evaluation and, then, of weighing the remaing evaluations equally. This protocol also correspond to an OWA operator, defined by the system of weights* $w = (0, 1/n - 2, ..., 1/n - 2, 0)$.

Concrete instances of this protocol are encountered in sports like ice skating and diving where an athlete's score is determined precisely by discarding both the highest and lowest among the judges' evaluations and by averaging the remaining evaluations equally. Another set of examples is provided by many rating systems like, for instance, TripAdvisor's

Example 5 (LIBOR). *Another concrete instance of the above protocol used to be the determination of the London Interbank Offered Rate (LIBOR). As reported by The Economist (July 7, 2012), the LIBOR used to be determined by soliciting 18 estimations, then removing the highest 4 and the lowest 4, and then averaging the remaining 10.*

When the set of experts is an interval of the real line or, more generally, a subset of a finite dimensional real vector space, the notion of equal reliability can be readily extended. For, in such cases, we can refer to the *uniform* measure on the set of experts, and replace the notion of permutation of a finite set with that of transformation that preserves the uniform measure. The next definition is due to Kadane and Wasserman [**19**]:

Definition 5 (see Kadane and Wasserman [**19**]). *Let* $([0, 1], \mathcal{L}, \lambda)$ *denote the unit interval along with the usual Lebesgue measure. A capacity* ν *on* $\mathcal{L}$ *is symmetric if* $\lambda(A) = \lambda(B) \Longrightarrow \nu(A) = \nu(B)$.

Just like in the case of permutations, a Choquet integral which is taken with respect to a symmetric (in the sense of the above definition) capacity ν has the property that $\int F d\nu = \int F \circ \tau d\nu$

for every λ-preserving transformation τ of the interval $[0,1]$. By means of this observation, we then have an analog of example 1 (the trivial proof is omitted):

PROPOSITION 5. *Let $\mathcal{I}=([0,1],\mathcal{L},\lambda)$ and assume that all experts are of class $\mathcal{NB}$. If ν is symmetric capacity on $\mathcal{L}$, then the preference $\succsim$ on $\mathcal{A}$ defined by*

$$\varphi \succsim \psi \qquad iff \qquad \int_{\mathcal{I}} F_{\varphi} d\nu \geq \int_{\mathcal{I}} F_{\psi} d\nu$$

is an AEO satisfying Assumption 0, AXIOM 1 and AXIOM 4.

We have seen that OWA operators provide us with important examples of aggregation procedures that satisfy Assumption 0, AXIOM 1 and AXIOM 4. Not all these procedures, however, are representable by OWA operators. The next example shows that the combination of two OWA operators – possibly corresponding to two different sets of experts – yields a protocol satisfying Assumption 0, AXIOM 1 and AXIOM 4.

EXAMPLE 6 (Two sources). *Suppose that the decision maker distinguishes between two sets of $\mathcal{NB}$ experts, $\mathcal{I}_1$ and $\mathcal{I}_2$ (both finite), possibly because he knows that the experts in one set have access to a source of information that is not accessible by the experts in the other set. The decision maker aggregates the opinions of the experts in the set $\mathcal{I}_1$ ($\mathcal{I}_2$) by using an OWA operator with system of weights w_1 (w_2). This yields the functionals I_{w_1} and I_{w_2}. Then, the decision maker evaluates alternative $\varphi \in \mathcal{A}$ by using the functional*

$$I(\varphi) = \max\{I_{w_1}(\varphi), I_{w_2}(\varphi)\}$$

The resulting preference is an AEO satisfying Assumption 0, AXIOM 1 and AXIOM 4 but does not belong to the class of preferences representable by OWA operators.

The full characterization of the class of protocols satisfying Assumption 0, AXIOM 1 and AXIOM 4 follows from a result of [**4**], which implies that (in the case of a finite set of experts) a general AEO satisfying Assumption 0, AXIOM 1 and AXIOM 4 is representable by a functional of the form

$$I(\varphi) = (1-\alpha(\varphi)) \min_{w\in W} I_{\bar{w}}(\varphi) + \alpha(\varphi) \max_{w\in W} I_w(\varphi)$$

where $\{I_w(\cdot)\}_{w\in W}$, $W \subseteq \Delta(\mathbb{R}^n)$, is a family of sub-additive OWA operators, $\{I_{\bar{w}}(\cdot)\}_{w\in W}$ is the family of conjugate OWA operators (i.e., defined by the conjugate capacities) and the function $\alpha(\cdot)$ is such that the resulting preference satisfies Assumption 0 and AXIOM 1 (see [**4**] for the explicit conditions).

5. An application to the theory of Risk Measures

Risk Measures are indicators of the variability associated with a portfolio or a company. Formally, a *risk* (a portfolio, a company, etc.) is modeled as a measurable mapping $f:(S,\Sigma) \longrightarrow \mathbb{R}$, and a *risk measure* is a functional $\rho : B(\Sigma) \longrightarrow \mathbb{R}$. As said in the Introduction, the idea of a risk measure as well as its use for regulatory purposes have been criticized on the grounds that the evaluation of a risk implicitly involves an estimate of the likelihood of all relevant events. For rare events, such as those associated with major crisis, this likelihood cannot be estimated (to any reasonable degree of precision) by using existing data and is, therefore, nothing other than a subjective assessment. Thus, in particular, to claim that the undertaking of a certain risk is justified by the fact that its value $\rho(f)$ is below a certain threshold is nothing more than stating a subjective opinion. As a consequence, regulatory measures based on the values taken by a risk measure (such a those of the Basel Accords that are based on the VaR) lack, generally speaking, any objective justification.

Risk measures are categorized according to their properties. A risk measure ρ is:

(1) Cash invariant if: $\forall f \in B(\Sigma)$ and $\forall \alpha \in \mathbb{R}, \quad \rho(f + \alpha\mathbf{1}) = \rho(f) - \alpha$;
(2) Positively homogeneous if: $\forall f \in B(\Sigma)$ and $\forall \lambda \geq 0, \quad \rho(\lambda f) = \lambda\rho(f)$;
(3) Monotone if: $f, g \in B(\Sigma) \quad and \quad f \leq g \quad \Longrightarrow \quad \rho(g) \leq \rho(f)$;
(4) Sub-additive if: $\forall f, g \in B(\Sigma), \quad \rho(f + g) \leq \rho(f) + \rho(g)$.

The first three properties are self-explanatory. The fourth recognizes the value of diversification: since diversification (weakly) reduces the variability of a portfolio, the risk of a diversified position must not exceed the sum of the risks of its components. Risk measures that satisfy these four properties are dubbed *coherent risk measures* (Artzner et al. [**5**]). The industry standard, the VaR, has been the target not only of the general criticism outlined above but also of the additional one stemming from the observation that the VaR is not sub-additive.

In this section, we look at risk measures from a different angle, which consists of viewing them as the result of an aggregation of experts' opinions. Having explicitly recognized the subjective nature of the assessments about the uncertainty as well as the possibility that these assessments might differ across experts (whatever the definition of expert), it makes sense to aggregate them and evaluate risks accordingly. From the viewpoint of the regulator, the aggregation procedure might provide the best surrogate for the lack of objectivity associated with a risk measure: the justifiability of a certain risk may be certified on the basis of the fact that all experts (or most experts, or most experts among those institutionally recognized, etc.) agree that that risk is justified.

DEFINITION 6. *We say that a risk measure ρ is an AEE (short for Aggregation of Experts' Evaluations) if there exists an AEO functional I such that for any $f \in B(\Sigma)$*

$$\rho(f) = I(-f)$$

From the results of the previous sections, we then obtain at once the following proposition[2] and the subsequent two corollaries.

PROPOSITION 6. *An AEE risk measure ρ satisfies properties (1), (2) and (3) above iff the corresponding AEO preference satisfies the conditions of Proposition 2 (or Theorem 1), that is*

$$\rho(f) = I(-f) = \int_{\tilde{\mathcal{C}}} -\tilde{\kappa}(f)d\Gamma \tag{5.1}$$

where $\tilde{\kappa}$ is defined as in Section 2.

[2]The necessity part in Proposition 6 was proven in [**3**], in a different context.

Corollary 4. *The VaR is an AEE risk measure.*

Corollary 5. *An AEE risk measure ρ is sub-additive iff the capacity Γ in (5.1) is submodular. In particular, an aggregation of coherent risk measures is coherent if it satisfies Preference for Polarization.*

Viewing risk measures as the result of an aggregation of experts' opinions has immediate implications in terms of the properties that might be deemed desirable in a risk measure as these depend not only on the experts' assessments but also on the aggregation procedure. For instance, the observation at the end of Subsection 3.1 can now be thought of as an example of aggregation of coherent risk measures (each representing an expert's opinion) which results in a non-coherent risk measure because the aggregation does not satisfy the axiom of Preference for Polarization.

According to this perspective, the lack of sub-additivity of a risk measure is not necessarily an issue about the value of diversification but possibly one about the rules of aggregation. Consequently, the desiderability of sub-additivity is no longer evident. To further elaborate on this point, let us recall the discussion of Section 2.1. By following that line of reasoning, one sees that the representations of the type (5.1) are of the form

$$\rho(f) = \tilde{\alpha}(-\tilde{\kappa}(f)) \inf_{\tilde{\mathcal{C}}}(-\tilde{\kappa}(f)) + (1 - \tilde{\alpha}(-\tilde{\kappa}(f))) \sup_{\tilde{\mathcal{C}}}(-\tilde{\kappa}(f)) \tag{5.2}$$

and, just like in Section 2.1, the coefficient $\tilde{\alpha}(-\tilde{\kappa}(f))$ may be interpreted as the decision maker's assessment of the experts' reliability in evaluating risk f. This shows that an aggregation that incorporates the information that the experts' reliability may be different depending on the risk being evaluated is one that may result in a non-coherent risk measure (because $\tilde{\alpha}(\cdot)$ is not constant). Conversely, any risk measure of the type (5.1) that produces a non-constant $\tilde{\alpha}(\cdot)$ in the representation (5.2) is suitable for this interpretation.

5.1. Dropping positive homogeneity. The above discussion has left out an important class of risk measures: the *convex risk measures* of Föllmer and Schied [**13**]. These are risk measures that, in addition to cash invariance and monotonicity, satisfy the property of convexity: $\forall f, g \in B(\Sigma)$, and for all $\lambda \in [0, 1]$

$$\rho(\lambda f + (1 - \lambda)g) \leq \lambda\rho(f) + (1 - \lambda)\rho(g)$$

but are not necessarily positively homogenous. In the absence of positive homogeneity, it is the property of convexity that expresses the value of diversification. Clearly, convex risk measures are not covered by Proposition 6 because the functional (5.1) is positively homogeneous.

Conceptually, however, there seems to be no impediment in extending our framework so to accommodate convex risk measures as well. Once this is done, the same observations made above would apply. We conclude the paper by conjecturing what the required extension of Proposition 2/Theorem 1 could look like. Föllmer and Schied [**13**] have shown that convex risk measures have the representation

$$\rho(f) = \sup_{P \in \tilde{\mathcal{C}}} [(-\tilde{\kappa}(f)) - \pi(P)]$$

where $\tilde{\mathcal{C}}$ is a unique, convex, weak*-compact set of probability measures, $\tilde{\kappa}$ is the same function defined above and π is a penalty function. It is easy to see that functionals $I : B(\Sigma) \longrightarrow \mathbb{R}$ of the type

$$I(f) = \int_{\tilde{\mathcal{C}}} [\tilde{\kappa}(f) - \Pi(P)] d\Gamma$$

where Γ is a capacity on the Borel sets of $\tilde{\mathcal{C}}$, define convex risk measures via the identity

$$\rho(f) = I(-f) = \int_{\tilde{\mathcal{C}}} [-\tilde{\kappa}(f) - \Pi(P)] d\Gamma$$

if Γ is submodular. This observation suggests that these functionals could constitute, in the non-submodular case, the appropriate generalizations of our functionals of Proposition 2/Theorem 1 and of the corresponding AEE risk measures. In order to obtain this extension, it is easy to see that one must replace translation invariance with a property of the Weak Certainty Independence type (see [**20**]), but we do not know if this property is also sufficient, along with monotonicity, in the non-submodular case.

Appendix: omitted proofs

Section 2

Proof of Theorem 1:

Let $\geq$ denote the ordering on $\kappa(B(\Sigma))$ defined by $E_\varphi \geq E_\psi$ iff $E_\varphi(P_e) \geq E_\psi(P_e)$ for all $e \in \mathcal{E}$. A functional $V : \kappa(B(\Sigma)) \longrightarrow \mathbb{R}$ is is *monotone* if for $E_\varphi, E_\psi \in \kappa(B(\Sigma))$

$$E_\varphi \geq E_\psi \qquad \Longrightarrow \qquad V(E_\varphi) \geq V(E_\psi)$$

Let us denote by $\tilde{\mathcal{C}}$ the closed convex hull of the set $\mathcal{C} = \{P_e\}_{e \in \mathcal{E}} \subset ba(\Sigma)$, $\tilde{\mathcal{C}} = \overline{co}\,\{P_e\}_{e \in \mathcal{E}}$, and let $\tilde{\kappa} : \varphi \longmapsto \tilde{E}_\varphi$, where $\tilde{E}_\varphi : \tilde{\mathcal{C}} \longrightarrow \mathbb{R}$ is defined by $\tilde{E}_\varphi(P) = \int u(\varphi) dP$.

Lemma 1. *Let $V : \kappa(B(\Sigma)) \longrightarrow \mathbb{R}$ be a functional which is monotone and equivariant with respect to positive affine transformations. Then, there exists a unique functional $\tilde{V} : \tilde{\kappa}(B(\Sigma)) \longrightarrow \mathbb{R}$ such that $V(E_\varphi) = \tilde{V}(\tilde{E}_\varphi)$, for all $\varphi \in B(\Sigma)$. Moreover, $\tilde{V}$ is monotone and equivariant with respect to positive affine transformations.*

Proof. By monotonicity and equivariance with respect to positive affine transformations, $(\sup_{\mathcal{C}} E_\varphi) V(\mathbf{1}) \geq V(E_\varphi) \geq (\inf_{\mathcal{C}} E_\varphi) V(\mathbf{1})$. Hence, for each E_φ there exists $\alpha(E_\varphi) \in [0,1]$ such that $V(E_\varphi) = [\alpha(E_\varphi) \inf_{\mathcal{C}} E_\varphi + (1 - \alpha(E_\varphi)) \sup_{\mathcal{C}} E_\varphi] V(\mathbf{1})$. Now, observe that

$$\inf_{\mathcal{C}} E_\varphi = \inf_{\tilde{\mathcal{C}}} \tilde{E}_\varphi \qquad \textit{and} \qquad \sup_{\mathcal{C}} E_\varphi = \sup_{\tilde{\mathcal{C}}} \tilde{E}_\varphi$$

and that the mapping $E_\varphi \longmapsto \tilde{E}_\varphi$ from $\kappa(B(\Sigma) \longrightarrow \tilde{\kappa}(B(\Sigma))$ is clearly one-to-one and onto. Hence, we can define $\tilde{\alpha} : \tilde{\kappa}(B(\Sigma)) \longrightarrow [0,1]$ by $\tilde{\alpha}(\tilde{E}_\varphi) = \alpha(E_\varphi)$. Then, $\tilde{V} : \tilde{\kappa}(B(\Sigma)) \longrightarrow \mathbb{R}$ defined by $\tilde{V}(\tilde{E}_\varphi) = [\tilde{\alpha}(\tilde{E}_\varphi) \inf_{\tilde{\mathcal{C}}} \tilde{E}_\varphi + (1 - \tilde{\alpha}(\tilde{E}_\varphi)) \sup_{\tilde{\mathcal{C}}} \tilde{E}_\varphi] V(\mathbf{1})$ is the unique functional satisfying $V(E_\varphi) = \tilde{V}(\tilde{E}_\varphi)$, for all $\varphi \in B(\Sigma)$. The second part is immediate. $\square$

By means of the lemma, the problem of representing V is transformed into that of representing $\tilde{V}$. The advantage of doing so is that the domain of $\tilde{V}$ is the space $A(\tilde{\mathcal{C}})$ of all weak*-continuous affine functions on the convex, weak*-compact set $\tilde{\mathcal{C}}$. $A(\tilde{\mathcal{C}})$ is a (closed) subspace of the Banach space $C(\tilde{\mathcal{C}})$ of all weak*-continuous functions on $\tilde{\mathcal{C}}$ equipped with the sup-norm. Moreover, $\tilde{\kappa}$ is the canonical linear mapping $B(\Sigma) \longrightarrow A(\tilde{\mathcal{C}})$. Thus, the proof of Theorem 1 becomes an easy consequence of [**2**, Cor. 1].

Proof of Theorem 1. Let $\succsim$ be an ABEO preference relation over $\mathcal{A}$. Thus, $\succsim$ is represented by a functional $I = V \circ \kappa$. $\succsim$ satisfies Assumption 0 and AXIOM 1 iff V is monotone and equivariant with respect to positive affine transformations. By Lemma 1,

$$I = V \circ \kappa = \tilde{V} \circ \tilde{\kappa}$$

where $\tilde{V}$ and $\tilde{\kappa}$ are as defined in Lemma 1. Now, let $\tilde{E}_\varphi, \tilde{E}_\psi \in \tilde{\kappa}(B(\Sigma)) = A(\tilde{\mathcal{C}})$. If $\tilde{E}_\varphi$ and $\tilde{E}_\psi$ are non-constant, then $\tilde{E}_\varphi$ and $\tilde{E}_\psi$ are comonotonic if and only if they are isotonic [**2**, Prop. 2]. In such a case, there exist $\lambda > 0$ and $\beta \in \mathbb{R}$ such that $\tilde{E}_\psi = \lambda \tilde{E}_\varphi + \beta \mathbf{1}$. Then, by the equivariance of $\tilde{V}$ with respect to positive affine transformations (Lemma 1, above)

$$\tilde{V}(\tilde{E}_\varphi + \tilde{E}_\psi) = \tilde{V}((1+\lambda)\tilde{E}_\varphi + \beta) = \tilde{V}(\tilde{E}_\varphi) + \tilde{V}(\tilde{E}_\psi)$$

That is, $\tilde{V}$ is comonotonic additive on its domain. By Lemma 1, $\tilde{V}$ is monotone as well. By [**2**, Cor. 1], these functionals can be represented by Choquet integrals. That is, there exists a capacity Γ on the Borel sets of $\tilde{\mathcal{C}}$ such that $\forall u \circ \varphi \in B(\Sigma)$

$$I(\varphi) = V \circ \kappa(\varphi) = \int_{\tilde{\mathcal{C}}} \tilde{\kappa}(\varphi) d\Gamma$$

Conversely, let $\succsim$ be defined by

$$\varphi \succsim \psi \qquad \Longleftrightarrow \qquad I(\varphi) = \int_{\tilde{\mathcal{C}}} \tilde{\kappa}(\varphi) d\Gamma \geq I(\psi) = \int_{\tilde{\mathcal{C}}} \tilde{\kappa}(\psi) d\Gamma$$

Then, $\kappa(\varphi) \geq \kappa(\psi)$ implies $\tilde{\kappa}(\varphi) \geq \tilde{\kappa}(\psi)$ which, in turn, implies $\varphi \succsim \psi$ by the monotonicity of the Choquet integral. We conclude that $\succsim$ satisfies Unanimity. Finally, by the equivariance with respect to positive affine transformations of the Choquet integral, Assumption 0 is also satisfied. □

Proof of Corollary 1. Assume that $\succsim$ is represented by a functional $I(\varphi) = V \circ \kappa(\varphi) = \int_{\tilde{\mathcal{C}}} \tilde{\kappa}(\varphi) d\Gamma$ with Γ a probability measure on the Borel sets of $\tilde{\mathcal{C}}$. Since for every $\varphi \in \mathcal{A}$, $\tilde{\kappa}(\varphi)$ is affine on $\tilde{\mathcal{C}}$, we have that for every $\varphi \in \mathcal{A}$

$$I(\varphi) = \tilde{\kappa}(\varphi)(P_\Gamma)$$

where $P_\Gamma \in \tilde{\mathcal{C}}$ is the barycenter of Γ. Since $P_\Gamma \in \tilde{\mathcal{C}}$, there exists a probability measure λ supported by the set of extreme points of $\tilde{\mathcal{C}}$, $\mathrm{supp}(\lambda) \subseteq ext\{\tilde{\mathcal{C}}\}$, such that for every $\varphi \in \mathcal{A}$

$$I(\varphi) = \int_{ext\{\tilde{\mathcal{C}}\}} \tilde{\kappa}(\varphi) d\lambda$$

Since $ext\{\tilde{\mathcal{C}}\} \subseteq \mathcal{E}$, the latter can be written as in integral on $\mathcal{E}$. In particular, in the case where $\mathcal{E}$ is finite, we obtain the expression in the statement. The converse is just the definition of linear opinion pool. $\square$

Prof of Proposition 1. Let $\varphi, \psi \in \mathcal{A}$ be such that $\varphi \sim \psi$. Let $\chi = \frac{1}{2}\varphi + \frac{1}{2}\psi$. If

$$\int u(\varphi)dP_e \geq \int u(\psi)dP_e \qquad for\, all \qquad e \in \mathcal{E}$$

then

$$\int u(\varphi)dP_e \geq \int u(\chi)dP_e \geq \int u(\psi)dP_e \qquad for\, all \qquad e \in \mathcal{E}$$

and, by unanimity, $\varphi \sim \chi \sim \psi$. The same conclusion obtains if we reverse the role of φ and ψ in the previous inequalities. If for some non-empty, strict subset $\mathcal{F}$ of $\mathcal{E}$ we have

$$\begin{aligned} \int u(\varphi)dP_e &\geq \int u(\psi)dP_e \qquad for\, all \qquad e \in \mathcal{F} \\ \int u(\psi)dP_e &\geq \int u(\varphi)dP_e \qquad for\, all \qquad e \in \mathcal{E} \diagdown \mathcal{F} \end{aligned}$$

then, we also have

$$\begin{aligned} \int u(\varphi)dP_e &\geq \int u(\chi)dP_e \geq \int u(\psi)dP_e \qquad for\, all \qquad e \in \mathcal{F} \\ \int u(\psi)dP_e &\geq \int u(\chi)dP_e \geq \int u(\varphi)dP_e \qquad for\, all \qquad e \in \mathcal{E} \diagdown \mathcal{F} \end{aligned}$$

Then, Preference for Compromise implies that $\chi \succsim \varphi \sim \psi$. We conclude that in all cases $\frac{1}{2}\varphi + \frac{1}{2}\psi \succsim \varphi \sim \psi$, which shows that the preference $\succsim$ satisfies the uncertainty aversion axiom in Gilboa-Schmeidler [**15**]. By assumption, $\succsim$ also satisfies the conditions of Theorem 1, which readily imply that $\succsim$ satisfies Gilboa-Schmeidler's axioms of C-Independence and Monotonicity. Finally, the sup-norm continuity of the functional representing $\succsim$, which follows again from Theorem 1, implies that $\succsim$ satisfies also the Archimedean axiom of [**15**]. Then, by [**15**], there exists a unique, convex, weak*-compact set of probability measures $\mathcal{CE} \subset ba(\Sigma)$ such that

$$\varphi \succsim \psi \qquad iff \qquad \min_{P \in \mathcal{CE}} \int_S u(\varphi)dP \geq \min_{P \in \mathcal{CE}} \int_S u(\psi)dP \tag{5.3}$$

To see that the inclusion $\mathcal{CE} \subset \tilde{\mathcal{C}} = \overline{co}\,\{P_e\}_{e \in \mathcal{E}}$ holds, suppose by the way of contradiction that $\exists \hat{P} \in \mathcal{CE} \setminus \tilde{\mathcal{C}}$. Then, by the Separation Theorem, there exists $u \circ \varphi \in B(\Sigma)$ such that

$$\int_S u(\varphi)d\hat{P} < \int_S u(\varphi)dP \qquad for\ all\ P \in \tilde{\mathcal{C}}$$

By assumption, $\succsim$ satisfies the conditions of Theorem 1. Then, $\succsim$ has the representation of Theorem 1 for some capacity Γ as well as the representation (5.3). But, by the monotonicity of the Choquet integral, it follows that

$$I(\varphi) = \int_{\tilde{\mathcal{C}}} \tilde{E}_\varphi d\Gamma = \int_{\tilde{\mathcal{C}}} \int_S u(\varphi) dP d\Gamma > \int_S u(\varphi) d\hat{P} \geq \min_{P \in \mathcal{CE}} \int_S u(\varphi) dP = I(\varphi)$$

a contradiction. $\square$

Proof of Corollary 2. By Proposition 1, $\succsim$ is represented by the functional

$$I(\varphi) = \min_{P \in \mathcal{CE}} \int_S u(\varphi) dP$$

First, replace each $P \in \mathcal{CE}$ with its barycenter, P_b. Then, just like in the proof of Corollary 1, observe that each P_b can be replaced by a measure λ_P, $P \in \mathcal{CE}$, such that $\text{supp}(\lambda_P) \subseteq ext\{\tilde{\mathcal{C}}\} \subseteq \mathcal{E}$. The assertion follows. $\square$

Section 3

Proof of Proposition 2. By AXIOM 1 and the factorization property $I = V \circ \gamma$, for every $\varphi \in \mathcal{A}$

$$\inf_{i \in \mathcal{I}} J_i(\varphi) \leq I(\varphi) \leq \sup_{i \in \mathcal{I}} J_i(\varphi)$$

Let $\mathcal{K} = \overline{co}\{\cup_{i \in \mathcal{I}} \mathcal{C}_i\}$ and, for every $\varphi \in \mathcal{A}$, let $\tilde{E}_\varphi : \mathcal{K} \longrightarrow \mathbb{R}$ be defined by $\tilde{E}_\varphi(P) = \int u \circ \varphi dP$. Then,

$$\bar{\Lambda}(\varphi) = \min_{\mathcal{K}} \tilde{E}_\varphi = \inf_{i \in \mathcal{I}} J_i(\varphi) \leq I(\varphi) \leq \sup_{i \in \mathcal{I}} J_i(\varphi) = \max_{\mathcal{K}} \tilde{E}_\varphi$$

and $\exists \alpha(\varphi) \in [0,1]$ such that

$$I(\varphi) = \alpha(\varphi) \min_{\mathcal{K}} \tilde{E}_\varphi + (1 - \alpha(\varphi)) \max_{\mathcal{K}} \tilde{E}_\varphi \tag{5.4}$$

If $\varphi, \psi \in \mathcal{A}$ are such that $\tilde{E}_\varphi = \tilde{E}_\psi$, then $J_i(\varphi) = J_i(\psi)$ for every $i \in \mathcal{I}$ and AXIOM 1 implies $I(\varphi) = I(\psi)$. Then, equation (5.4) implies $\alpha(\varphi) = \alpha(\psi)$. It then follows that I factors as $I = \tilde{V} \circ \kappa$, where $\kappa : \varphi \longmapsto \tilde{E}_\varphi$.

Since all experts are of class $\mathcal{NB}$ and $\succsim$ satisfies Assumption 0 and AXIOM 1, $\tilde{V}$ is equivariant with respect to positive affine transformations and monotone on the set of sup-norm continuous affine functions on $\mathcal{K}$. By [**2**, Cor. 1], $\tilde{V}$ has a monotone comonotonic additive extension to the space of bounded, measurable functions on $\mathcal{K}$, which yields that for every $\varphi \in \mathcal{A}$

$$I(\varphi) = \int_{\tilde{\mathcal{C}}} \tilde{E}_\varphi d\Gamma$$

where Γ is a capacity on the Borel sets of $\mathcal{K}$ and the integral is taken in the sense of Choquet. $\square$

It is worth observing that alternative representations of the Choquet type obtain as follow. Let $\succsim$ be an AEO preference and let $I = V \circ \gamma$ be its representing functional. By definition, AEO preferences are complete and transitive. Since each i is of class $\mathcal{NB}$, each J_i satisfies C-independence (Definition 3). Thus, for each $\varphi \in \mathcal{A}$, $F_{\alpha\varphi+(1-\alpha)\mathbf{1}} = \alpha F_\varphi + (1-\alpha)\mathbf{1}$. This along with Assumption 0 implies that $\succsim$ satisfies the axiom of C-Independence in [**15**]. If $\varphi(s) \succsim \psi(s)$ for every $s \in S$, then $J_i(\varphi) \geq J_i(\psi)$ for every $i \in \mathcal{I}$ because each i is of class $\mathcal{NB}$ (thus, satisfies (1) in Definition 3). By AXIOM 1, $\varphi \succsim \psi$. By an elementary argument, these properties imply that $\succsim$ is also Archimedean. Thus, $\succsim$ satisfies the first five axioms in [**15**], and by [**2**, Theorem 2] $\succsim$ has the representation

$$I(\varphi) = \int_{\mathcal{C}} \int_{S} u(\varphi) dP d\Gamma' \tag{5.5}$$

where $\mathcal{C}$ is a set of probability charges, Γ' is a capacity on the Borel sets of $\mathcal{C}$ and the integral is taken in the sense of Choquet. Finally, by Krein-Milmann, $\mathcal{C}$ is the (closed) convex hull of its extreme points, $\mathcal{C} = \overline{co}\{P_e\}_{e\in\mathcal{E}}$, and each $e \in \mathcal{E}$ can be interpreted as a Bayesian expert.

Proof of Proposition 3. Assume that $\succsim$ satisfies Assumption 0, AXIOM 1 and AXIOM 2. Since $\succsim$ is an AEO satisfying Assumption 0 and AXIOM 1 and since the MEU class is contained in the $\mathcal{NB}$ class, then $\succsim$ satisfies the first five axioms in [**15**], as shown following the proof of Proposition 2. Next, let $\varphi, \psi \in \mathcal{A}$ be such that $\varphi \sim \psi$, and let $\chi = \frac{1}{2}\varphi + \frac{1}{2}\psi$. We want to show that $\chi \succsim \psi \sim \varphi$.

If $J_i(\varphi) \geq J_i(\psi)$ for every $i \in \mathcal{I}$, then by the concavity of each J_i we have that $J_i(\chi) \geq J_i(\psi)$ for every $i \in \mathcal{I}$. By AXIOM 1, $\chi \succsim \psi \sim \varphi$, and the same conclusion holds if we reverse the role of φ and ψ.

If there exists a $\mathcal{J}$ strictly contained in $\mathcal{I}$ such that $J_i(\varphi) \geq J_i(\psi)$ for every $i \in \mathcal{J}$ and $J_i(\psi) \geq J_i(\varphi)$ for every $i \in \mathcal{I} \diagdown \mathcal{J}$, then, by the concavity of each J_i, we can then partition $\mathcal{I}$ into three subsets, $\mathcal{I} = \mathcal{I}_1^0 \cup \mathcal{I}_2^0 \cup \mathcal{I}_3^0$, so that

$$\begin{array}{rclr} J_i(\varphi) & \geq & J_i(\chi) \geq J_i(\psi) & \quad for\ all\ i \in \mathcal{I}_1^0 \\ J_i(\psi) & \geq & J_i(\chi) \geq J_i(\varphi) & \quad for\ all\ i \in \mathcal{I}_2^0 \\ J_i(\chi) & > & \max\{J_i(\psi), J_i(\varphi)\} & \quad for\ all\ i \in \mathcal{I}_3^0 \end{array}$$

If either $\mathcal{I}_1^0$ or $\mathcal{I}_2^0$ is empty, the assertion follows from unanimity; if $\mathcal{I}_3^0$ is empty, it follows from Preference for Compromise. If $\mathcal{I}_1^0, \mathcal{I}_2^0$ and $\mathcal{I}_3^0$ are all non-empty, proceed as follows. Since $u \circ \varphi$ and $u \circ \psi$ are bounded functions and J_i is monotone for each i, there exists a $\kappa_1 > 0$ such that $\forall i \in \mathcal{I}_3^0$

$$J_i(\varphi + \kappa_1\mathbf{1}) \geq J_i(\chi - \kappa_1\mathbf{1})$$

Define

$$\begin{array}{rcl} \mathcal{I}_1^{\kappa_1} & = & \{i \in \mathcal{I}\,|J_i(\varphi + \kappa_1\mathbf{1}) \geq J_i(\chi - \kappa_1\mathbf{1}) \geq J_i(\psi - \kappa_1\mathbf{1})\} \\ \mathcal{I}_2^{\kappa_1} & = & \{i \in \mathcal{I}\,|J_i(\psi - \kappa_1\mathbf{1}) \geq J_i(\chi - \kappa_1\mathbf{1}) \geq J_i(\varphi + \kappa_1\mathbf{1})\} \\ \mathcal{I}_3^{\kappa_1} & = & \{i \in \mathcal{I}\,|J_i(\chi - \kappa_1\mathbf{1}) > \max\{J_i(\psi - \kappa_1\mathbf{1}), J_i(\varphi + \kappa_1\mathbf{1})\}\} \end{array}$$

Notice that $\mathcal{I}_1^0 \cup \mathcal{I}_3^0 \subset \mathcal{I}_1^{\kappa_1}$, which is thus non-empty. Now, if $\mathcal{I}_3^{\kappa_1} = \varnothing$, $\chi - \kappa_1\mathbf{1}$ is a compromise between $\varphi + \kappa_1\mathbf{1}$ and $\psi - \kappa_1\mathbf{1}$ and by AXIOM 2 is preferred to the worst of the two, which is

$\psi - \kappa_1 \mathbf{1}$ by the assumption $\psi \sim \varphi$ and AXIOM 1. Thus, $\chi - \kappa_1 \mathbf{1} \succsim \psi - \kappa_1 \mathbf{1} \Longleftrightarrow$ (by Assumption 0) $\chi \succsim \psi \sim \varphi$. If $\mathcal{I}_2^{\kappa_1} = \varnothing$, $\chi - \kappa_1 \mathbf{1} \succsim \psi - \kappa_1 \mathbf{1}$ by AXIOM 1, and the same conclusion obtains. If both $\mathcal{I}_3^{\kappa_1}$ and $\mathcal{I}_2^{\kappa_1}$ are non-empty, we can iterate the procedure. Since $u \circ \varphi$ and $u \circ \psi$ are bounded functions and J_i is monotone for each i, there exists a $\bar{\kappa}$ such that $J_i(\varphi + \bar{\kappa}\mathbf{1}) \geq J_i(\psi - \bar{\kappa}\mathbf{1})$ for all $i \in \mathcal{I}$, and we conclude that there exists a $\kappa^* \leq \bar{\kappa}$ such that $\mathcal{I}$ can be partitioned into the three subsets

$$\begin{aligned}
\mathcal{I}_1^{\kappa^*} &= \{i \in \mathcal{I} \,| J_i(\varphi + \kappa^* \mathbf{1}) \geq J_i(\chi - \kappa^* \mathbf{1}) \geq J_i(\psi - \kappa^* \mathbf{1})\} \\
\mathcal{I}_2^{\kappa^*} &= \{i \in \mathcal{I} \,| J_i(\psi - \kappa^* \mathbf{1}) \geq J_i(\chi - \kappa^* \mathbf{1}) \geq J_i(\varphi + \kappa^* \mathbf{1})\} \\
\mathcal{I}_3^{\kappa^*} &= \{i \in \mathcal{I} \,| J_i(\chi - \kappa^* \mathbf{1}) > \max\{J_i(\psi - \kappa^* \mathbf{1}), J_i(\varphi + \kappa^* \mathbf{1})\}\}
\end{aligned}$$

and at least one between $\mathcal{I}_2^{\kappa^*}$ and $\mathcal{I}_3^{\kappa^*}$ is empty. Then, the conclusion $\chi \succsim \psi \sim \varphi$ follows in the same way we saw above. We conclude that $\succsim$ satisfies all the axioms in [**15**], and the claimed representation follows. The extreme points of $\tilde{\mathcal{C}}$ can then be interpreted as Bayesian experts.

To show the inclusion $\tilde{\mathcal{C}} \subseteq \mathcal{K} = \overline{co}\left\{\bigcup_{i \in \mathcal{I}} \mathcal{C}_i\right\}$, for each $\varphi \in \mathcal{A}$ let

$$\hat{\varphi} = \min_{P \in \bigcup_{i \in \mathcal{I}} \mathcal{C}_i} \int_S u(\varphi) dP = \min_{i \in \mathcal{I}} J_i(\varphi)$$

Then, for each $\varphi \in \mathcal{A}$ and for all $i \in \mathcal{I}$, we have $J_i(\varphi) \geq J_i(\hat{\varphi}\mathbf{1})$ and AXIOM 1 implies that $\varphi \succsim \hat{\varphi}\mathbf{1}$. Suppose now that $\tilde{\mathcal{C}} \not\subseteq \mathcal{K} = \overline{co}\left\{\bigcup_{i \in \mathcal{I}} \mathcal{C}_i\right\}$. Then, $\exists \tilde{P} \in \tilde{\mathcal{C}} \diagdown \mathcal{K}$. By the separation theorem, $\exists \varphi \in \mathcal{A}$ such that

$$I(\varphi) = \min_{P \in \tilde{\mathcal{C}}} \int_S u(\varphi) dP \leq \int_S u(\varphi) d\tilde{P} < \min_{P \in \mathcal{K}} \int_S u(\varphi) dP = u(\hat{\varphi}\mathbf{1}) \leq I(\varphi)$$

a contradiction.

To show the second part, simply observe that any $P' \in \mathcal{K}$ can be written as the barycenter of a measure $\lambda_{P'}$ on the Borel sets of $\mathcal{K}$. That is,

$$\int_S u(\varphi) dP' = \int_{\mathcal{K}} \int_S u(\varphi) dP d\lambda_{P'}$$

and I can be written in the form

$$I(\varphi) = \min_{\lambda \in \Lambda} \int_{\mathcal{K}} \int_S u(\varphi) dP d\lambda$$

$\square$

The following proof of Proposition 4 uses a fictitious construct, which consists of gluing together n copies of the set of states S, one for each expert, and of viewing an expert's ranking as the preference that obtains conditional on the copy assigned to that expert.

PROOF OF PROPOSITION 4. Let $\mathcal{I} = \{1, 2, ..., n\}$. For each $i \in \mathcal{I}$, let $S_i = S$ denote a copy of the space S, and let $\Omega = \oplus S_i$ denote the disjoint union of the S_i's. We are going to view an alternative $\varphi \in \mathcal{A}$ as a mapping $\varphi : \Omega \longrightarrow X$ (X the space of consequences used throughout the paper) which produces the consequence $\varphi(s_i)$ if $s_i \in S_i$ realizes.

For each $i \in \mathcal{I}$, let us define a preference $\succsim_i^*$ on $\mathcal{A}$ by

$$\varphi \succsim_i^* \psi \qquad \Longleftrightarrow \qquad J_i(\varphi) \geq J_i(\psi)$$

where the probabilities $P \in \mathcal{C}_i$ appearing the functional J_i are understood as being supported by S_i. The statement $\varphi \succsim_i^* \psi$ is to be interpreted as "φ is preferred to ψ conditionally on S_i".

Next, let us define a preference relation $\succsim^*$ on $\mathcal{A}$ by

$$\varphi \succsim^* \psi \qquad \Longleftrightarrow \qquad I(\varphi) \geq I(\psi)$$

where I is the functional of Proposition 3. By Proposition 3, $\succsim^*$ satisfies all the axioms in [**15**] and there exists a unique, convex, weak*-compact set Λ of probability charges on $(\Omega, \otimes_1^n \Sigma)$ such that for every $\varphi \in \mathcal{A}$ has the representation

$$\tilde{I}(\varphi) = \min_{\lambda \in \Lambda} \int_{\Omega} u \circ \varphi d\lambda \tag{5.6}$$

We now consider the family of preferences $\{\succsim^*, \succsim_i^*\}_{i \in \mathcal{I}}$. For $\varphi, \psi \in \mathcal{A}$, denote by $\varphi i \psi$ the alternative in $\mathcal{A}$ which is equal to φ on S_i and equal to ψ on $\Omega \diagdown S_i$. Now observe that for each $i \in \mathcal{I}$ and for all $\varphi, \psi \in \mathcal{A}$,

$$\varphi \sim_i^* \varphi i \psi$$

($\sim_i^*$ is the symmetric part of $\succsim_i^*$) and that

$$\varphi \succsim_i^* \psi \qquad \forall i \in \mathcal{I} \qquad \Longrightarrow \qquad \varphi \succsim^* \psi$$

because $\succsim^*$ is defined by means of the functional $I(\cdot)$ which satisfies both the factorization property $I = V \circ \gamma$ and AXIOM 1. Thus, the family $\{\succsim^*, \succsim_i^*\}_{i \in \mathcal{I}}$ satisfies both Consequentialism and Dynamic Consistency with respect to the partition $\{S_i\}_{i \in \mathcal{I}}$, and all the preferences are of MEU type. Then, it follows from Epstein-Schneider [**12**] that the set Λ in (5.6) is *rectangular* (see [**12**]). By the rectangularity of Λ,

$$\begin{aligned}
\tilde{I}(\varphi) &= \min_{\lambda \in \Lambda} \int_{\Omega} u \circ \varphi d\lambda = \min_{\lambda \in \Lambda} \sum_{i=1}^{n} \lambda(S_i) \min_{\lambda \in \mathcal{C}_i} \int_{\Omega} u \circ \varphi d\lambda(\cdot \mid S_i) \\
&= \min_{\tilde{\lambda} \in \Delta(\mathbb{R}^n)} \sum_{i=1}^{n} \tilde{\lambda}_i J_i(\varphi)
\end{aligned}$$

□

Section 5

PROOF OF PROPOSITION 6. It is easy to see that if I satisfies the conditions of Proposition 2/Theorem 1, then ρ satisfies properties (1), (2) and (3). The converse is shown in Amarante [**3**]. □

References

[1] Alon S. and G. Gayer (2016), Utilitarian Preferences with Multiple Priors, *Econometrica* **84**, 1181-1201.

[2] Amarante M. (2009), Foundations of Neo-Bayesian Statistics, *Journal of Economic Theory* **144**, 2146-73.

[3] Amarante M. (2016), A representation of risk measures, *Decisions in Economics and Finance* **39**, 95-103.

[4] Amarante M. (2017), Mm-OWA: a generalization of OWA operators, mimeo.

[5] Artzner P., F. Delbaen, J-M Eber and D. Heath (1999), Coherent measures of risk, *Mathematical Finance* **9**, 203-28.

[6] Blackorby C., D. Donaldson and J. A. Weymark, 1999, Harsanyi's social aggregation theorem for state-contingent alternatives, *Journal of Mathematical Economics* **32**, 365–387.

[7] Chambers C. (2007), Ordinal aggregation and quantiles, *Journal of Economic Theory* **137**, 416-31.

[8] Coulhon T. and P. Mongin, 1989, Social choice theory in the case of von Neumann – Morgenstern utilities, *Social Choice and Welfare* **6**, 175-87.

[9] Crès H., I. Gilboa and N. Vieille (2011), Aggregation of multiple prior opinions, *Journal of Economic Theory*, **146**, 2563-2582.

[10] Dhillon A. and J-F Mertens (1999), Relative Utilitarianism, *Econometrica* **67**, 471-98.

[11] Duffie, D. and J. Pan (1997), An overview of value at risk, *The Journal of Derivatives* **4**, 7-49.

[12] Epstein L. and M. Schneider (2003), Recursive multiple priors, *Journal of Economic Theory* **113**, 1-31.

[13] Föllmer H. and A. Schied (2002), Convex measures of risk and trading constraints. *Finance and Stochastics* **6**, 429-447.

[14] Ghirardato P., F. Maccheroni and M. Marinacci (2004), Differentiating ambiguity and ambiguity attitude, *Journal of Economic Theory*, **118**, 133-173.

[15] Gilboa I. and D. Schmeidler (1989), Maxmin expected utility with a non-unique prior, *Journal of Mathematical Economics* **18**, 141-53.

[16] Grabisch M., OWA Operators and Nonadditive Integrals, in Recent Developments in the Ordered Weighted Averaging Operators: Theory and Practice, R. R. Yager, J. Kacprzyk and G. Beliakov Eds, Studies in Fuzziness and Soft Computing, Volume 265, pp. 3-15, Springer 2011.

[17] Harsanyi J.C. (1955), Cardinal welfare, individualistic ethics, and interpersonal comparisons of utility, *Journal of Political Economy* **63**, 309–321.

[18] Jorion P. (2007) Value at risk: The new benchmark for managing financial risk, McGraw-Hill, Third ed.

[19] Kadane J.B. and L. Wasserman (1996), Symmetric, coherent, Choquet capacities, *The Annals of Statistics* **24**, 1250-64.

[20] Maccheroni F., M. Marinacci and A. Rustichini (2006), Ambiguity aversion, robustness and the variational representation of preferences, *Econometrica* **74**, 1447-1498.

[21] Markowitz H. M. (1952), Portfolio Selection, *Journal of Finance* **7**, 77-91.

[22] Markowitz, H. M. (1959), Portfolio Selection: Efficient Diversification of Investments, New York: John Wiley & Sons.

[23] Mesiar R., A. Stupnanova and R. R. Yager (2015), Generalizations of OWA Operators, IEEE Transactions on Fuzzy Systems, 23, 2154-2162.

[24] Murofushi T. and M. Sugeno, Some quantities represented by the Choquet integral, *Fuzzy Sets and Systems* 1993, 56, 229-235.

[25] Nascimento L. (2012), The ex-ante aggregation of opinions under uncertainty, *Theoretical Economics* **7**, 535-70.

[26] Phelps R. R. (1966), Lectures on Choquet theorem, van Nostrand.

[27] Schmeidler D. (1986), Integral Representation without Additivity, *Proceedings of the AMS* **97**, 255-61.

[28] Sprumont Y. (2013), On relative egalitarianism, *Social Choice and Welfare* **40**, 1015-1032.

[29] Stone M. (1961), The linear opinion pool, *Annals of Mathematical Statistics* **32**, 1339–1342.

[30] Yager RR (1988), On ordered weighted averaging aggregation operators in multicriteria decision making, *IEEE Trans Systems Man and Cybernet*,18(1):183–190.

Université de Montréal et CIREQ
E-mail address: massimiliano.amarante1@gmail.com