Aggregation of multiple prior opinions

Hervé Crès, Itzhak Gilboa,, Nicolas Vieille

Available online at www.sciencedirect.com

SciVerse ScienceDirect

Journal of Economic Theory 146 (2011) 2563–2582

JOURNAL OF Economic Theory

www.elsevier.com/locate/jet

Notes

Aggregation of multiple prior opinions ☆

Hervé Crès [a], Itzhak Gilboa [b,c,*], Nicolas Vieille [b]

[a] *Sciences-Po, Paris, France*
[b] *HEC, Paris, France*
[c] *Tel-Aviv University, Israel*

Received 2 November 2009; final version received 18 October 2010; accepted 25 June 2011

Available online 12 October 2011

Abstract

Experts are asked to provide their advice in a situation of uncertainty. They adopt the decision maker's utility function, but each has a potentially different set of prior probabilities, and so does the decision maker. The decision maker and the experts maximize the minimal expected utility with respect to their sets of priors. We show that a natural Pareto condition is equivalent to the existence of a set Λ of probability vectors over the experts, interpreted as possible allocations of weights to the experts, such that (i) the decision maker's set of priors is precisely all the weighted-averages of priors, where an expert's prior is taken from her set and the weight vector is taken from Λ; (ii) the decision maker's valuation of an act is the minimal weighted valuation, over all weight vectors in Λ, of the experts' valuations.

JEL classification: D7; D8

Keywords: Aggregation of opinions; Ambiguity; Multiple priors

☆ We thank Thibault Gajdos, Philippe Mongin, Ben Polak, Jean-Christophe Vergnaud, four anonymous reviewers, and the associate editor for many detailed comments and references. Gilboa gratefully acknowledges ISF Grant 396/10 and ERC Grant 269754.

* Corresponding author at: HEC, Paris, France and Tel-Aviv University, Israel.
E-mail addresses: herve.cres@sciences-po.fr (H. Crès), tzachigilboa@gmail.com (I. Gilboa), vieille@hec.fr (N. Vieille).

0022-0531/$ – see front matter
doi:10.1016/j.jet.2011.06.018

1. Introduction

1.1. The problem

A decision maker is offered a partnership in a business venture. The offer appears very attractive, but, after she talks to some friends, it's pointed out to her that the project will be a serious failure if the globe becomes warmer by 3–4 degrees Celsius. She decides to study the matter and to find out what, according to the experts, is the probability of such a warming taking place within, say, the next ten years.

It turns out that different experts have different opinions. Indeed, the phenomenon of global warming that we presumably witness today is not similar enough to anything that we have observed in the past. The cycles of the globe's temperature take a very long time, and the conditions that prevailed in past warmings differ significantly from those in the present. Thus, one cannot resort to simple empirical frequencies in order to assess probabilities. More sophisticated econometric techniques should be invoked to come up with such assessments, but they tend to depend on assumptions that are not shared by all. In short, the experts do not agree, and we cannot assume that the event in question has an "objective" probability. What should the decision maker do? What is a rational way of aggregating the opinions of the experts?

1.2. The Bayesian case

Consider first the benchmark case in which all experts are Bayesian. Specifically, let p_i be the subjective probability of the event according to expert $i = 1, \ldots, n$. It seems very natural to take the arithmetic average of the assessments of the experts. This approach may be a little too simple because it treats all experts equally. More generally, we may consider a weighted average of the experts' opinions. For every vector of non-negative weights $\lambda = (\lambda_i)_i$ summing up to 1, we may define

$$p_0 = p_0^\lambda = \sum_{i=1}^{n} \lambda_i p_i. \tag{1}$$

Moreover, in a more general case where p_i are probability vectors, distributions on the real line, or general probability measures, their weighted average p_0^λ is well defined, and suggests itself as a natural candidate for the decision maker's beliefs. This rule for aggregation of probabilistic assessments has been dubbed *linear opinion pool* by Stone [37], and is attributed to Laplace (see [2,29,16], for a survey).

We will henceforth interpret $p_0, p_1, \ldots, p_n$ as probability measures on a state space, and refer to

$$p_0(A) = p_0^\lambda(A) = \sum_{i=1}^{n} \lambda_i p_i(A)$$

when we discuss the probability of a specific event A.

The weight λ_i can be viewed as the degree of confidence the decision maker has in expert i. Another interpretation is the following: suppose that there exists an "objective" uncertainty about the state that will obtain, and the "true" probability is known to be one of $p_1, \ldots, p_n$. The decision maker knows this structure, but does not know which of $p_1, \ldots, p_n$ is the actual probability, or the "real" data generating process. If the decision maker has Bayesian beliefs over the objective probability, given by the weights $\lambda = (\lambda_i)_i$, then p_0 is her derived subjective probability

over the state space. Thus, we can think of the decision maker as if she believed that one expert is guaranteed to generate the "correct" probability, and attaching the probability λ_i to the event that expert i has access to "the truth". We will refer to this interpretation as "the truth metaphor".[1]

The averaging of probabilities resembles the averaging of utilities in social choice theory. Correspondingly, such an averaging can be derived from a Pareto, or a unanimity condition à la Harsanyi [24]. Specifically, the decision maker's probability p_0 is of the form (1) if and only if the following holds: for every two choices a and b, and every utility function u, if all experts agree that the expected utility (u) of a is at least as high as that of b, so should the decision maker.[2]

1.3. Uncertainty with Bayesian experts

In terms of elegance and internal coherence, the Bayesian paradigm is hardly matched by any other approach to modeling uncertainty. Moreover, it relies on powerful axiomatic foundations laid by Ramsey [32], de Finetti [9,10], and Savage [33]. Yet, it has been criticized on descriptive and normative grounds alike. Following the seminal contributions of Knight [27], Ellsberg [11], and Schmeidler [34,35], there is a growing body of literature dealing with more general models representing uncertainty.[3] Importantly, many authors view these models as not necessarily less rational than the Bayesian benchmark. In particular, since the Bayesian approach calls for the formulation of subjective beliefs, it has been argued that the very fact that these beliefs differ across individuals might indicate that it is perhaps not rational to insist on one of them.[4] In our context, the decision maker may ask herself, "If the experts fail to agree on the probability, how can I be so sure that there is one probability that is 'correct'? Maybe it's safer to allow for a *set* of possible probabilities, rather than pick only one?"

In this paper we focus on the maxmin expected utility ("MEU") model, suggested by Gilboa and Schmeidler (GS [23]), which is arguably the simplest multiple prior model that defines a complete ordering over alternatives.[5] Given a set of probabilities C, each possible act f is evaluated by

$$J(f) = \min_{p \in C} EU_p(f)$$

where $EU_p(f)$ is the expected utility of act f according to the probability p. Uncertainty aversion is built into this decision rule by the min operator, evaluating each act by its worst-case expected utility (over $p \in C$).

Applied to our context, the decision maker may, for example, be very cautious and consider as the set C all the probability assessments of the experts, or, equivalently, their convex hull.

[1] We find this interpretation useful, but it should be stressed that our model does not capture notions such as "truth" or "objectivity".

[2] Such derivations are restricted to the case in which all experts share the same utility function, and they require some richness conditions. Hylland and Zeckhauser [25], Mongin [30], Blackorby, Donaldson and Mongin [4], Chambers and Hayashi [8] and Zuber [39] provided impossibility results for the simultaneous aggregation of utilities and probabilities. Gilboa, Samet and Schmeidler [22] restricted unanimity to choices over which there are no differences of belief, and derived condition (1) under certain assumptions. However, Gajdos, Tallon and Vergnaud [13] showed that in the presence of uncertainty aggregation of preferences is impossible even if the agents have the same beliefs.

[3] See Gilboa [17] and Gilboa and Marinacci [18] for surveys of axiomatic foundations of the Bayesian approach, their critiques, and alternative models.

[4] See [19–21].

[5] Sets of probabilities that are used to define incomplete preferences were studied by Bewley [3].

This may be a little extreme, however. If, for example, nine experts evaluate $P(A)$ at 0.4 and one at 0.1, the range of probabilities for the event A will be $[0.1, 0.4]$. Clearly, the same interval would result from nine experts providing the assessment $P(A) = 0.1$ and only one the assessment 0.4. Consequently, the decision maker may adopt a set of probabilities that is strictly smaller than the entire convex hull of the experts' beliefs. For instance, the decision maker may consider

$$C = \left\{ p = \sum_{i=1}^{n} \lambda_i p_i \;\middle|\; (\lambda_1, \ldots, \lambda_n) \in \Lambda \right\}$$

where $\Lambda \subset \Delta(\{1, \ldots, n\})$ is a set of weight vectors over the n experts. If Λ is a singleton, $\Lambda = (\lambda_1, \ldots, \lambda_n)$, the decision maker's behavior will be Bayesian, with the probability given by (1). If, by contrast, $\Lambda = \Delta(\{1, \ldots, n\})$, so that the decision maker allows for any possible vector of weights on the experts' opinions, the decision maker will behave according to the MEU model, with extreme uncertainty aversion as discussed above. In between these extremes, a non-singleton, proper subset Λ of $\Delta(\{1, \ldots, n\})$ allows the decision maker to (i) assign different weights to different experts; (ii) take into account how many experts provided a certain assessment; and (iii) leave room for a healthy degree of doubt.

1.4. Uncertainty averse experts

If, however, we accept the possibility that rationality, or at least scientific caution, may favor a set of probabilities over a single one, then we should also allow the experts to provide their assessments by sets of, rather than by single probabilities. Indeed, if the experts use standard statistical techniques, they may come up with confidence sets that cannot be shrunk to singletons without compromising the notion of "confidence". Thus we re-phrase the question and ask, how can beliefs be aggregated, where "beliefs" are modeled by sets of probabilities?

The procedure suggested above consisted in representing beliefs by a class of weighted averages of experts' beliefs, with varying weight vectors. It has a natural extension to the case of non-Bayesian experts: let $\Lambda \subset \Delta(\{1, \ldots, n\})$ be a set of weight vectors over the experts. Assume that each expert has beliefs that are modeled as a set of probabilities C_i. Then, let the decision maker entertain the beliefs given by the set

$$C_0 = \left\{ p = \sum_{i=1}^{n} \lambda_i p_i \;\middle|\; \lambda \in \Lambda,\ p_i \in C_i \right\}. \tag{2}$$

Assume that the decision maker uses this set in the context of the maxmin expected utility model. As in the case in which C_i are singletons, this formula allows the decision maker a wide range of attitudes towards the experts' opinions. If Λ is a singleton, the decision maker does not add any uncertainty aversion of her own: she takes a fixed average of all probabilities the experts deem possible, and the freedom of choosing a probability in C_0 is entirely due to the fact that the experts fail to commit to a single probability measure. By contrast, if Λ is the entire simplex $\Delta(\{1, \ldots, n\})$, the decision maker behaves as if any probability that at least one expert deems possible is indeed possible. And different sets Λ between these extremes allow for different attitudes towards the uncertainty generated by the choice of the expert.

One attractive feature of the decision rule proposed here is the following. Assume that each expert i uses the maxmin EU principle with a set of probabilities C_i as above. Given an act f,

expert i evaluates it by

$$J_i(f) = \min_{p \in C_i} EU_p(f).$$

Suppose that the experts do not report their entire sets of beliefs C_i, but only their bottom-line evaluation $J_i(f)$. Using the "truth metaphor", the decision maker then faces uncertainty about which expert is closest to the "truth". As above, the decision maker may consider n epistemic states of nature, where in state i expert i is right. Thus, if the decision maker knew which of the experts had access to "truth", she would follow that expert and evaluate f by $J_i(f)$. However, the decision maker does not know which expert has the correct assessment. Facing this uncertainty, the decision maker might adopt the maxmin EU rule, with a set of probabilities $\Lambda \subset \Delta(\{1, \ldots, n\})$. The decision maker should therefore evaluate act f by

$$J_0(f) = \min_{\lambda \in \Lambda} \sum_{i=1}^{n} \lambda_i J_i(f). \tag{3}$$

It turns out that this evaluation coincides with the rule proposed above for the same set Λ.[6] In other words, if the decision maker had access to the probabilities used by each expert, C_i, and were she to use the MEU rule relative to the set C_0 given in (2), she would have arrived at precisely the same conclusion as she would if she followed the rule (3) for the same set Λ. Thus, the set of weight vectors over the experts, Λ, can be interpreted in two equivalent ways, as in the Bayesian case: as a set of weights used to average the experts' beliefs, and as a set of probabilities over which expert is "right". In particular, the decision maker can use the rule (3), applying to the bottom-line evaluations of the experts, knowing that she would have arrived at the same evaluation had she asked the experts for their entire sets of probabilities.

Gajdos and Vergnaud [15] also deal with the aggregation of expert opinions under uncertainty. While their approach differs from ours in several significant ways (to be discussed below), they also axiomatize a functional of the form (3).

The MEU model has been criticized for its inability to disentangle objective, given uncertainty, which is presumably a feature of the decision problem at hand, from the subjective taste for uncertainty, which is a trait of the decision maker. For example, if we observe a set C that is a singleton, we cannot tell whether the decision maker has very precise information about the problem, allowing her to formulate a Bayesian prior, or has a natural tendency to treat uncertain situations as if they were risky. Gajdos, Hayashi, Tallon and Vergnaud [14] explicitly model both "hard" information and the set of priors that governs behavior, allowing for the latter to be a proper subset of the former.[7]

In our context, we interpret the sets C_i of the experts as their objective information, ignoring whatever attitudes towards uncertainty they may have. By contrast, the set Λ may reflect both the decision maker's information about the experts' reliability and experience, and her attitude towards uncertainty. Thus, it is possible that, given the same expert advice and the same information about the experts, a certain decision maker will choose a larger set Λ than will another decision maker, who is less averse to uncertainty.

[6] This fact, which is rather straightforward, is formally stated and proved in the sequel.

[7] Another approach to separate uncertainty from uncertainty attitudes is the "smooth" model [26,31,36], where the decision maker is assumed to aggregate (in a non-linear way) various expected utility evaluations of an act rather than focus on the minimal one.

The remainder of the paper is organized as follows. We first define the formal framework in Section 2.1. The aggregation rule discussed above is formally stated in Section 2.2, where we also state the equivalence between applying the MEU approach to the experts' valuations and applying it to the original problem, with the convex combinations of the experts' probabilities. We then turn to axiomatize this rule. To this end, we first explain the axiomatization informally in Section 2.3. Finally, the main result is stated in Section 2.4 and proved in Appendix A. Section 3 concludes with a discussion of related literature.

2. Model and results

2.1. Set-up

We use a version of the Anscombe–Aumann [1] model as re-stated by Fishburn [12].

Let X be a set of *outcomes*. Let L denote the set of von Neumann–Morgenstern [38] *lotteries*, that is, the distributions on X with finite support. L is endowed with a mixing operation: for every $P, Q \in L$ and every $\alpha \in [0,1]$, $\alpha P + (1-\alpha)Q \in L$ is given by

$$\big(\alpha P + (1-\alpha)Q\big)(x) = \alpha P(x) + (1-\alpha)Q(x).$$

The set of *states of the world* is S and it is endowed with a σ-algebra of *events* Σ. The set of *acts*, F, consists of the Σ-measurable simple functions from S to L. It is endowed with a mixture operation as well, performed pointwise. That is, for every $f, g \in F$ and every $\alpha \in [0,1]$, $\alpha f + (1-\alpha)g \in F$ is given by

$$\big(\alpha f + (1-\alpha)g\big)(s) = \alpha f(s) + (1-\alpha)g(s).$$

The decision maker ($i = 0$) and n experts ($i = 1, \ldots, n$) have binary relations $\succsim_i \subset F \times F$, interpreted as preference relations. The relations $\succ_i$, $\sim_i$ are defined as usual, namely, as the asymmetric and symmetric parts of $\succsim_i$, respectively.

We extend the relations $(\succsim_i)_i$ to L as usual. Thus, for $P, Q \in L$, $P \succsim_i Q$ means $f_P \succsim_i f_Q$ where, for every $R \in L$, $f_R \in F$ is the constant act given by $f_R(s) = R$ for all $s \in S$. The set of all constant acts is denoted F_c.

We assume that each of $\succsim_i$ satisfies the axioms of GS [23]. We further assume that all $(\succsim_i)_i$ agree on F_c, equivalently, on L. Thus, there exists an affine $u : L \to \mathbb{R}$ that represents $\succsim_i$ on L for $i = 0, \ldots, n$. Clearly, $range(u)$ is a convex subset of $\mathbb{R}$. We assume without loss of generality (i) that u is not a constant; and (ii) that 0 is in the interior of $range(u)$.

From GS [23] it follows that for each $i = 0, \ldots, n$ there exists a convex and weak*-closed set of finitely additive measures on (S, Σ), C_i, such that

$$f \succsim_i g \quad \text{iff} \quad J_i(f) \geqslant J_i(g)$$

where

$$J_i(f) = J_{u,C_i}(f) \equiv \min_{p \in C_i} \int u(f)\, dp. \tag{4}$$

Moreover, (u, C_i) are the unique pair that represents $\succsim_i$ as in (4) (up to an increasing affine transformation of u).

2.2. The aggregation rule

We start by observing the following.

Proposition 1. *For a convex and closed set* $\Lambda \subseteq \Delta(\{1,\dots,n\})$,

$$J_0(f) = \min_{\lambda\in\Lambda} \sum_{i=1}^{n} \lambda_i J_i(f) \tag{5}$$

iff

$$C_0 = \left\{ p = \sum_{i=1}^{n} \lambda_i p_i \;\middle|\; \lambda \in \Lambda,\ p_i \in C_i \right\}. \tag{6}$$

Thus, when one proposes that the decision maker choose a convex and closed set of weights $\Lambda \subset \Delta(\{1,\dots,n\})$, set

$$C_0 = \left\{ p = \sum_{i=1}^{n} \lambda_i p_i \;\middle|\; \lambda \in \Lambda,\ p_i \in C_i \right\}$$

and then define, as above,

$$J_0(f) = \min_{p\in C_0} \int u(f)\,dp,$$

one may equivalently propose that the decision maker employ the MEU approach on the experts' valuations, using the same set of weights $\Lambda \subset \Delta(\{1,\dots,n\})$.

2.3. Motivating the axiomatization

In the Bayesian setting the averaging over the experts' opinions follows from a unanimity principle. As in Harsanyi's [24] result, if all preference functionals are linear, then, under certain richness conditions, unanimity implies that the functional representing society's preferences is a linear combination of the functionals representing individuals' preferences. But in a non-Bayesian setting, this is no longer the case. Unanimity implies (i) that society's functional can be written as a function of the functionals of the individuals; and (ii) that this function is non-decreasing. Assuming that all functionals are continuous, one may conclude that there exists a continuous and non-decreasing function φ such that, for every act f,

$$J_0(f) = \varphi\big(J_1(f),\dots,J_n(f)\big) \tag{7}$$

where φ is defined over the range of the J_i's.

To derive a representation as in (3), one needs to know more about the function φ. First, it has to satisfy (1-)homogeneity and Shift whenever defined:

(i) φ is *homogeneous* if for $a = (a_1,\dots,a_n) \in \mathbb{R}^n$, $\gamma > 0$,

$$\varphi(\gamma a) = \gamma\varphi(a);$$

(ii) φ satisfies *Shift* if for $a = (a_1, \ldots, a_n) \in \mathbb{R}^n$, and $c \in \mathbb{R}$,

$$\varphi(a + \vec{c}) = \varphi(a) + c$$

where $\vec{c} = (c, \ldots, c) \in \mathbb{R}^n$.

Second, φ has to be concave.

The fact that $J_0, J_1, \ldots, J_n$ are all MEU functionals implies that each of them satisfies the corresponding properties, appropriately formulated for the space of acts. It turns out that this is sufficient to determine that φ satisfies homogeneity and Shift. Intuitively, this is similar to arguing that if φ maps linear functionals into a linear functional, φ itself has to be linear. However, concavity does not have this property: even in the case $n = 1$ the fact that a functional J is concave and that $\varphi(J)$ is concave does not imply that φ is concave.

One may, of course, stop here, and consider functionals as in (7) where φ need not be concave, rather than insist on the MEU-style aggregation of the experts' evaluations. However, the following example illustrates why we find concavity of φ natural in this setting.

Let there be two states 1, 2 and two experts. Expert 1 believes that state 1 is at least as likely as state 2, while expert 2 believes that state 2 is at least as likely as state 1. That is, $C_1 = \{p \in \Delta(S), \text{s.t. } p(1) \geqslant \frac{1}{2}\}$, and $C_2 = \{p \in \Delta(S), \text{s.t. } p(1) \leqslant \frac{1}{2}\}$. Set $C_0 = C_1 \cap C_2$: the decision maker believes that both states are equally likely. As a consequence,

$$J_0(f) = \max\big(J_1(f), J_2(f)\big) = \max_{\lambda \in \Delta(\{1,\ldots,n\})} \big\{\lambda_1 J_1(f) + \lambda_2 J_2(f)\big\}.$$

Clearly, unanimity holds, because max is a non-decreasing function. Moreover, the decision maker has concave preferences, which happen to be linear in this case. However, we find the decision maker's confidence in the probability $(0.5, 0.5)$ poorly justified: each expert admits that there is some uncertainty about the state of the world. The decision maker does not commit to follow the advice of a single expert or to use a fixed averaging of their evaluations. Rather, she behaves as if there is uncertainty about which expert is to be trusted – for some acts f, $J_0(f) = J_1(f)$, for others, $J_0(f) = J_2(f)$. And yet, miraculously, the uncertainty about the experts "cancels out" the uncertainty about the state of the world, and the agent behaves in a Bayesian way. There is nothing wrong about Bayesian beliefs when they are based on hard evidence, but in this case the Bayesian beliefs result from an optimistic mix of the pessimistic views of the experts.

We now seek an axiom that would rule out this type of aggregation. We wish to express the intuition that the decision maker is averse to uncertainty at the level of the experts, and not only at the level of the state of the world. We start with an example.

Assume that, for three acts, f, f_1, f_2, all experts agree that a 50%–50% average between the evaluations (or certainty equivalents) of f_1 and f_2 cannot do better than f. That is, for every $i = 1, \ldots, n$

$$J_i(f) \geqslant \frac{1}{2} J_i(f_1) + \frac{1}{2} J_i(f_2).$$

It stands to reason that the decision maker would reach the same conclusion, namely that

$$J_0(f) \geqslant \frac{1}{2} J_0(f_1) + \frac{1}{2} J_0(f_2).$$

Our main axiom will be a generalization of this condition. Before we state it explicitly, we explain it in this simple case.

To see why this axiom reflects uncertainty aversion, it may be useful to understand why we do not require the same condition for the opposite inequality. Suppose, again, that there are two experts and two states of the world. Expert $i = 1, 2$ assigns probability 1 to state i. Assume that f_i yields a payoff of 1 in state i and 0 in the other state $(3 - i)$. Thus,

$$J_1(f_1) = 1, \qquad J_1(f_2) = 0,$$
$$J_2(f_1) = 0, \qquad J_2(f_2) = 1.$$

By contrast, act f guarantees the payoff 0.5 in both states. Each of the two experts therefore believes that a 50%–50% average between the evaluations of acts f_1, f_2 is just as good as act f:

$$\frac{1}{2}J_i(f_1) + \frac{1}{2}J_i(f_2) = J_i(f).$$

But the decision maker need not accept this conclusion. Clearly, f is evaluated by 0.5 also by the decision maker, as the two experts agree on its evaluation. However, each act f_i has a worst case of 0 (if expert i happens to be wrong and expert $3 - i$ happens to get it right). Each expert, using the same set of probabilities for the evaluation of f_1 and of f_2, finds the average $\frac{1}{2}J_i(f_1) + \frac{1}{2}J_i(f_2)$ sufficiently attractive because one of the two values $\{J_i(f_1), J_i(f_2)\}$ is 1. But the decision maker, not being sure which expert is right, takes into account the worst case for each act separately. And then she finds that each of these acts has a worst case of zero, yielding

$$\frac{1}{2}J_0(f_1) + \frac{1}{2}J_0(f_2) < J_0(f).$$

The difference of opinions between the experts can therefore reduce the average

$$\frac{1}{2}J_0(f_1) + \frac{1}{2}J_0(f_2)$$

below each of the averages

$$\frac{1}{2}J_i(f_1) + \frac{1}{2}J_i(f_2).$$

By contrast, consider again the condition we started with. Assume that for each expert we have

$$J_i(f) \geqslant \frac{1}{2}J_i(f_1) + \frac{1}{2}J_i(f_2)$$

when we consider the evaluation by the decision maker, we follow the same reasoning as above: the average of evaluations on the right-hand side can only be lower due to the fact that the decision maker evaluates each act separately. Hence, a decision maker who is uncertainty averse with respect to the experts should satisfy

$$J_0(f) \geqslant \frac{1}{2}J_0(f_1) + \frac{1}{2}J_0(f_2).$$

Observe that the aversion of uncertainty at the level of the experts can also be interpreted as aversion to the divergence of opinions across experts. This is closer to the intuition of Gajdos and Vergnaud [15].

The condition as discussed above considers the averaging of the evaluations of only two acts f_1, f_2 and compares them with a third act f. If the range of the vector function $(J_1(\cdot), \ldots, J_n(\cdot))$ is a convex subset of $\mathbb{R}^n$, this condition will suffice to conclude that φ is concave. But if this

range (which is the domain of φ) fails to be convex, the condition does not suffice for the derived conclusion. Thus, we will require a stronger condition, involving any convex combination of finitely many acts: if each expert evaluates an act f above the weighted average of the evaluations of other acts $f_1, \ldots, f_m$, so should the decision maker. We refer to this condition as Expert Uncertainty Aversion (EUA), as it reflects the decision maker's aversion to her uncertainty about the expert who "has access to truth".

Our main result states that axiom EUA is equivalent to the decision rule proposed above.

2.4. Main result

To state our main axiom, we introduce the following notation. For act $f \in F$ and relation $\succsim_i$, $c_i^f \in F_c$ is a $\succsim_i$-certainty equivalent of f, that is, a constant act such that $f \sim_i c_i^f$. Such certainty equivalents obviously exist, but they will typically not be unique. However, it is easy to see that the validity of following condition is independent of the choice of these certainty equivalents.

Expert Uncertainty Aversion (EUA). For every acts $f \in F$, $f_k \in F$, $k = 1, \ldots, K$, and every numbers $\alpha_k \geqslant 0$ such that $\sum \alpha_k = 1$, if

$$f \succsim_i \sum_k \alpha_k c_i^{f_k} \quad \text{for } i = 1, \ldots, n$$

then

$$f \succsim_0 \sum_k \alpha_k c_0^{f_k}.$$

Observe that, since $c_i^{f_k} \in F_c$ for all i, k, their mixtures, $\sum_k \alpha_k c_i^{f_k}$ are also in F_c. Thus, EUA states that, if each expert thinks that f is at least as desirable as a certain constant act, then the decision maker also thinks that f is at least as desirable as a certain constant act. However, the constant acts involved will typically vary across the different experts, as well as between them and the decision maker. What relates the constant acts on the right-hand side is the fact that each of them is obtained by the same mixture (α_k) of certainty equivalents of the same acts (f_k). Since, however, for each expert and for the decision maker we have a distinct relation $\succsim_i$, these certainty equivalents will typically vary.

It is easy to see that EUA is equivalent to the following condition: for every acts $f \in F$, $f_k \in F$, $k = 1, \ldots, K$, and every numbers $\alpha_k \geqslant 0$ such that $\sum \alpha_k = 1$, if

$$J_i(f) \geqslant \sum_k \alpha_k J_i(f_k) \quad \text{for } i = 1, \ldots, n$$

then

$$J_0(f) \geqslant \sum_k \alpha_k J_0(f_k).$$

The EUA condition clearly implies a more standard condition, namely:

Unanimity. For every $f, g \in F$, if $f \succsim_i g$ for all $i = 1, \ldots, n$, then $f \succsim_0 g$.

Our main result is the following.

Theorem 1. *The following are equivalent*:

(i) $(\succsim_i)_{i=0}^n$ *satisfy* ***EUA***.
(ii) *There is a convex and closed set* $\Lambda \subseteq \Delta(\{1,\dots,n\})$ *such that*

$$J_0(f) = \min_{\lambda\in\Lambda} \sum_{i=1}^{n} \lambda_i J_i(f).$$

Combining this result with the proposition above, we can state

Corollary 2. *The following three conditions are equivalent*:

(i) $(\succsim_i)_{i=0}^n$ *satisfy* ***EUA***.
(ii) *There is a convex and closed set* $\Lambda \subseteq \Delta(\{1,\dots,n\})$ *such that*

$$J_0(f) = \min_{\lambda\in\Lambda} \sum_{i=1}^{n} \lambda_i J_i(f).$$

(iii) *There is a closed and convex set* $\Lambda \subseteq \Delta(\{1,\dots,n\})$ *such that*

$$C_0 = \left\{ p = \sum_{i=1}^{n} \lambda_i p_i \;\middle|\; \lambda \in \Lambda,\ p_i \in C_i \right\}.$$

The set Λ in (ii) or in (iii) is not unique in general. For instance, when all experts' beliefs coincide, condition (i) implies that $J_0 = J_i$ for each $i = 1,\dots,n$. Then the conclusion of either (ii) or (iii) holds irrespective of the set Λ.

3. Discussion and related literature

In terms of motivation and content, this paper is close to Gajdos and Vergnaud [15]. As mentioned above, they also deal with the aggregation of beliefs of uncertainty averse experts, and their suggested rule also takes the form suggested in (3). However, their model and axiomatization differ from the present ones in several ways. First, Gajdos and Vergnaud [15] follow the set-up of Gajdos, Hayashi, Tallon and Vergnaud [14] in allowing sets of probabilities to be a component of the object of choice. Whereas Gajdos et al. [14] deal with a single decision maker, who has a single set of probabilities, Gajdos and Vergnaud [15] deal with two experts and with two sets of probabilities, one for each expert. Thus, the decision maker has preferences over triples of the form (f, P, Q) where f is an act, P is the set of probabilities of the first expert, and Q – of the second. In this set-up, Gajdos and Vergnaud derive the same aggregation rule for two experts who are treated symmetrically (resulting in a set Λ that is an interval in $[0, 1]$ with a midpoint at $1/2$).[8] They also provide a definition and characterization of the relation "decision maker 1 is more averse to conflict (between experts' opinions) than decision maker 2".

A related strand in the literature has to do with "judgment aggregation". This title refers to the aggregation of binary opinions, to be viewed as truth functions over a set of propositions. Starting with an equivalent of the Condorcet paradox, List and Pettit [28] derive an Arrow-style

[8] The two papers were independently developed.

impossibility theorem. Many further impossibility results were derived since.[9] Our approach differs from the judgment aggregation approach in that (i) we assume that beliefs are modeled as probabilities, or sets thereof, rather than as truth functions; and (ii) by considering the average as a basic aggregation procedure, we do not follow an "independence" axiom that is common in this literature and that is essential for the impossibility results.[10]

Chambers and Echenique [7] also study the aggregation of uncertainty averse preferences. In their case, however, the objects of choice are allocations of an aggregate bundle among the members of a household, so that the problem involves the aggregation of preferences and not only of beliefs.

Appendix A. Proofs

A.1. Proof of Proposition 1

We first assume that (6) holds and prove (5). Let there be given $f \in F$. Choose $\lambda^f \in \Lambda$ to minimize $\sum_{i=1}^n \lambda_i J_i(f)$, and choose $p_i^f \in C_i$ for $i = 1, \ldots, n$ such that $J_i(f) = \int u(f)\, dp_i^f$.

Clearly, $p_0^f \equiv \sum_{i=1}^n \lambda_i^f p_i^f \in C_0$. Hence

$$J_0(f) \leqslant \int u(f)\, dp_0^f = \sum_{i=1}^n \lambda_i^f \int u(f)\, dp_i^f = \min_{\lambda \in \Lambda} \sum_{i=1}^n \lambda_i J_i(f).$$

To see the converse inequality, let $p_0^f \in C_0$ be such that

$$J_0(f) = \int u(f)\, dp_0^f.$$

By (6), there exist $p_i^f \in C_i$ for $i \geqslant 1$, and $\lambda^f \in \Lambda$, such that $p_0^f \equiv \sum_{i=1}^n \lambda_i^f p_i^f$. Hence

$$\sum_{i=1}^n \lambda_i^f J_i(f) \leqslant \sum_{i=1}^n \lambda_i^f \int u(f)\, dp_i^f = \int u(f)\, dp_0^f = J_0(f).$$

Next assume that (5) holds. Define a set of measures $\hat{C}_0 = \{p = \sum_{i=1}^n \lambda_i p_i \mid p_i \in C_i,\ \lambda \in \Lambda\}$ and a corresponding functional $\hat{J}$, namely

$$\hat{J}(f) := \min_{p \in \hat{C}_0} \int u(f)\, dp.$$

The first part of the proof ((6) implies (5)) can be invoked to verify that $\hat{J}(f) = \min_{\lambda \in \Lambda} \sum_{i=1}^n \lambda_i J_i(f)$ for all $f \in F$. Hence $\hat{J} = J_0$. Uniqueness of the maxmin representation then implies $\hat{C}_0 = C_0$.

---

[9] See a forthcoming Symposium in the *Journal of Economic Theory*.

[10] In the absence of probabilities one may still use averaging, say, of relative rankings. However, this type of aggregation of opinions would violate independence-like axioms, much as Borda's rule for the aggregation of preferences violates Arrow's independence axiom.

Another non-probabilistic approach to uncertainty aversion has been suggested by Bossert [5] and Bossert and Slinko [6].

A.2. Proof of Theorem 1

The fact that (ii) implies (i) is immediate. We therefore turn to the converse direction, namely, that (i) implies (ii).

We denote by $B = B(J)$ the range of the vector $J = (J_1, \ldots, J_n)$ of the experts' evaluations and assume, without loss of generality, that u assumes the value zero. It is an immediate consequence of Unanimity that, for a given act f, $J_0(f)$ is only a function of $J(f)$. We state this fact as a lemma.

Lemma 1. *There is a function $\phi : B \to \mathbf{R}$ such that, for each $f \in F$,*

$$J_0(f) = \phi\big(J(f)\big).$$

Proof. Indeed, let $f, g \in F$ be two acts such that $J(f) = J(g)$. By Unanimity, and since $f \sim_i g$ for each i, one has both $f \succsim_0 g$ and $g \succsim_0 f$. □

The proof of the implication (i) ⇒ (ii) is organized along the following steps.

We prove first that ϕ is monotone (non-decreasing), homogeneous and satisfies Shift whenever defined.[11] We then perform three extensions of ϕ to supersets of B. First, we consider the positive cone spanned by B, *cone*(B), extend ϕ to this positive cone by homogeneity, and prove that it retains the three properties (monotonicity, homogeneity, and Shift). Second, we consider the convex hull of *cone*(B), to be denoted by D. We show that D is a convex cone invariant under translation along the unit vector. We extend ϕ to a concave function ψ on D, using axiom EUA. We then continue to prove that ψ retains monotonicity, homogeneity, and Shift on D. Finally, we further extend ψ from D to all of $\mathbb{R}^n$, retaining monotonicity, concavity, homogeneity, and Shift. Finally, we apply the logic of the proof of GS [23] to show that there is a closed set $\Lambda \subseteq \Delta(\{1, \ldots, n\})$ such that, for all $(x_1, \ldots, x_n) \in \mathbb{R}^n$,

$$\psi(x_1, \ldots, x_n) = \min_{\lambda \in \Lambda} \sum_{i=1}^{n} \lambda_i x_i.$$

Lemma 2. *For $b \in B$ and $0 < \alpha < 1$, we have $\alpha b \in B$. The map ϕ is monotone and homogeneous on B.*

Proof. We start with monotonicity. Let $a, b \in B$ be given, such that $a \leqslant b$. Let f and g be two acts such that $J(f) = a$ and $J(g) = b$. It follows from Unanimity that $J_0(f) \leqslant J_0(g)$. This implies that ϕ is monotone.

Let now $b \in B$ and $0 < \alpha \leqslant 1$ be given. We wish to prove that $\alpha b \in B$ and that $\phi(\alpha b) = \alpha\phi(b)$. Let g be an act such that $J(g) = b$, let $z \in L$ be such that $u(z) = 0$, and define an act f by $f = \alpha g + (1 - \alpha) f_z$, where f_z is the constant act that yields z in each state. Since u is affine, one has $u(f(s)) = \alpha u(g(s))$, for each state s. It follows from the functional forms of $J_0, J_1, \ldots, J_n$ that $J_0(f) = \alpha J_0(g)$ and that $J_i(f) = \alpha J_i(g)$ for each expert i. This means that $\alpha b = J(f) \in B$.

[11] Throughout the proof, "monotonicity" refers to weak monotonicity, that is, to the function being non-decreasing.

Furthermore,

$$\phi(\alpha b) = \phi\big(J(f)\big) = J_0(f) = \alpha J_0(g) = \alpha\phi\big(J(g)\big) = \alpha\phi(b),$$

as desired. □

We let $cone(B)$ stand for the cone spanned by B, and still denote by ϕ the extension of ϕ (by homogeneity) to the set $cone(B)$.

Lemma 3. *The extension of the map ϕ to $cone(B)$ is monotone and homogeneous.*

Proof. Homogeneity is immediate from the definition. To see that monotonicity holds, assume that $a, b \in B$ are such that, for some $\alpha, \beta > 0$, $\alpha a \leqslant \beta b$. In light of Lemma 2, there is no loss of generality in assuming that $\beta \geqslant 1$. Consider the point $d = \frac{\alpha}{\beta}a$ and observe that $d \leqslant b$. Choose $\gamma > 0$ such that $\gamma < \min(1, \frac{\beta}{\alpha})$. Thus, $\gamma d, \gamma b \in B$. We also have $\gamma d \leqslant \gamma b$, and thus $\phi(\gamma d) \leqslant \phi(\gamma b)$ by monotonicity of ϕ on B. Homogeneity then yields $\phi(d) \leqslant \phi(b)$ and $\phi(\alpha a) \leqslant \phi(\beta b)$. □

Next we wish to show that ϕ satisfies Shift, and that $cone(B)$ is closed under shifts. More explicitly:

Lemma 4. *For all $c \in \mathbb{R}$,*

- $\vec{c} + cone(B) \subseteq cone(B)$.
- *For all $x \in cone(B)$, one has $\phi(x + \vec{c}) = \phi(x) + c$.*

Proof. We prove the two statements together. Let $x \in cone(B)$, and $c \in \mathbb{R}$. We will prove that $x + \vec{c} \in cone(B)$, and that $\phi(x + \vec{c}) = \phi(x) + c$. By homogeneity, we may assume without loss of generality that $2x \in B$ and that $2c$ is in the range of u. Let $f \in F$ be an act such that $J(f) = 2x$, and let $z \in L$ be such that $u(z) = 2c$. Let $h \in F$ be the act defined by $h = \frac{1}{2}f + \frac{1}{2}f_z$. Since u is affine, and using the functional form of the J_i's, one has

$$J_i(h) = \frac{1}{2}J_i(f) + c, \tag{8}$$

for all experts, $i = 1, \ldots, n$, as well as for the decision maker, $i = 0$.

Applying (8) to all experts yields $J(h) = \frac{1}{2}J(f) + \vec{c} = x + \vec{c}$. In particular, $x + \vec{c} \in cone(B)$, and $\phi(x + \vec{c}) = \phi(J(h)) = J_0(h)$. On the other hand, and when applying (8) to the decision maker, one gets

$$J_0(h) = \frac{1}{2}J_0(f) + c = \frac{1}{2}\phi\big(J(f)\big) + c = \frac{1}{2}\phi(2x) + c = \phi(x) + c,$$

where the last equality holds since ϕ is homogeneous. Combining both equalities yields $\phi(x + \vec{c}) = \phi(x) + c$. □

We now come to the main part of the proof, which is the extension of ϕ to a convex domain, so that it be concave on it. To highlight this step, we change the notation of the domain (to D) and the function (to ψ).

Let, then, D stand for the convex hull of the $cone(B)$. Define $\psi : D \to \mathbf{R}$ as follows. For $x \in D$, we set $\psi(x) := \sup \sum_k \alpha_k \phi(x_k)$, where the supremum is taken over all finite families $x_k \in cone(B), \alpha_k \geqslant 0$ $(k = 1, \ldots, K)$, such that $\sum_k \alpha_k = 1$ and $x \geqslant \sum_k \alpha_k x_k$. The following lemma shows, by standard arguments, that ψ is monotone and concave, and, using the EUA axiom, that ϕ coincides with ψ wherever (ϕ is) defined. Observe that if ϕ were already known to be defined on a convex domain, one could use EUA directly to prove that ϕ is concave.

Lemma 5. *The map ψ coincides with ϕ on $cone(B)$, and it is monotone and concave on D.*

Proof. We start with the first statement. Observe first that for every $x \in cone(B)$, there exists $\eta_* \in (0, 1]$, such that $\eta_* x \in B$. Then, as Lemma 2 states, $\eta x \in B$, for *every* $\eta \in (0, \eta_*]$.

Let $x \in cone(B)$ be given. Clearly, $\psi(x) \geqslant \phi(x)$. We now prove that $\phi(x) \geqslant \psi(x)$. Let a finite family $x_k \in cone(B)$, $\alpha_k \geqslant 0$ $(k = 1, \ldots, K)$ be given, with $\sum_k \alpha_k = 1$ and $x \geqslant \sum_k \alpha_k x_k$. There is $\eta > 0$ small enough such that $\eta x \in B$, and $\eta x_k \in B$, for each k. Let $f, f_k \in K$ be acts such that $J(f) = \eta x$ and $J(f_k) = \eta x_k$ for each k. Since $x \geqslant \sum_k \alpha_k x_k$, and using **EUA**, one has $J_0(f) \geqslant \sum_k \alpha_k J_0(f_k)$, that is, $\phi(\eta x) \geqslant \sum_k \alpha_k \phi(\eta x_k)$. By homogeneity, $\phi(x) \geqslant \sum_k \alpha_k \phi(x_k)$. Since the family (x_k, α_k) is arbitrary, the inequality $\phi(x) \geqslant \psi(x)$ follows, as desired.

To see that ψ is monotone, assume that $x' \geqslant x$. Then the set of points and weights $\{(x_k), (\alpha_k)\}$ used in the definition of $\psi(x')$ is a superset of the corresponding set for $\psi(x)$, and the supremum over the former can only be larger than the supremum over the latter.

We now prove that ψ is concave. Let $x, x' \in D$ be given. We will prove that $\psi(\frac{1}{2}x + \frac{1}{2}x') \leqslant \frac{1}{2}\psi(x) + \frac{1}{2}\psi(x')$. Let two finite families $x_k \in B, \alpha_k \geqslant 0$ $(k = 1, \ldots, K)$ and $x'_l \in B, \alpha'_l \geqslant 0$ $(l = K+1, \ldots, K')$ be given, with $\sum_k \alpha_k = \sum_l \alpha'_l = 1$, $x \geqslant \sum_k \alpha_k x_k$ and $x' \geqslant \sum_l \alpha'_l x'_l$. Consider the finite family $\tilde{x}_k \in B$, $\tilde{\alpha}_k \geqslant 0$, $k = 1, \ldots, K'$, where $\tilde{x}_k = x_k$ if $k \leqslant K$, and $\tilde{x}_k = x'_k$ if $k > K$, while $\tilde{\alpha}_k = \frac{1}{2}\alpha_k$ if $k \leqslant K$, and $\tilde{\alpha}_k = \frac{1}{2}\alpha'_k$ if $k > K$.

Plainly,

$$\frac{1}{2}x + \frac{1}{2}x' \geqslant \sum_k \tilde{\alpha}_k \tilde{x}_k,$$

hence

$$\psi\left(\frac{1}{2}x + \frac{1}{2}x'\right) \geqslant \sum_k \tilde{\alpha}_k \phi(\tilde{x}_k) = \frac{1}{2}\sum_{k=1}^{K} \alpha_k \phi(x_k) + \frac{1}{2}\sum_{l=K+1}^{K'} \alpha'_l \phi\big(x'_l\big).$$

Taking the supremum over families $x_k, \alpha_k, x'_l, \alpha'_l$ then yields

$$\psi\left(\frac{1}{2}x + \frac{1}{2}x'\right) \geqslant \frac{1}{2}\psi(x) + \frac{1}{2}\psi\big(x'\big),$$

as desired. □

The proof of the next lemma is routine, and is given in Appendix B.

Lemma 6. *The map ψ is homogeneous, and satisfies Shift.*

We now have a convex cone D containing the diagonal, and a function ψ on it that is monotone, concave, homogeneous and satisfies Shift. The set D is a superset of B, and ψ is an extension of ϕ. Thus, a minimum-average representation of ψ will serve as a minimum-average

representation of ϕ. To obtain such a representation, we wish to know that D is of full dimensionality. Since this is not guaranteed to be the case, we further extend the domain and the function.

Lemma 7. *The function ψ can be extended to $\mathbb{R}^n$ retaining monotonicity, concavity, homogeneity, and Shift.*

Proof. Extend ψ to $\mathbb{R}^n$ by defining

$$\psi'(y) = \sup_{x\in D,\, x\leqslant y} \psi(x).$$

To see that ψ' is well defined, consider $y \in \mathbb{R}^n$. Denote

$$y^* = \max_{i\leqslant n} y_i,$$
$$y_* = \min_{i\leqslant n} y_i$$

and observe that $\overrightarrow{y_*} \leqslant y \leqslant \overrightarrow{y^*}$. Because $\overrightarrow{y_*} \in D$, the set $\{x \in D | x \leqslant y\}$ is non-empty and $\psi'(y) \geqslant \psi(\overrightarrow{y_*}) = y_*$. On the other hand, any $x \in D$ that satisfies $x \leqslant y$ also satisfies $x \leqslant \overrightarrow{y^*}$, where $\overrightarrow{y^*} \in D$ and $\psi(\overrightarrow{y^*}) = y^*$. Hence (by monotonicity of ψ on D) any such x satisfies $\psi(x) \leqslant \psi(\overrightarrow{y^*}) = y^*$. This means that $\psi'(y)$ is well defined and that it satisfies $y_* \leqslant \psi'(y) \leqslant y^*$.

Next, since ψ is known to be non-decreasing on D, $\psi'(y) = \psi(y)$ for $y \in D$. Hence ψ' is an extension of ψ, and for simplicity we will denote it also by ψ.

We claim that ψ continues to satisfy the following properties on all of $\mathbb{R}^n$: (i) monotonicity; (ii) concavity; (iii) homogeneity; (iv) Shift. See Appendix B for details. □

We finally repeat the GS argument.

Lemma 8. *There is a closed and convex set $\Lambda \subseteq \Delta(\{1,\dots,n\})$ such that, for all $x \in \mathbb{R}^n$,*

$$\psi(x) = \min_{\lambda\in\Lambda} \sum_{i=1}^{n} \lambda_i x_i.$$

Proof. Consider $x^* \in \mathbb{R}^n$. Using a supporting hyperplane theorem, there is $(q^*, \gamma^*) \in \mathbb{R}^n \times \mathbb{R}$ such that

$$\langle q^*, x\rangle + \gamma^* \geqslant \psi(x), \quad \forall x \in \mathbb{R}^n,$$
$$\langle q^*, x^*\rangle + \gamma^* = \psi(x^*).$$

Monotonicity implies that $q \geqslant 0$, for if $q_i < 0$ for some $i \leqslant n$, we get $\psi(x^* + e_i) < \psi(x^*)$ where e_i is the i-th unit vector, contradicting monotonicity.

Next, set $c = \psi(x^*)$ and observe that ψ is constant along the line

$$\{\alpha x^* + (1-\alpha)\overrightarrow{c} \mid \alpha \in \mathbb{R}\}.$$

We conclude that $\langle q^*, x\rangle$ also has to be constant along this line. And this implies that $\langle q^*, \overrightarrow{c}\rangle = c$ and q^* is a probability vector in $\Delta(\{1,\dots,n\})$. Finally, homogeneity implies that $\gamma^* = 0$.

Thus, for $x^* \in \mathbb{R}^n$ we have established the existence of $q^* \in \Delta(\{1,\dots,n\})$ such that

$$\begin{aligned} &\langle q^*, x\rangle \geqslant \psi(x), \quad \forall x \in \mathbb{R}^n, \\ &\langle q^*, x^*\rangle = \psi(x^*). \end{aligned} \tag{9}$$

Finally, we define

$$\Lambda := clco\big(\{q^* \mid x^* \in \mathbb{R}^n\}\big)$$

to be the closed convex hull of the corresponding set of vectors q^*, and we observe that, for every $x \in \mathbb{R}^n$,

$$\psi(x) = \min_{\lambda\in\Lambda} \sum_{i=1}^{n} \lambda_i x_i. \qquad \square$$

In particular, $\phi(x) = \min_{\lambda\in\Lambda} \sum_{i=1}^{n} \lambda_i x_i$ for all $x \in B$. This concludes the proof of (ii).

Appendix B. Proof details

B.1. Proof of Lemma 6

We define ψ by $\psi(x) := \sup \sum_k \alpha_k \phi(x_k)$, where the supremum is taken over all finite families $x_k \in cone(B)$, $\alpha_k \geqslant 0$ $(k = 1,\dots,K)$, such that $\sum_k \alpha_k = 1$ and $x \geqslant \sum_k \alpha_k x_k$.

Homogeneity: Let there be given x and $\beta > 0$. Since β may be larger or smaller than 1, it suffices to show that

$$\psi(\beta x) \geqslant \beta\psi(x).$$

Let there be given $x_k \in cone(B)$, $\alpha_k \geqslant 0$ $(k = 1,\dots,K)$, such that $\sum_k \alpha_k = 1$ and $x \geqslant \sum_k \alpha_k x_k$. Define $x'_k = \beta x_k \in cone(B)$. Then

$$\beta x \geqslant \sum_k \alpha_k \beta x_k = \sum_k \alpha_k x'_k$$

and it follows that

$$\psi(\beta x) \geqslant \sum_k \alpha_k \phi\big(x'_k\big) = \sum_k \alpha_k \phi(\beta x_k) = \beta \sum_k \alpha_k \phi(x_k).$$

As this holds for any pair of sequences $\{x_k\}$, $\{\alpha_k\}$, and $\psi(x)$ is the supremum over the respective $\sum_k \alpha_k \phi(x_k)$, the conclusion follows.

Shift: Let there be given x and $c \in \mathbb{R}$. Since c may be positive or negative, it suffices to show that

$$\psi(x + \vec{c}) \geqslant \psi(x) + c.$$

Consider $x_k \in cone(B)$, $\alpha_k \geqslant 0$ $(k = 1,\dots,K)$, such that $\sum_k \alpha_k = 1$ and $x \geqslant \sum_k \alpha_k x_k$. Define $x'_k = x_k + \vec{c} \in cone(B)$. Observe that

$$x + \vec{c} \geqslant \sum_k \alpha_k x'_k = \sum_k \alpha_k (x_k + \vec{c})$$

hence

$$\psi(x+\vec{c}) \geqslant \sum_k \alpha_k \phi(x_k+\vec{c}) = \sum_k \alpha_k \phi(x_k) + c.$$

As this holds for any pair of sequences $\{x_k\}$, $\{\alpha_k\}$, and $\psi(x)$ is the supremum over the respective $\sum_k \alpha_k \phi(x_k)$, the conclusion follows.

B.2. Details of proof of Lemma 7

We show that ψ satisfies monotonicity, concavity, homogeneity and Shift:

Monotonicity: Consider $y, z \in \mathbb{R}^n$ such that $y \geqslant z$. Then any $x \in D$ such that $x \leqslant z$ also satisfies $x \leqslant y$ and it follows that $\psi(y) \geqslant \psi(z)$.

Concavity: Let $y, z \in \mathbb{R}^n$ and $\alpha \in [0,1]$. We wish to show that

$$\psi\big(\alpha y + (1-\alpha)z\big) \geqslant \alpha\psi(y) + (1-\alpha)\psi(z).$$

For $\varepsilon > 0$, let $x_y, x_z \in D$ be such that $y \geqslant x_y$, $z \geqslant x_z$ and

$$\psi(y) < \psi(x_y) + \varepsilon,$$
$$\psi(z) < \psi(x_z) + \varepsilon.$$

This means that

$$\alpha\psi(y) + (1-\alpha)\psi(z) < \alpha\psi(x_y) + (1-\alpha)\psi(x_z) + \varepsilon.$$

Since D is convex, $\alpha x_y + (1-\alpha)x_z \in D$. Because ψ is concave on D,

$$\psi\big(\alpha x_y + (1-\alpha)x_z\big) \geqslant \alpha\psi(x_y) + (1-\alpha)\psi(x_z)$$

and thus

$$\psi\big(\alpha x_y + (1-\alpha)x_z\big) > \alpha\psi(y) + (1-\alpha)\psi(z) - \varepsilon.$$

Next, observe that $y \geqslant x_y$, $z \geqslant x_z$ imply

$$\alpha y \geqslant \alpha x_y,$$
$$(1-\alpha)z \geqslant (1-\alpha)x_z$$

and

$$\alpha y + (1-\alpha)z \geqslant \alpha x_y + (1-\alpha)x_z.$$

By definition of ψ,

$$\psi\big(\alpha y + (1-\alpha)z\big) \geqslant \psi\big(\alpha x_y + (1-\alpha)x_z\big)$$

and thus

$$\psi\big(\alpha y + (1-\alpha)z\big) > \alpha\psi(y) + (1-\alpha)\psi(z) - \varepsilon$$

for all $\varepsilon > 0$, which means that ψ is concave.

Homogeneity: Consider $y \in \mathbb{R}^n$ and $\alpha > 0$. We wish to show that

$$\psi(\alpha y) = \alpha\psi(y).$$

It suffices to show

$$\psi(\alpha y) \geqslant \alpha\psi(y)$$

(because α can be larger or smaller than 1). Consider, then, $\varepsilon > 0$ and $x_y \in D$ such that $y \geqslant x_y$ and

$$\psi(y) < \psi(x_y) + \varepsilon.$$

Since ψ is homogeneous on D,

$$\psi(\alpha x_y) = \alpha\psi(x_y).$$

Moreover,

$$\alpha y \geqslant \alpha x_y$$

and thus

$$\psi(\alpha y) \geqslant \psi(\alpha x_y) = \alpha\psi(x_y) > \alpha\psi(y) - \alpha\varepsilon.$$

This being true for all $\varepsilon > 0$, we conclude that $\psi(\alpha y) \geqslant \alpha\psi(y)$.

Shift: Let there be given $y \in \mathbb{R}^n$ and $c \in \mathbb{R}$. We wish to show that

$$\psi(y + \vec{c}) = \psi(y) + c.$$

It suffices to show that (for all $y \in \mathbb{R}^n$ and $c \in \mathbb{R}$)

$$\psi(y + \vec{c}) \geqslant \psi(y) + c$$

observe that, for every $x \in D$ such that $x \leqslant y$, we have $x + \vec{c} \in D$ and $x + \vec{c} \leqslant y + \vec{c}$. Thus,

$$\psi(y + \vec{c}) \geqslant \psi(x + \vec{c}) = \psi(x) + c$$

where the last equality follows from the fact that ψ satisfies Shift on D. Taking the supremum on the right-hand side, we obtain the conclusion that $\psi(y + \vec{c}) \geqslant \psi(y) + c$.

References

[1] F.J. Anscombe, R.J. Aumann, A definition of subjective probability, Ann. Math. Statist. 34 (1963) 199–205.
[2] M. Bacharach, Normal Bayesian dialogues, J. Amer. Statistical Assoc. 74 (1979) 837–846.
[3] T. Bewley, Knightian decision theory: Part I, Decis. Econ. Finance 25 (2002) 79–110 (Working paper, 1986).
[4] C. Blackorby, D. Donaldson, P. Mongin, Social aggregation without the expected utility hypothesis, mimeo, 2000.
[5] W. Bossert, Uncertainty aversion in nonprobabilistic decision models, Math. Soc. Sci. 34 (1997) 191–203.
[6] W. Bossert, A. Slinko, Relative uncertainty aversion and additively representable set rankings, Int. J. Econ. Theory 2 (2006) 105–122.
[7] C.P. Chambers, F. Echenique, When does aggregation reduce uncertainty aversion, mimeo, 2009.
[8] C.P. Chambers, T. Hayashi, Preference aggregation under uncertainty: Savage vs. Pareto, Games Econ. Behav. 54 (2006) 430–440.
[9] B. de Finetti, Sul significato soggettivo della probabilità, Fund. Math. 17 (1931) 298–329.
[10] B. de Finetti, La prevision: ses lois logiques, ses sources subjectives, Ann. Inst. H. Poincaré 7 (1937) 1–68.
[11] D. Ellsberg, Risk, ambiguity and the Savage axioms, Quart. J. Econ. 75 (1961) 643–669.
[12] P.C. Fishburn, Utility Theory for Decision Making, John Wiley and Sons, 1970.

[13] T. Gajdos, J.-M. Tallon, J.-C. Vergnaud, Representation and aggregation of preferences under uncertainty, J. Econ. Theory 141 (2008) 68–99.
[14] T. Gajdos, T. Hayashi, J.-M. Tallon, J.-C. Vergnaud, Attitude toward imprecise information, J. Econ. Theory 140 (2008) 27–65.
[15] T. Gajdos, J.-C. Vergnaud, Decisions with conflicting and imprecise information, mimeo, 2009.
[16] C. Genest, J. Zidek, Combining probability distributions: A critique and an annotated bibliography, Statist. Sci. 1 (1986) 68–99.
[17] I. Gilboa, Theory of Decision under Uncertainty, Econ. Soc. Monogr. Ser., Cambridge University Press, 2009.
[18] I. Gilboa, M. Marinacci, Ambiguity and the Bayesian approach, in: Advances in Economics and Econometrics: Theory and Applications, Tenth World Congress of the Econometric Society, forthcoming.
[19] I. Gilboa, A. Postlewaite, D. Schmeidler, Probabilities in economic modeling, J. Econ. Perspect. 22 (2008) 173–188.
[20] I. Gilboa, A. Postlewaite, D. Schmeidler, Is it always rational to satisfy Savage's axioms?, Econ. Philos. 25 (2009) 285–296.
[21] I. Gilboa, A. Postlewaite, D. Schmeidler, Rationality of belief, Synthese (2010), forthcoming.
[22] I. Gilboa, D. Samet, D. Schmeidler, Utilitarian aggregation of beliefs and tastes, J. Polit. Economy 112 (2004) 932–938.
[23] I. Gilboa, D. Schmeidler, Maxmin expected utility with a non-unique prior, J. Math. Econ. 18 (1989) 141–153.
[24] J.C. Harsanyi, Cardinal welfare, individualistic ethics, and interpersonal comparison of utility, J. Polit. Economy 63 (1955) 309–321.
[25] A. Hylland, R. Zeckhauser, The impossibility of Bayesian group decision making with separate aggregation of beliefs and values, Econometrica 6 (1979) 1321–1336.
[26] P. Klibanoff, M. Marinacci, S. Mukerji, A smooth model of decision making under ambiguity, Econometrica 73 (2005) 1849–1892.
[27] F.H. Knight, Risk, Uncertainty, and Profit, Houghton Mifflin, Boston, New York, 1921.
[28] C. List, P. Pettit, Aggregating sets of judgments: An impossibility result, Econ. Philos. 18 (2002) 89–110.
[29] K.J. McConway, Marginalization and linear opinion pools, J. Amer. Statistical Assoc. 76 (1981) 410–414.
[30] P. Mongin, Consistent Bayesian aggregation, J. Econ. Theory 66 (1995) 313–351.
[31] R.F. Nau, Uncertainty aversion with second-order utilities and probabilities, Manage. Sci. 52 (2006) 136–145.
[32] F.P. Ramsey, Truth and probability, in: The Foundation of Mathematics and Other Logical Essays, Harcourt, Brace and Co., New York, 1931.
[33] L.J. Savage, The Foundations of Statistics, John Wiley and Sons, New York, 1954; second addition in Dover, 1972.
[34] D. Schmeidler, Integral representation without additivity, Proc. Amer. Math. Soc. 97 (1986) 255–261.
[35] D. Schmeidler, Subjective probability and expected utility without additivity, Econometrica 57 (1989) 571–587.
[36] K. Seo, Ambiguity and second-order belief, Econometrica 67 (2009) 1575–1605.
[37] M. Stone, The linear opinion pool, Ann. Math. Statist. 32 (1961) 1339–1342.
[38] J. von Neumann, O. Morgenstern, Theory of Games and Economic Behavior, Princeton University Press, Princeton, NJ, 1944.
[39] S. Zuber, Harsanyi's theorem without the sure-thing principle, mimeo, 2009.