DISAGREEMENT AVERSION

Antoine Bommier*, Adrien Fabre*, Arnaud Goussebaïle* and Daniel Heyen‡

October 16, 2021

WORK IN PROGRESS. DO NOT CIRCULATE NOR CITE.

Link to latest version

Abstract

Decision-makers rely on experts who often disagree. Aversion to expert disagreement is usually modelled with ambiguity-averse preferences which rest on a unanimity principle: if according to all experts one choice is better than another, so should the decision-maker. Such unanimity among experts however can be spurious, masking substantial disagreement on the underlying reasons. We introduce a novel notion of disagreement aversion to distinguish spurious from genuine unanimity. We develop a model centered around the cautious aggregation of expert beliefs that is able to capture that novel notion of disagreement aversion. We provide formal results and illustrate them in applications.

JEL: D81; D83; D71.

Keywords: Disagreement Aversion; Uncertainty Aversion; Ambiguity; Belief aggregation.

---

*ETH Zürich. Fabre: afabre@ethz.ch (corresponding author).

†Technische Universität Kaiserslautern.

‡We are grateful to Adam Dominiak, Peter Wakker, and Hans Gersbach for their generous comments. We also thank the seminar participants at the Paris School of Economics and ETH Zürich. We are grateful to Rémi Peyre for sharing mathematical insights. We are grateful to Carmel Connell for the proofreading.

Experts disagree. They disagree on topics as different as climate change, macroeconomic outlooks, and the Covid-19 pandemic. The diversity of scientific opinions forces decision-makers to make choices without precise knowledge of the likelihood of the possible outcomes. In such a situation, decision-makers may be averse to expert disagreement and prefer "safer" actions on which experts differ less. This paper introduces a novel notion of aversion to expert disagreement and illustrates its implications.

To fix ideas, consider the following simple example summarized in Table 1. The CEO of a firm needs to decide whether to implement a project or not. The project's net return depends on contingencies the CEO has no control over: the net return is $+10$ in the "good" state, 0 in the "neutral" state, and -10 in the "bad" state (we assume that the CEO is risk-neutral, or that the numbers in Table 1 are utils). The CEO consults two experts for their views about the likelihood of each state. Whereas the first expert is convinced that the neutral state will occur for sure, the second expert assigns a null probability to this event and thinks that the good and bad states are equally likely. The interesting aspect in this example is that, according to both experts, implementing the project is as good as not implementing it, since both options provide the same expected utility. However, the experts provide very different reasons for this conclusion. Intuitively, the CEO may infer from expert disagreement that there is no robust understanding of what the project's outcomes could be and consequently prefer the no project option about which experts agree.

Table 1 – Leading example of a CEO's decision whether to implement a project. Payoffs are in utils.

	Payoffs		Beliefs	
	No project	Project	Expert 1	Expert 2
State 1 ("bad")	0	-10	0	0.5
State 2 ("neutral")	0	0	1	0
State 3 ("good")	0	10	0	0.5

There are several papers in the decision theory literature on how to take decisions when facing imprecise or conflicting information. A key reference is Gajdos et al. (2008) who model the relation between received information and subjective beliefs used by the decision-maker, proposing a non-Bayesian framework that allows for aversion toward imprecise information. Extensions by Crès, Gilboa and Vieille (2011) and Gajdos and Vergnaud (2013) include the possibility that experts themselves are uncertain about their probability judgment ("within-agent ambiguity", Baillon, Cabantous and Wakker (2012)), so that the information available to the decision-maker consists of several sets of probabilities. Other contributions in the decision theory literature specifically concerned with expert disagreement are Nascimento (2012), Stanca (2021), Hill (2012) and Mongin (1995, 1998). The landscape is completed by more applied contributions, which use multiple-prior ambiguity models for the aggregation of heterogeneous experts' beliefs (see for example the review in Berger et al. (2021)). An interesting aspect however is that all these papers, whether in decision theory or the applied literature, keep as a fundamental assumption a unanimity principle: if according to all experts implementing the project is just as good as doing nothing, the decision-maker should be indifferent between the two options as well.[1] That experts may provide contradictory reasons – as is the case in our example – for why implementing the project is just as good as doing nothing has no impact on the decision-maker's willingness to implement the project.

A few papers have criticized this unanimity principle, with Mongin (2016) and Gilboa, Samet and Schmeidler (2004) as notorious examples.[2] Both challenge the unanimity principle (or Pareto condition) when people hold different beliefs: "Society should not necessarily endorse a unanimous choice when it is based on contradictory beliefs" (Gilboa, Samet and Schmeidler 2004). In particular, Mongin (2016) argues that "spurious unanimity", i.e. agreement on expressed preferences – such as both experts' indifference between implementing and not implementing the project in the introductory example – need not be preserved on the social level when there is disagreement on the underlying reason. Responding to this criticism, we propose a framework able to deal with spurious agreement among experts. As in Gajdos et al. (2008), we consider the case where experts all have precise beliefs

[1]Various terminology is used to refer to this unanimity principle. For example Gajdos et al. (2008) calls it "dominance", it is named "unanimity" in Crès, Gilboa and Vieille (2011) while Stanca (2021) refers to it as the "Pareto condition".

[2]Other examples include Kets and Siniscalchi (2016); Machina (2014); Skiadas (2013).

and therefore focus on “conflicting information” (Gajdos and Vergnaud 2013), i.e. disagreement *between* experts.

This paper makes three contributions. The first contribution is to formalize a novel notion of disagreement aversion that has traction even in the presence of spurious agreement, as in our introductory example. Intuitively, disagreement aversion is a tendency to prefer choices on which experts have consensus. One possible interpretation of "consensus" on a choice option is that all experts agree on the resulting expected utility. However, such *utility-consensus* may be “spurious” and obtained in spite of fundamental heterogeneity in expert views. For this reason, we introduce a stronger notion of consensus, *distribution-consensus*, which requires that all experts not only agree on the expected utility, but even have consensus on the distribution of outcomes. The different notions of aversion to experts disagreement defined in the literature, such as “ambiguity aversion” or “imprecision aversion”, all rest on the unanimity principle and are thus intimately tied to the concept of utility-consensus. The existing notions of disagreement aversion, therefore, do not have bite if experts have profound disagreement but happen to agree on the expected utility of a choice option. Our novel notion of disagreement aversion, in contrast, captures a decision-maker’s aversion to a lack of distribution-consensus. Going back to the introductory example, both the “no project” and the “project” options are utility-consensual; therefore, under any model that fulfills the unanimity principle, the decision-maker is necessarily indifferent between both options. In contrast, only the “no project” option is distribution-consensual. Accordingly, as experts have no distribution-consensus on the project option, a decision-maker averse to disagreement (in our sense) prefers to abstain from the project.

Our second contribution is to introduce a model that allows for disagreement aversion. Like most previous contributions, we borrow from the ambiguity aversion literature. Since disagreement aversion requires abandoning the unanimity principle, we cannot rely on the most common ambiguity models. Indeed, those "monotone models" assume a property of monotonicity (in terms of states) which in our setting translates to the unanimity principle (in terms of experts). We instead build upon the dual model of ambiguity aversion by Bommier (2017), yielding a “probability-averaging model”. In such a model the decision-maker evaluates a choice option based on an *equivalent* probability distribution of its outcomes which aggregates the distribution functions provided by the experts. Importantly, the aggregation is specific to the choice option to be evaluated, leaving the possibility of a prudent aggregation that puts more weight on more pessimistic views, outcome by outcome. Formally, this procedure is obtained by aggregating pointwise the decumulative distribution functions provided by all experts into a single decumulative function. Intuitive characteristics of the aggregator (or "averager") such as concavity ensure that the decision-rule is disagreement-averse according to our novel notion. We contrast our probability-averaging model with the approach that consists in using monotone ambiguity models to account for aversion to expert disagreement. Both approaches rely on some form of cautious aggregation where the degree of pessimism (or cautiousness) is what drives aversion to expert disagreement. The fundamental difference is that in the model we propose, the aggregation occurs on the level of experts' beliefs, while monotone ambiguity models suggest an aggregation that occurs on the level of utility values, yielding what we call "utility-averaging models". Although structurally different, these approaches can be compared in terms of ambiguity aversion. In particular, once utility is normalized to take values in $[0,1]$, like probabilities, we show that, for a given strictly concave averager and given risk preferences, our disagreement-averse specification exhibits greater ambiguity aversion than the corresponding utility-averaging models. Compared to the existing literature our paper can therefore be seen as bringing the notion of aversion to expert disagreement one step further.

The third contribution is to highlight the implications of disagreement aversion in concrete applications. We show that under general conditions, increases in disagreement aversion result in more cautious choices. We then apply this general result to climate mitigation and precautionary savings. We also compare our findings to those obtained when considering common "utility-averaging models" models exhibiting ambiguity aversion. When experts' beliefs can be unambiguously ranked from the most optimistic to the least optimistic (in terms of first-order stochastic dominance), ambiguity aversion in utility-averaging models also leads to more cautious choices. However, in the general case when experts' beliefs cannot be ranked in terms of first-order stochastic dominance, the result does not hold for utility-averaging models: an increase in ambiguity aversion could in fact result in less cautious choices. We illustrate this difference in a consumption-saving example: precautionary savings increase with disagreement aversion in our framework but decrease with ambiguity aversion in utility-averaging models.

Our paper is related to several strands of literature. There is a vast literature in mathematics and management on *opinion pooling*, in particular linear opinion pools in which the decision-maker uses a weighted average of expert opinions (DeGroot and Mortera 1991; Genest and Zidek 1986; Jose, Grushka-Cockayne and Lichtendahl 2013; Larrick and Soll 2006; McConway 1981; Morris 1977; Stone 1961). In line with that literature, we aggregate beliefs of experts. The key difference is that the literature on opinion pooling does not assume any ordering on the set of outcomes. With the notions of good and bad outcomes thus undefined, there cannot be anything such as a cautious aggregation of beliefs as is key in our contribution.

In line with other papers (Basili and Chateauneuf 2020; Berger, Emmerling and Tavoni 2016; Berger et al. 2021; Crès, Gilboa and Vieille 2011; Gajdos and Vergnaud 2013; Hansen and Sargent 2001; Heal and Millner 2018; Millner, Dietz and Heal 2013), we find it natural to borrow from the decision-theoretic literature on ambiguity aversion to model disagreement aversion.[3] The key difference is that to feature aversion for the lack of distribution-consensus, we depart from the standard unanimity principle and follow the dual approach by Bommier (2017).

Finally, our paper is connected to the general discussion in economics and social choice on how to extend the well-known preference aggregation result in Harsanyi (1955) from risk to uncertainty (Alon and Gayer 2016; Brandl 2020; Danan et al. 2016; Hylland and Zeckhauser 1979; Mongin 1995), and in particular to the debate on how to weaken the Pareto condition (Gilboa, Samet and Schmeidler 2004; Gilboa, Samuelson and Schmeidler 2014; Mongin 2016). However, like Stanca (2021), we exclusively focus on heterogeneity in beliefs. The decision-maker is indeed assumed to have her own risk preferences, and there is thus no need to discuss heterogeneity in tastes.

We proceed as follows. Section 1 defines the notion of disagreement aversion, introduces probability-averaging decision-rules, and clarifies the relation to

[3] See Machina and Siniscalchi (2014) for an excellent review of the ambiguity aversion literature.

common "utility-averaging" models that rest on the unanimity principle. Section 2 compares disagreement aversion and ambiguity aversion in the context of concrete applications. Section 3 concludes.

1 Theoretical framework

1.1 Setting

We consider a decision-maker who has to take decisions in a setting of uncertainty. Decisions will be driven by preferences over uncertain *prospects*. Experts provide scientific knowledge, but different experts may disagree. The decision-maker's preferences will naturally depend on the beliefs of all experts, which we call the *expertise*. Our analysis bears on *decision-rules* which describe how the expertise shapes the decision-maker's preferences. Formal definition of these concepts are provided below.

Prospects Let $X = [X^-, X^+]$, a closed interval of $\mathbb{R}$, be the space of *outcomes*. Given a set of *states of the world* Ω, we define (simple) *prospects* as mappings from states of the world to outcomes.[4] For any prospect α, we will denote by K_α the number of values taken by α, and by $(\alpha_1, \dots, \alpha_{K_\alpha})$ the values taken by α in increasing ordering (i.e. $\alpha_1 < \dots < \alpha_{K_\alpha}$). For any outcome x, the *sure prospect* with outcome x is the prospect that equals x in all states of the world. With a common abuse of notation, such a sure prospect will also be denoted by x. To avoid confusion, Greek letters α, β will be used for possibly non-sure prospects, while latin letters such as x will be reserved for outcomes and the corresponding sure prospects.

Expertise We consider N *experts*. Expert i holds *belief* P_i, a subjective probability measure on the set of the states of the world Ω. We call the list of experts' beliefs $\mathscr{P} = (P_1, \dots, P_N)$ the *expertise*. Making again a slight abuse of notation, for all

[4] Prospects are therefore just acts *à la* Savage (1954). We nevertheless opted for a different terminology to avoid the confusion with the notion of acts in the two-stage setting of Anscombe and Aumann (1963) where outcomes are lotteries, i.e. probabilistic objects.

belief P, the expertise $(P,\ldots,P)$ where all experts share the same belief will also be denoted P. Confusion will be avoided by using calligraphic symbols (like $\mathscr{P}$) for expertise where experts may disagree, and roman symbols (like P or P_i) for expertise where all experts agree. Given some belief P_i and an outcome x, we denote by $F_{\alpha}^{P_i}(x)$ the probability that α yields an outcome larger or equal to x, i.e

$$F_{\alpha}^{P_i}(x) = P_i\left(\{\omega \in \Omega \,|\, \alpha(\omega) \geq x\}\right).$$

The mapping $x \to F_{\alpha}^{P_i}(x)$ is thus the decumulative distribution function of the lottery generated by the prospect α when holding belief P_i.

Decision-rule A decision-rule $\succcurlyeq : \mathscr{P} \mapsto \succcurlyeq^{\mathscr{P}}$ is a mapping that associates any expertise $\mathscr{P}$ with a preference relation (i.e. a weak order) $\succcurlyeq^{\mathscr{P}}$ over the set of prospects. For any expertise $\mathscr{P}$, we denote by $\succ^{\mathscr{P}}$ and $\sim^{\mathscr{P}}$ the strict order and indifference relation corresponding to the weak order $\succcurlyeq^{\mathscr{P}}$.

With our notation, the preference relation $\succcurlyeq^{P_i}$ is, formally speaking, the preference relation that the decision-maker uses when all experts have the same beliefs as expert i.

Consensual prospects Central in our analysis is the notion of disagreement aversion (see section 1.2 below). A notion of disagreement aversion must ceteris paribus rely on a notion of disagreement, or lack of consensus, which could be formalized in different ways.

A first possibility to define consensus involves comparing how experts evaluate prospects in terms of certainty-equivalents. We will say that a prospect α is *utility-consensual* if the prospect's appeal does not depend on the expert the decision-maker relies on, i.e. when there exists $x \in X$ such that $\alpha \sim^{P_i} x$ for all i. Note that the formulation makes use of the indifference relations $\sim^{P_i}$, which means that the notion of utility-consensual actually depends on the decision-rule which is considered. Crucial for our analysis, a prospect may look utility-consensual while experts still fundamentally disagree on the risk implied by such a prospect. An illustration is given in our introductory example: no matter which expert the decision-maker relies on, the project is seen as good as doing nothing, but one expert thinks that the project is risky and the other not.

As stressed by Mongin (2016), the decision-maker may want to treat cases of complete expert agreement and cases of "spurious agreement" differently. This leads us to introduce a second notion of consensus. We say that a prospect α is *distribution-consensual* when experts agree on the distribution of outcomes it entails (i.e $F_\alpha^{P_i} = F_\alpha^{P_j}$ for all experts i and j). The concept of disagreement aversion introduced in section 1.2 will directly rely on the notion of distribution-consensus. In contrast to utility-consensus, distribution-consensus is defined independently from the decision rule. Note also that all sure prospects are both utility- and distribution-consensual, reflecting that experts' beliefs are irrelevant for the evaluation of sure prospects.

Risk preferences Throughout the paper, we will assume that distribution-consensual prospects are evaluated through a rank-dependent expected utility (RDU) model. In other words, we assume that when experts all agree on the distribution of outcomes implied by a prospect, the prospect is evaluated through RDU preferences. RDU is a well-known generalization of the expected utility (EU) theory, introduced by Quiggin (1982), where the probabilities of a prospect are transformed. We could of course have restricted our setting to EU risk preferences (one just needs replace the function f by the identity function in all mathematical expressions that follow). However, since it does not make the analysis significantly more complex, we thought it interesting to keep the additional flexibility of the RDU model which allows to express extra sensitivity to rare and extreme events. This may be valuable for modelling social preferences in the context of uncertainty, especially when considering unlikely but dramatic outcomes (e.g., a major nuclear catastrophe).

As it was emphasized by Castagnoli and Calzi (1996) for the EU model, a given framework may be presented in different (but equivalent) ways, suggesting however quite different readings and interpretations. The same is true for the RDU framework. For convenience, we use the approach of Castagnoli and Calzi (1996) in Definition 1 below (and mention further below, in equation (2) an equivalent formulation which might look more familiar).

Definition 1 (RDU on distribution-consensual prospects)**.** *A decision-rule $\succcurlyeq$ is said RDU on distribution-consensual prospects if there exist two increasing bijections $f:[0,1]\to[0,1], u:X\to[0,1]$ such that for any expertise $\mathscr{P}$, the weak order $\succcurlyeq^{\mathscr{P}}$ restricted on distribution-consensual prospects is represented by:*

$$U^{\mathscr{P}}(\alpha)=\sum_{k=1}^{K_\alpha}\sigma_k f(p_k) \tag{1}$$

where $\sigma_k=u(\alpha_k)-u(\alpha_{k-1})\geq 0$ for $k\in\{2,\ldots,K_\alpha\}$, $\sigma_1=u(\alpha_1)$ and $p_k=F_\alpha^{P_1}(\alpha_k)$.[5]

As a remark, note that through a summation by part in (1), one may also write the RDU utility as follows:

$$U^{\mathscr{P}}(\alpha)=\sum_{k=1}^{K_\alpha-1}u(\alpha_k)\left(f(p_k)-f(p_{k+1})\right)+u(\alpha_{K_\alpha})f(p_{K_\alpha}). \tag{2}$$

This popular formulation is just another way to present the RDU specification.

It directly follows from (1), or from (2), that for any sure prospect x and any expertise $\mathscr{P}$, one has $U^{\mathscr{P}}(x)=u(x)$.[6] Note also that we normalize the utility function by imposing that $\mathrm{Im}(u)=\mathrm{Im}(f)=[0,1]$. This implies that the representation is then unique, but of course without loss of generality.[7]

1.2 Disagreement aversion

Disagreement aversion, which is a central concept in our paper, can be seen as a "dislike" for prospects which are not distribution-consensual. It is formally defined

[5]This definition seems to give a special role to the beliefs of expert 1, since we define p_k using belief P_1. Remember, however, that for distribution-consensual prospects, the functions $F_\alpha^{P_i}$ are independent of i. Thus, we also have $p_k=F_\alpha^{P_i}(\alpha_k)$ for all $1\leq i\leq N$.

[6]When α is a sure prospect providing outcome x in all states, one has $K_\alpha=1$, $\alpha_1=x$, and $p_1=1$.

[7]This normalization will also be very convenient for comparing utility-averaging and probability-averaging decision-rules in Section 1.5 since having utility levels and transformed probabilities covering the same interval $[0,1]$ will make it possible to consider identical averager function, be it to aggregate utilities or to aggregate probabilities.

as follows:

Definition 2 (Disagreement aversion)**.** *A decision-rule* $\succcurlyeq$ *is said* disagreement-averse *if for every expertise* $\mathscr{P}$*, every prospect* α *that is not distribution-consensual, and every sure prospect x, the following implication holds:*

$$\left(x \succcurlyeq^{P_i} \alpha,\, \forall i\right) \Rightarrow x \succ^{\mathscr{P}} \alpha.$$

Note that for decision-rules which are RDU on distribution-consensual prospects, it directly follows from Definition 1 that for all distribution-consensual prospects α, one has $\left(x \succcurlyeq^{P_i} \alpha,\, \forall i\right) \Rightarrow x \succcurlyeq^{\mathscr{P}} \alpha$. Disagreement aversion states that this implication extends to all prospects: intuitively, if all experts think that α (consensual or not) is not better than x, disagreement among experts cannot make the decision-maker strictly prefer α to x. In addition, disagreement aversion requires that when α is not distribution-consensual, then the decision-maker strictly prefers the sure prospect x.

In the most standard way, we can follow Yaari (1969)'s general approach — initially introduced to compare risk aversion — to compare disagreement aversion. Formally:

Definition 3 (Comparative disagreement aversion)**.** *A decision-rule* $\succcurlyeq_A$ *exhibits greater disagreement aversion than a decision-rule* $\succcurlyeq_B$*, if for every expertise* $\mathscr{P}$*, every prospect* α *and every distribution-consensual prospect* β*,*

$$\alpha \succ_A^{\mathscr{P}} \beta \Rightarrow \alpha \succ_B^{\mathscr{P}} \beta; \quad \alpha \succcurlyeq_A^{\mathscr{P}} \beta \Rightarrow \alpha \succcurlyeq_B^{\mathscr{P}} \beta$$

and

$$\alpha \sim_A^{\mathscr{P}} \beta \Rightarrow \alpha \succ_B^{\mathscr{P}} \beta \text{ if } \alpha \text{ is not distribution-consensual}$$

Intuitively, what this definition says is that if according to the decision-rule $\succcurlyeq_A$ the prospect α looks preferable to the distribution-consensual β (despite potential disagreement about the riskiness of α), it must also be the case when considering a decision-rule exhibiting lower disagreement aversion. Moreover, in case α is not consensual, and seen just as good as β when using the more disagreement-averse decision-rule, it must be seen as strictly better than β under the less disagreement averse decision-rule. Note that if $\succsim_A$ and $\succsim_B$ are comparable in terms of disagreement aversion, in the sense that one exhibits more disagreement aversion than the other, they must necessarily agree on the ranking of distribution-consensual prospects (see proof in Appendix A.3). This parallels the literature on risk aversion, where agents are comparable in terms of risk aversion only if they agree on the ranking of deterministic outcomes (see Kihlstrom and Mirman 1974).

Violation of the Pareto condition Disagreement aversion directly conflicts with the Pareto (or unanimity) condition. The fundamental reason is that utility-consensual prospects are not necessarily distribution-consensual. In order to formally stress the tension between disagreement aversion and the Pareto condition, we introduce the notion of Paretian decision-rule.

Definition 4 (Paretian decision-rule)**.** *A decision-rule $\succsim$ is Paretian if for every expertise $\mathscr{P} = (P_1, \ldots, P_N)$ and all prospects α, β:*

$$\left(\beta \succsim^{P_i} \alpha, \forall i\right) \Rightarrow \beta \succsim^{\mathscr{P}} \alpha$$

The following result immediately follows.

Lemma 1. *There is no decision-rule which is Paretian, RDU on distribution-consensual prospects, and exhibits disagreement aversion.*

Proof. Suppose that the decision-rule is RDU on distribution-consensual prospects. One can construct an expertise $\mathscr{P} = (P_1, \ldots, P_N)$, a non distribution-consensual prospect α and a sure prospect x such that $x \sim^{P_i} \alpha$, $\forall i$, e.g. using the beginnings of the proofs of Proposition 1 (in Appendix A.1) and Proposition 10 (in Appendix A.4). If the decision rule is Paretian we then have $x \sim^{\mathscr{P}} \alpha$, contradicting $x \succ^{\mathscr{P}} \alpha$, and thus disagreement aversion. ∎

Lemma 1 emphasizes that allowing for disagreement aversion requires departing from the Pareto condition. Given that disagreement aversion is a plausible normative property, we argue that the Pareto condition is not as appealing when aggregating experts' beliefs as when betting on a horse race.[8] To model disagreement

[8]Following Anscombe and Aumann (1963), horse lotteries is often used as an analogy for decision-making under uncertainty.

aversion, we cannot exploit the major ambiguity models discussed in the literature (e.g. in Machina 2014; Strzalecki 2013), as they rely on a monotonicity axiom which in our setting translates to the Pareto condition.

An exception in the ambiguity aversion literature is Bommier (2017), who explicitly relaxes the monotonicity axiom.We will borrow from this paper to suggest a class of decision-rules exhibiting disagreement aversion.

1.3 Probability-averaging decision-rules

The following definition introduces the class of *probability-averaging decision rules* with which disagreement may matter, possibly to exhibit disagreement-aversion, disagreement-loving or some ambiguous patterns of disagreement sensitivity. Below the definition, we explain how such specification may be derived from a formal set of axioms.

Definition 5 (Probability-averaging decision-rule)**.** *A decision-rule* $\succcurlyeq$ *is said probability-averaging, with representation* (u, f, I)*, if there exist two increasing bijections* $u : X \to [0,1]$*,* $f : [0,1] \to [0,1]$ *, and a continuous, component-wise strictly increasing function* $I : [0,1]^N \to [0,1]$ *fulfilling* $I(q,\ldots,q) = q$ *such that for any expertise* $\mathscr{P}$*, the weak order* $\succcurlyeq^{\mathscr{P}}$ *is represented by:*

$$U^{\mathscr{P}}(\alpha) = \sum_{k=1}^{K_\alpha} \sigma_k I\left(f\left(p_k^1\right),\ldots,f\left(p_k^N\right)\right) \tag{3}$$

where $\sigma_k = u(\alpha_k) - u(\alpha_{k-1}) \geq 0$ *for* $k \in \{2,\ldots,K_\alpha\}$*,* $\sigma_1 = u(\alpha_1)$ *and* $p_k^i = F_\alpha^{P_i}(\alpha_k)$*.*

We will refer to u as the utility index, to f as the probability transformation function and to I as the averager.

Remark 1. The representation is unique, in the sense that a decision-rule admits only one (u, f, I) representation. A probability-averaging decision rule is necessarily RDU on consensual prospect, with representation as in equation (1).

Remark 2. Defining $\widetilde{I} : \left(p^1,\ldots,p^N\right) \mapsto \widetilde{I}\left(p^1,\ldots,p^N\right) = f^{-1}\left(I\left(f\left(p^1\right),\ldots,f\left(p^N\right)\right)\right)$, equation (3) simply rewrites:

$$U^{\mathscr{P}}(\alpha) = \sum_{k=1}^{K_\alpha} \sigma_k f\left(\widetilde{I}\left(p_k^1, \ldots, p_k^N\right)\right). \tag{4}$$

Thus, as shown in Figure 1, a probability-averaging decision-rule can be seen as one where the values of the decumulative distribution functions provided by each expert are "averaged" through the function $\widetilde{I}$ in order to provide some equivalent decumulative distribution function (i.e. some equivalent belief). The prospect is then evaluated using the RDU specification using this equivalent decumulative distribution function. It is worth noting that the way some belief P_i ends up giving a decumulative distribution function depends on the prospect that is evaluated (mathematically, the p_k^i depends on both P_i and α). Thus, aggregating decumulative distributions functions associated to a given prospect is not equivalent to a mere aggregation of beliefs which would be made independently of whether states provide high or low pay-offs.[9]

Axiomatization An axiomatization of probability-averaging decision-rules is provided in online Appendix B. The main axioms are:

- Monotonicity with respect to first-order stochastic dominance (M-FSD). This axiom is weaker than the monotonicity axiom assumed in models based on the unanimity principle (Pareto condition), which rules out disagreement aversion. Our M-FSD axiom only requires that a decision-rule prefers a prospect α to a prospect β ($\alpha \succcurlyeq \beta$) if α dominates β with respect to first-order stochastic dominance, i.e. $F_\alpha^{P_i} \geq F_\beta^{P_i}, \forall i$, while the Pareto condition requires that it be the case as soon as $U^{P_i}(\alpha) \geq U^{P_i}(\beta), \forall i$.

- RDU on distribution-consensual prospects. Equivalently, this axiom can be replaced by comonotonic mixture independence (see Chateauneuf 1999).

[9] This is a fundamental difference with the linear opinion pooling where the decision-maker uses a convex combination of expert beliefs to forms their own belief, independently of the prospect to be evaluated.

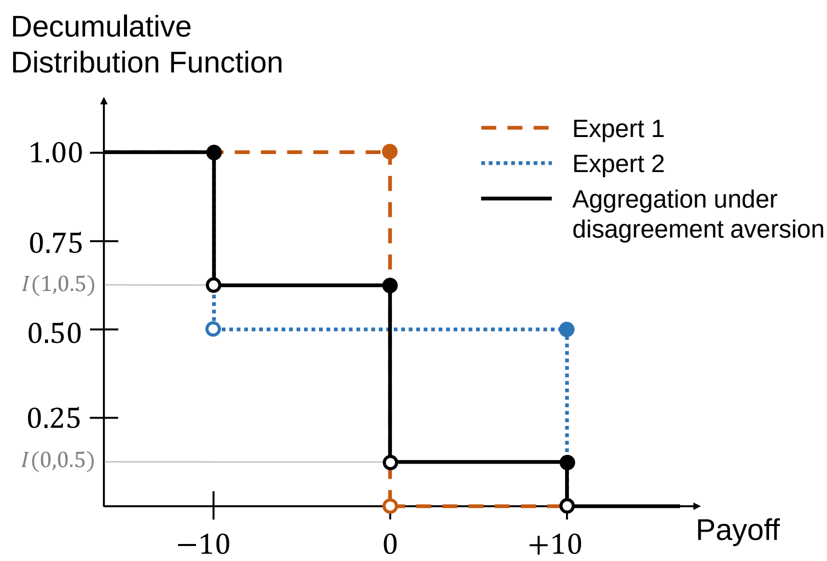

Figure 1 – Aggregation of experts' beliefs.

- The comonotonic sure-thing principle (Axiom 4, absent from models with the Pareto condition). This specifies that a common outcome of two prospects at some state of the world can be changed without impacting the comparison between those prospects, as long as the change does not affect the rankings of each prospect's outcomes.
- A usual continuity axiom.
- An axiom of "level independence" which restricts the type of disagreement aversion that the decision-rule can exhibit; namely, we require that when experts disagree over the probability of an outcome, the way disagreement is resolved does not depend on the outcome but just on the probabilities. This axiom is introduced to obtain a simple representation. One could however relax it, to obtain a broader class of probability-averaging decision-rules.[10]

[10]One would then obtain a representation similar to the one given in the Theorem 2 of Bommier (2017).

1.4 Characterization of disagreement aversion

We first state what it requires to have a decision-rule that exhibits disagreement aversion.

Proposition 1. *A probability-averaging decision-rule with representation* (u, f, I) *exhibits disagreement aversion if and only if the averager* I *is such that for any matrix* $(q^i_k)_{1\leq k\leq K, 1\leq i\leq N} \in [0,1]^{N\times K}$ *such that* $k \geq l \Rightarrow q^i_k \leq q^i_l$*, and any vector* $(\sigma_1, \ldots, \sigma_K) \in [0,1]^K$ *such that* $\sum_{k=1}^K \sigma_k \leq 1$*, one has:*

$$\sum_{k=1}^{K} \sigma_k I\left(q^1_k, \ldots, q^N_k\right) \leq \max_{1\leq i\leq N} \sum_{k=1}^{K} \sigma_k q^i_k$$

where the inequality is strict whenever one has $q^i_k \neq q^j_k$ *and* $\sigma_k > 0$ *for some indices* i, j, k*.*

Proof. See Appendix A.1. ■

To our knowledge, there is no available name for the property stated in the above proposition. One can however provide simpler conditions on I that are sufficient (but not necessary) to obtain disagreement aversion.

Proposition 2. *Each of the conditions below is sufficient to imply disagreement aversion:*

- **Condition 1**: *There exist numbers* $(\lambda_i)_{1\leq i\leq N}$ *(also called weights), which are non-negative and sum to 1 such that* $I(p_1, \ldots, p_N) \leq \sum_{i=1}^N \lambda_i p_i$ *for all* $(p_1, \ldots, p_N) \in [0,1]^N$*, with a strict inequality when* $p_i \neq p_j$ *for some indices* i *and* j*.*

- **Condition 2**: *The function* $(p_1, \ldots, p_N) \to I(p_1, \ldots, p_N)$ *is strictly concave except on the diagonal.*[11]

Proof. See Appendix A.2. ■

[11] By strictly concave except on the diagonal, we mean that the strict concavity inequality holds for any pair of vectors of which at least one is not constant

Condition 1 is very convenient, as it makes it possible to borrow from the literature on ambiguity aversion where such an inequality (considered then for averagers that average utilities rather than probabilities) is the one that usually defines ambiguity aversion (e.g., Cerreia-Vioglio et al. 2011). For example, one can think of smooth "KMM" kind of averagers of the form:[12]

$$I_S(p_1,\ldots,p_N) = \psi^{-1}(\sum_{i=1}^{N}\lambda_i\psi(p_i)) \tag{5}$$

for some increasing smooth function $\psi : [0,1] \to [0,1]$. Disagreement aversion is then obtained when ψ is strictly concave. Alternatively, one can use averagers of the max-min kind,

$$I_M(p_1,\ldots,p_N) = \min_{(\lambda_i)_{1\leq i\leq N}\in\chi} \left\{\sum_{i=1}^{N}\lambda_i p_i\right\}, \tag{6}$$

where χ is a closed and convex set of weights. This averager will exhibit disagreement aversion whenever χ does not reduce to a singleton. These are just two possibilities, the literature on ambiguity aversion offers many averagers, such as those related to the variational model of Maccheroni, Marinacci and Rustichini (2006) or to the model of Siniscalchi (2009).

The last result of this section is about comparative disagreement aversion.

Proposition 3. *Consider two probability-averaging decision-rules $\succcurlyeq_A$ and $\succcurlyeq_B$ with representation (u_A, f_A, I_A) and (u_B, f_B, I_B). Then $\succcurlyeq_A$ exhibits greater disagreement aversion than $\succcurlyeq_B$ if and only if $u_A = u_B$, $f_A = f_B$ and $I_A(\vec{p}) < I_B(\vec{p})$ for every non-constant vector $\vec{p} = (p_1,\ldots,p_N) \in [0,1]^N$.*

Proof. See Appendix A.3. ■

The degree of disagreement aversion is thus fully determined by the averager, with "smaller" averagers yielding more disagreement-averse decision-rules. This shares a lot of similarity with results from the ambiguity aversion literature, where

[12] More precisely, such averagers correspond to the Second-Order EU model (see e.g., Nau 2006) the most common specification for the models introduced by Klibanoff, Marinacci and Mukerji (2005).

ambiguity aversion is driven by how "pessimistic" the averagers (which aggregate utilities in that literature) are. This is very convenient, as we can readily import knowledge from the ambiguity literature to determine what it means to increase disagreement aversion. For example, if we consider smooth KMM averagers as in equation (5), we readily know that increasing disagreement aversion is equivalent to increasing the concavity of the function ψ. Similarly, if we consider max-min kind averager defined as in equation (6), we know that increasing disagreement aversion involves using a larger set of weights χ.

1.5 Relation with models that rest on the unanimity principle

In order to explain how our work can connect with previous contributions, we introduce a class of models that rest on the unanimity principle.

Definition 6 (Utility-averaging decision-rule)**.** *A decision-rule $\succcurlyeq$ is said to be utility-averaging with representation (u, f, I), if there exist two increasing bijections $u : X \to [0,1]$, $f : [0,1] \to [0,1]$, and a continuous, component-wise strictly increasing function $I : [0,1]^N \to [0,1]$ fulfilling $I(q, \ldots, q) = q$ such that for any expertise $\mathscr{P}$, the weak order $\succcurlyeq^{\mathscr{P}}$ is represented by:*

$$U^{\mathscr{P}}(\alpha) = I\left(\sum_{k=1}^{K_\alpha} \sigma_k f\left(p_k^1\right), \ldots, \sum_{k=1}^{K_\alpha} \sigma_k f\left(p_k^N\right)\right) \tag{7}$$

where $\sigma_k = u(\alpha_k) - u(\alpha_{k-1}) \geq 0$ for $k \in \{2, \ldots, K_\alpha\}$, $\sigma_1 = u(\alpha_1)$ and $p_k^i = F_\alpha^{P_i}(\alpha_k)$.

The qualification of *utility-averaging* comes from the fact that (7) can be rewritten as $U^{\mathscr{P}}(\alpha) = I\left(U^{P_1}(\alpha), \ldots, U^{P_N}(\alpha)\right)$ where $U^{P_i}(\alpha) = \sum_{k=1}^{K_\alpha} \sigma_k f\left(p_{k,}^i\right)$ is the RDU utility obtained when holding belief P_i. The averager I being increasing, such decision-rules are necessarily utility-averaging. This class of utility-averaging models embeds all major models listed in Machina (2014). The representation (7) is actually very similar to that of Cerreia-Vioglio et al. (2011), a very general specification that encompasses most ambiguity models.[13] In those ambiguity models, the

[13]Compared to the Monotonic, Bernoullian and Archimedean (MBA) model of Cerreia-Vioglio et al. (2011), our utility-averaging model evaluates risk using RDU instead of EU. In so doing, we drop the Bernoullian assumption in MBA and consider an even broader class of models.

averager I is what characterizes ambiguity aversion.

Comparing equation (3) and equation (7), we see that probability-averaging and utility-averaging decision rules only differ by the stage where the averager is applied. In the former case, I aggregates the probabilities, while in the latter, I aggregates the utility levels. When I is linear, both representations are equivalent, and the decision-rule is both probability-averaging and utility-averaging. The reciprocal is also true, so that:

Proposition 4. *A decision-rule is both probability-averaging and utility-averaging if and only if it admits a (probability-averaging or utility-averaging) representation* (u, f, I) *with a linear averager* I*.*

This result, whose proof is given in Appendix A.4, shows that being simultaneously probability-averaging and utility-averaging requires to have a linear averager I, which is usually considered as characterizing ambiguity neutrality. The probability-averaging and utility-averaging approaches differ as soon as non-linear averagers are assumed. Interestingly, they can be compared in terms of ambiguity aversion, a fundamental concept in the ambiguity literature. In our setting, comparative ambiguity aversion can be defined as follows:

Definition 7 (Comparative ambiguity aversion)**.** *A decision-rule* $\succcurlyeq_A$ *exhibits greater ambiguity aversion than a decision-rule* $\succcurlyeq_B$ *if for every expertise* $\mathscr{P}$*, every prospect* α *and every distribution-consensual prospect* β*,*

$$\alpha \succ_A^{\mathscr{P}} \beta \Rightarrow \alpha \succ_B^{\mathscr{P}} \beta; \quad \alpha \succcurlyeq_A^{\mathscr{P}} \beta \Rightarrow \alpha \succcurlyeq_B^{\mathscr{P}} \beta$$

and

$$\alpha \sim_A^{\mathscr{P}} \beta \Rightarrow \alpha \succ_B^{\mathscr{P}} \beta \text{ if } \alpha \text{ is not utility-consensual}$$

This definition is very similar to Definition 3 but relies on a different notion of consensus. While Definition 3 refers to the aversion for the lack of distribution consensus, Definition 7 refers to the aversion for a lack of utility consensus.[14]

[14]Remark that although the notion of utility consensus depends on the decision-rule (and in particular, on u and f), we do not need to specify whether "utility-consensual" refers to A or B in Definition 7, as both choices would be equivalent. This is because decision-rules share the func-

We first state a result (already well-known for utility-averaging models – see Ghirardato, Maccheroni and Marinacci 2004) showing that for both probability-averaging decision-rules and utility-averaging decision-rules, the level of ambiguity aversion is dictated by the averager I. Formally:

Proposition 5. *Consider two decision-rules $\succcurlyeq_A$ and $\succcurlyeq_B$, which are either both probability-averaging or both utility-averaging with representations (u_A, f_A, I_A) and (u_B, f_B, I_B). Then $\succcurlyeq_A$ exhibits greater ambiguity aversion than $\succcurlyeq_B$ if and only if $u_A = u_B$, $f_A = f_B$ and $I_A(\vec{p}) < I_B(\vec{p})$ for every non-constant vector $\vec{p} = (p_1, \ldots, p_N)$.*

Proof. See Appendix A.3. ■

Proposition 5 show how two decision-rules of the same kind can be compared in terms of ambiguity aversion. The following result shows that it is also possible to compare probability-averaging decision-rules with utility-averaging decision-rules. Namely, under a concavity condition on the averager I, we show that a probability-averaging decision-rules exhibits greater ambiguity aversion and greater disagreement aversion than the utility-averaging decision rule which uses the same utility index, the same probability transformation functions and the same averager. To be fully precise, the result is valid only when we exclude the special case of prospects whose outcomes are all extremal (i.e. $\alpha_k \in \{X^-, X^+\}, \forall k$), because both rules coincide on these prospects. For the purpose of the proposition, we thus say that a decision-rule exhibits greater disagreement aversion (resp. ambiguity aversion) *except on prospects with only extremal outcomes* if it fulfills a slightly weaker version of Definition 3 (resp. Definition 7) where α is replaced by any prospect with at least one non-extremal outcome (i.e. $0 < \sigma_k < 1$ for some k).

Proposition 6. *Consider a probability-averaging decision-rule $\succcurlyeq_{PA}$ with representation (u, f, I) and the utility-averaging decision rule $\succcurlyeq_{UA}$ with representation (u, f, I) for the same functions u, f, and I. If I is strictly concave except on the diagonal then $\succcurlyeq_{PA}$ exhibits greater ambiguity aversion and greater disagreement aversion than $\succcurlyeq_{UA}$ except on prospects with only extremal outcomes.*

tions u and f as soon as they are comparable in terms of ambiguity aversion, a property proven in AppendixA.3 using the first conditions of the Definition 7.

Proof. See Appendix A.6. ■

Combining several of the above results, Figure 2 provides a global picture on how probability-averaging and utility-averaging decision-rules relate to each other. Starting from the linear pooling case where the decision maker's utility is a linear average of experts' utilities, the probability-averaging and utility-averaging frameworks offer two different ways to introduce ambiguity aversion. The utility-averaging framework does so while preserving the unanimity principle. This leaves, however, no room for disagreement aversion. The probability-averaging framework allows for further ambiguity aversion, by aggregating probabilities first, and then computing utilities. This is what produces disagreement aversion.

Our motivation for formalizing the notion of disagreement aversion and suggesting a model that takes it into account is not purely theoretical, but rooted in implications that this may have when considering concrete issues. Accounting for disagreement aversion may indeed advocate for more precautionary decisions when experts disagree and provide insights on how scientific expertise should be used for decision making. We develop these aspects in the following section.

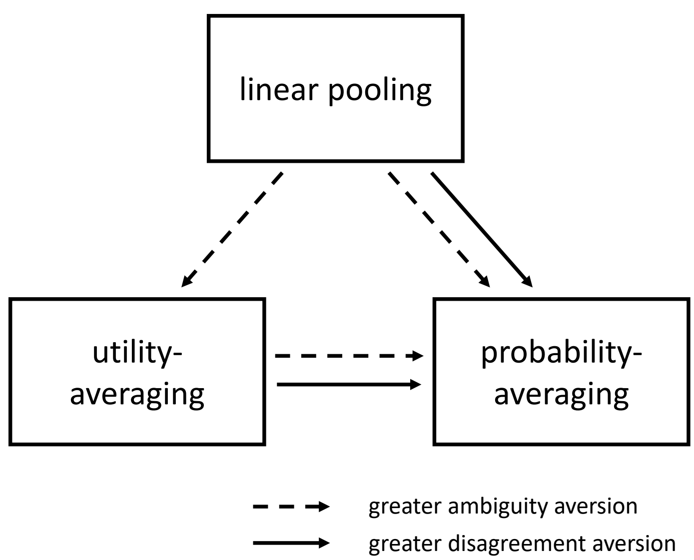

Figure 2 – The relation between decision-rules when I is strictly concave.

2 Applications

We consider a decision-maker whose ex-post utility $u(a,\omega)$ depends on action $a \in \mathbb{R}$ and a contingency $\omega \in \mathbb{R}$ upon which they have no control. We assume that the function $(a,\omega) \mapsto u(a,\omega)$ is increasing in ω, which is without loss of generality when contingencies can be ranked from "adverse" to "favorable" independently of the decision-maker's action. Ex-ante, when the decision-maker has to decide about a, the value of the contingency ω is uncertain. There are N experts who provide potentially diverging views about the distribution of ω. To make the link with the theory section 1, we assume that the distributions provided by the experts have finite support, with $\omega \in \{\omega_1, \dots, \omega_K\}$ where $(\omega_k)_{1 \leq k \leq K}$ is an increasing sequence of real numbers ($k < l \Rightarrow \omega_k < \omega_l$). Each expert $i \in \{1, \dots, N\}$ provides a probability vector $(\pi_k^i)_{1 \leq k \leq K}$, where π_k^i is the probability of $\omega = \omega_k$ according to expert i. We assume that there is some disagreement among experts ($\pi_k^i \neq \pi_k^j$ for some i, j, k).

2.1 A general result on the impact of disagreement aversion

We assume here that decision-makers use probability-averaging decision-rules as introduced in Definition 5, with the only difference that we no longer require the utility to be normalized to values in $[0,1]$. Since our aim is to discuss the impact of disagreement aversion, we will consider two decision-makers, A and B, who only differ by their averagers I_A and I_B. Formally speaking, the decision-maker $\tau = A, B$ has the following program:

$$\max_{a} \sum_{k=1}^{K} \sigma_k I_\tau \left(f\left(p_k^1\right), \dots, f\left(p_k^N\right) \right), \tag{8}$$

where $p_k^i = \sum_{l=k}^{K} \pi_l^i$, $\sigma_k = u(a,\omega_k) - u(a,\omega_{k-1})$ for $k > 1$ and $\sigma_1 = u(a,\omega_1)$. The function f is an increasing bijection over $[0,1]$ and the averager I_τ is a continuous and componentwise strictly increasing function $I_\tau : [0,1]^N \to [0,1]$ fulfilling $I_\tau(q, \dots, q) = q$. We assume that the decision maker's problem always has a finite (thus interior) solution denoted $a_\tau^* \in \mathbb{R}$.

Proposition 7. *Assume that $(a,\omega) \mapsto u(a,\omega)$ is twice continuously differentiable,* *and that probability-averaging decision-maker A exhibits greater disagreement aversion than B, i.e.* $I_A(p_1,\dots,p_N) < I_B(p_1,\dots,p_N)$ *for all non-constant vectors* $(p_1,\dots,p_N) \in [0,1]^N$. *Then:*

- *If* $\frac{\partial^2 u(a,\omega)}{\partial a\, \partial \omega} > 0$ *for all* a *and* ω*, then* $a_A^* < a_B^*$.
- *If* $\frac{\partial^2 u(a,\omega)}{\partial a\, \partial \omega} < 0$ *for all* a *and* ω*, then* $a_A^* > a_B^*$.

Proof. See Appendix A.7. ■

This proposition shows that when the cross-derivative of the function u has a constant sign, an increase of disagreement aversion has an unambiguous impact on the optimal action. The reason is the following. If the cross-derivative is positive, an increase of a widens the distance between any two possible ex-post utility levels. This makes the disagreement between experts (i.e. the difference in their distribution functions) more significant. This, in turn, implies that an increase of disagreement aversion leads to a decrease of the optimal a. Vice versa for a negative cross-derivative. It is worth noting that for obtaining this result we did not have to assume a particular functional form for the averager. It thus holds for KMM types or max-min averagers, as with other averagers that can be found in the ambiguity aversion literature.

We illustrate the general result in Proposition 7 with two examples, climate policy and precautionary savings.

2.1.1 A climate mitigation example

As highlighted by the literature (e.g., Meinshausen et al. 2009), experts in climate physics substantially disagree on the climate sensitivity, i.e. by how much global average temperatures increase as a result of increased greenhouse gas levels. The climate sensitivity is a key parameter for determining the optimal level of greenhouse gas abatement. Should disagreement among experts lead to choose a higher or lower emission abatement target?

In order to give insights on this question we consider a two-period model. In period 1, the decision-maker is endowed with wealth w_1 of which an amount $C(a)$, increasing in a, can be taken to finance an abatement level a, leaving $w_1 - C(a)$

for consumption. In period 2, consumption equals wealth w_2 minus climate-related damages. These damages depend on the abatement level a chosen in period 1 and the climate sensitivity $\theta \in \mathbb{R}$, which is ex-ante uncertain. We denote by $(a,\theta) \mapsto D(a,\theta)$ the damage function and assume that $\frac{\partial D}{\partial a} < 0$, $\frac{\partial D}{\partial \theta} > 0$ and $\frac{\partial^2 D}{\partial a\, \partial \theta} < 0$; the last inequality means that abatement is more efficient in reducing damages when climate sensitivity is high. The ex-post utility is given by

$$u(a,\theta) = v(w_1 - C(a)) + \beta v(w_2 - D(a,\theta))$$

where v is the (increasing and concave) instantaneous utility function and $\beta \in (0,1)$ the time discount factor. One can easily check that

$$\frac{\partial u(a,\theta)}{\partial \theta} < 0 \ ; \ \frac{\partial^2 u(a,\theta)}{\partial a\, \partial \theta} > 0.$$

In order to be in the setting of Proposition 7, we set $\omega = -\theta$, so that utility increases with ω. It then directly follows from Proposition 7 that an increase of disagreement aversion leads to a strict increase of emission abatement. The interpretation is simple: Both in terms of ex-post wealth and ex-post utility, the marginal benefit of emission abatement is higher in bad states of nature than in good states of nature. Thus abatement is able to reduce the distance between ex-post utilities, which implies that more disagreement aversion leads to an increase of optimal abatement.

2.1.2 A precautionary savings example

We now consider a standard precautionary savings problem in a two-period setting. The decision-maker receives income y in period 1 and has uncertain income ω in period 2. The decision-maker chooses the amount a saved in period 1. We assume a deterministic rate of interest, r, so that saving a in period 1 yields $(1+r)a$ in period 2. Choosing an amount of saving a and receiving income ω in period 2 provides an (ex-post) intertemporal utility:

$$u(a,t) = v(y-a) + \beta v(\omega + (1+r)a)$$

where v is the (increasing and concave) instantaneous utility function and $\beta \in (0,1)$ the time discount factor. One has

$$\frac{\partial u(a,\omega)}{\partial \omega} > 0 \ ; \ \frac{\partial^2 u(a,\omega)}{\partial a \, \partial \omega} < 0.$$

Thus a direct application of Proposition 7 shows that an increase in disagreement aversion leads to a strict increase of precautionary savings. While the marginal benefit of precautionary savings is similar across states of nature in terms of ex-post wealth, it is higher in bad states of nature than in good states of nature in terms of ex-post utility. Thus precautionary savings reduce the distance between ex-post utilities, which implies that an increase of disagreement aversion leads to higher precautionary savings.

2.2 Utility-averaging versus probability-averaging

This Section highlights how our model differs from common models of ambiguity aversion in applications. As explained in Section 1.5, probability-averaging decision-rules and common models of ambiguity aversion differ as the former suggests to aggregate probabilities first, and then compute expected utilities, while the latter suggests to compute expected utilities first, and then to aggregate utilities. We will show in this section, that both procedures may yield qualitatively similar results in some cases, but very different, possibly contrary results, in others.

2.2.1 The impact of ambiguity aversion under utility-averaging

We consider here two decision-makers, A and B who use utility-averaging decision-rules but only differ by the averagers I_A and I_B they are using. With the setting described at the beginning of Section 2, the decision-maker $\tau = A, B$ has the following program:

$$\max_a I_\tau \left(\sum_{k=1}^{K} \sigma_k f\left(p_k^1\right), \ldots, \sum_{k=1}^{K} \sigma_k f\left(p_k^N\right) \right). \tag{9}$$

where the σ_k and the p_k^i are defined as in Section 2.1. We assume that the decision-maker's program always admits a finite (thus interior) solution $a_\tau^* \in \mathbb{R}$.

We know from Proposition 5 that decision-maker A exhibits greater ambiguity aversion than decision-maker B if $I_A(u_1,\ldots,u_N) < I_B(u_1,\ldots,u_N)$ for non-constant vectors of utility levels $(u_1,\ldots,u_N)$. This is however generally insufficient to derive general conclusions about the impact on the optimal action. As noticed by Gollier (2011), Millner, Dietz and Heal (2013), Berger, Emmerling and Tavoni (2016), Berger (2014) or Peter (2019), a result can nevertheless be provided under specific assumptions regarding the averagers I_A and I_B and how experts' beliefs compare:

Proposition 8. *Assume that:*

- *I_A and I_B have the smooth "KMM" form (5) with twice continuously differentiable functions ψ_A and ψ_B.*

- *ψ_A is more concave than ψ_B (so that utility-averaging decision-maker A is more ambiguity-averse than B).*

- *$p_k^1 = \sum_{l=k}^K \pi_l^1 \leq \ldots \leq p_k^N = \sum_{l=k}^K \pi_l^N$ for all k (meaning that the experts' beliefs can be ordered in terms of first-order stochastic dominance).*

- *The function $(a,\omega) \mapsto u(a,\omega)$ is twice continuously differentiable and such that $\frac{\partial^2 u(a,\omega)}{\partial a\, \partial \omega} > 0$ (resp. $\frac{\partial^2 u(a,\omega)}{\partial a\, \partial \omega} < 0$) for all $(a,\omega) \in \mathbb{R}^2$.*

Then $a_A^* < a_B^*$ (resp. $a_A^* > a_B^*$).

Proof. See Appendix A.8. ■

The result stated in Proposition 8 looks qualitatively similar to the one stated in Proposition 7 regarding disagreement-averse decision-rules. The interpretation is also qualitatively similar. Assuming for instance that the cross-derivative of the function u is positive and experts ordered in the sense of first-order stochastic dominance, an increase of a widens the distance between the expected utility levels computed with any two expert beliefs. Thus, the higher the a the more significant the lack of unanimity between experts, which implies that an increase of ambiguity aversion leads to a decrease of the optimal a. This result requires however an important additional assumption relative to Proposition 8. Indeed Proposition 8 assumes that experts' beliefs are comparable in terms of first-order stochastic dominance. This means that experts can unambiguously be ranked in terms of optimism (or pessimism), ruling out cases where some experts look more optimistic than others on some aspects but less optimistic on other aspects. For example, it rules out the case discussed in our introductory example where expert 2 looks more optimistic than expert 1 as she predicts that the project may generate positive returns but also look more pessimistic than expert 1 as she foresees cases where the project would generate negative returns. In such a case, Proposition 8 has no bite.

The following section provides an example where utility-averaging and probability-averaging decision-rules may provide opposite conclusions as experts' beliefs cannot be ordered in terms of first-order stochastic dominance.

2.2.2 More about the precautionary savings example

Let us consider again the precautionary example described in Section 2.1.2, while specifying further the utility functions, the income values, and expert beliefs.

To simplify, assume that instantaneous utilities are quadratic, with $v(c) = c - \frac{1}{8}c^2$, the discount factor is $\beta = 1$ and the probability transformation function f is the identity function. Consider the case where first period income is $y = 2$, while the second period income can only take one of the following three values $\omega_1 = 0$, $\omega_2 = 1$ or $\omega_3 = 3$. The rate of interest is $r = 0$. Assume that there are only two experts. According to expert 1, the likelihood of these three values for the second period income are given by the probability vector $(\pi_1^1, \pi_2^1, \pi_3^1) = (0, 1, 0)$. Expert 2 disagrees and holds belief $(\pi_1^2, \pi_2^2, \pi_3^2) = (0.6, 0, 0.4)$. Finally, consider the KMM kind of averager:

$$(x_1, x_2) \in \mathbb{R} \mapsto I(x_1, x_2) = -\frac{1}{\lambda} \log\left(\frac{1}{2}e^{-\lambda x_1} + \frac{1}{2}e^{-\lambda x_2}\right) \tag{10}$$

with $\lambda \geq 0$. This averager is defined over $\mathbb{R}^N$ and will be used either to aggregate probabilities (for probability-averaging decision-rules) or to aggregate utility values (for utility-averaging decision-rules). The limit case when $\lambda \to 0$, provides $I(x_1, x_2) = \frac{1}{2}x_1 + \frac{1}{2}x_2$, corresponding to a linear opinion pooling with symmetric weights. In all cases, increasing λ involves increasing ambiguity aversion. This also generates an increase in disagreement aversion when using probability-averaging decision-rules.

Denote by $a^*_{PA,\lambda}$ the optimal amount of saving when using the probability-averaging decision-rule with averager (10) and by $a^*_{UA,\lambda}$ the optimal action when using a utility-averaging decision-rule with that same averager.

Proposition 9. *The optimal saving level $a^*_{PA,\lambda}$ is increasing with λ (i.e. with disagreement aversion) and the optimal saving level $a^*_{UA,\lambda}$ is decreasing with λ (i.e with ambiguity aversion).*

Proof. See Appendix A.9. ■

The above result, which would actually extend to all KMM kinds of averager, indicates that we have here an example where disagreement aversion leads to larger precautionary savings while ambiguity aversion in the utility-averaging models is found to have the opposite result.

To gain intuition, we report in Figure 3 how $a^*_{PA,\lambda}$ and $a^*_{UA,\lambda}$ vary with λ, as well as the optimal amount of savings a^*_1 and a^*_2 that would be chosen if relying exclusively on the beliefs of expert 1 or on those of expert 2. Expert 2 predicts a lower marginal benefit of precautionary savings than expert 1 in terms of expected utility. This explains why $a^*_2 = 0.4$ is lower than $a^*_1 = 0.5$. Besides, expert 2 predicts a lower expected utility than expert 1. Indeed, although expert 2 predicts a higher second-period income, she also predicts a more risky second period income, which overall yields a lower expected utility. Thus an increase of precautionary savings widens the distance between the expected utility levels computed with the two expert beliefs. This finally implies that an increase of ambiguity aversion leads to a decrease of the optimal precautionary savings.

As a consequence of the utility-averaging decision-rule, $a^*_{UA,\lambda}$ always remains in between a^*_2 and a^*_1. By contrast, probability-averaging decision-rules can entail a saving level $a^*_{PA,\lambda}$ that is larger than both a^*_1 and a^*_2. This actually occurs when disagreement aversion is strong enough. Naturally, one may wonder whether it can be rational to save more than what both experts suggest. If the decision-maker could be sure that one expert is right, without knowing necessarily which one, such a decision would be hardly tenable, as saving a^*_1 would yield a higher utility for

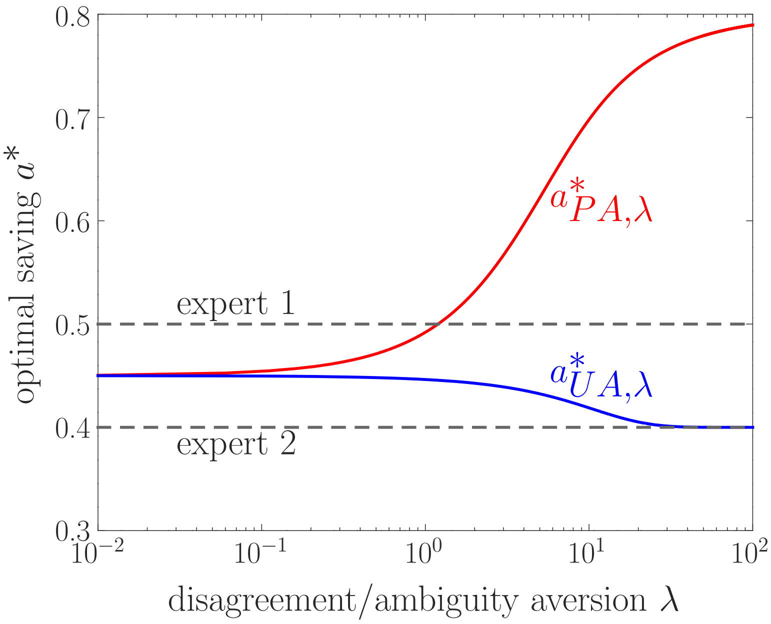

Figure 3 – Optimal saving level as a function of λ for the probability-averaging decision-rule ($a^*_{PA,\lambda}$ in red) and the utility-averaging decision-rule ($a^*_{UA,\lambda}$ in blue), as well as optimal saving levels with expert 1's beliefs and expert 2's beliefs ($a^*_1 = 0.5$ and $a^*_2 = 0.4$ in dashed lines).

sure. However, is there any reason to think that at least one expert is right? A decision-maker who observes that experts disagree may conclude that there is some uncertainty in scientific knowledge and that most likely none of the experts is fully right. The underlying principle of the probability-averaging decision-rule is that the decision-maker uses the diversity of expert opinions to infer "the range of possibilities", and then decides to form her own belief using an aggregation procedure which is deliberately cautious in order to reflect disagreement aversion.

3 Conclusion

Decision-makers rely on scientific expertise. It is only natural, and even a sign of healthy science, that experts disagree. But how should decision-makers deal with expert disagreement? Most prominent is the Bayesian approach: The decision-maker forms a (weighted) average of expert views. While consistent and wellunderstood, the Bayesian approach has been criticized for implying that a situation of scientific consensus and a situation of wide expert disagreement around the same average belief are treated the same.

In this paper, we have developed a model of cautious aggregation of beliefs, exhibiting a novel notion of disagreement aversion. In line with existing contributions that have paved the way to deviate from the Bayesian approach and model aversion to expert disagreement, we borrow from the literature on ambiguity aversion. But instead of the pessimistic aggregation of utility values that is standard in the literature, our model rests on the pessimistic aggregation of expert beliefs. We demonstrate that our model is more sensitive to disagreement than common models: It allows to be averse to "spurious unanimity", a situation in which experts agree on the action the decision-maker should take but offer potentially widely different reasons for their assessment.

Our paper has relevance for the long-standing discussion about the appropriate institutional interplay of scientific expertise and policy-makers, in particular the interplay of risk assessment and risk management (National Research Council U.S.). Our work strengthens the view that the realms of expertise and decision should be kept separate. According to our approach, experts provide policy-makers only with their scientific assessment, not with a preference ordering over prospects. Experts should submit their scientific assessment as clearly as possible, and even deliberate among themselves, but decisions should not be fully delegated to expert committees. Indeed, observing why experts may agree or disagree on a decision (i.e. accounting for the discussion that may take place in committees of experts) is a valuable source of information that can be used by a decision-maker – in addition to the recommendations that expert committees could provide.

References

Aczél, J., and Hansjorg Oser. 2006. *Lectures on functional equations and their applications.* Mineola, N.Y:Dover Publications. 39

Alon, Shiri, and Gabi Gayer. 2016. "Utilitarian Preferences

With Multiple Priors." *Econometrica*, 84(3): 1181–1201. _eprint: https://onlinelibrary.wiley.com/doi/pdf/10.3982/ECTA12676. 6

Anscombe, F. J., and R. J. Aumann. 1963. "A Definition of Subjective Probability." *The Annals of Mathematical Statistics*, 34(1): 199–205. ArticleType: research-article / Full publication date: Mar., 1963 / Copyright © 1963 Institute of Mathematical Statistics. 7, 12

Baillon, Aurélien, Laure Cabantous, and Peter P. Wakker. 2012. "Aggregating imprecise or conflicting beliefs: An experimental investigation using modern ambiguity theories." *Journal of Risk and Uncertainty*, 44(2): 115–147. 3

Basili, Marcello, and Alain Chateauneuf. 2020. "Aggregation of experts' opinions and conditional consensus opinion by the Steiner point." *International Journal of Approximate Reasoning*, 123: 17–25. 6

Berger, Loïc. 2014. "Precautionary saving and the notion of ambiguity prudence." *Economics Letters*, 123(2): 248–251. 26

Berger, Loïc, Johannes Emmerling, and Massimo Tavoni. 2016. "Managing catastrophic climate risks under model uncertainty aversion." *Management Science*. 6, 26

Berger, Loïc, Nicolas Berger, Valentina Bosetti, Itzhak Gilboa, Lars Peter Hansen, Christopher Jarvis, Massimo Marinacci, and Richard D. Smith. 2021. "Rational policymaking during a pandemic." *Proceedings of the National Academy of Sciences*, 118(4). Publisher: National Academy of Sciences Section: Perspective. 3, 6

Bommier, Antoine. 2017. "A dual approach to ambiguity aversion." *Journal of Mathematical Economics*, 71: 104–118. 4, 6, 13, 15, 45, 47, 49

Brandl, Florian. 2020. "Belief-Averaged Relative Utilitarianism." 32. 6

Castagnoli, E., and M. Li Calzi. 1996. "Expected utility without utility." *Theory and Decision*, 41(3): 281–301. 9

Cerreia-Vioglio, Simone, Paolo Ghirardato, Fabio Maccheroni, Massimo Marinacci, and Marciano Siniscalchi. 2011. "Rational preferences under ambiguity." *Economic Theory*, 48(2): 341–375. 17, 18

Chateauneuf, Alain. 1999. "Comonotonicity axioms and rank-dependent expected utility theory for arbitrary consequences." *Journal of Mathematical Economics*, 32(1): 21–45. 14, 45, 48, 49

Crès, Hervé, Itzhak Gilboa, and Nicolas Vieille. 2011. "Aggregation of multiple prior opinions." *Journal of Economic Theory*, 146(6): 2563–2582. 3, 6

Danan, Eric, Thibault Gajdos, Brian Hill, and Jean-Marc Tallon. 2016. "Robust Social Decisions." *American Economic Review*, 106(9): 2407–2425. 6

DeGroot, Morris H., and Julia Mortera. 1991. "Optimal Linear Opinion Pools." *Management Science*, 37(5): 546–558. 6

Gajdos, Thibault, and Jean-Christophe Vergnaud. 2013. "Decisions with conflicting and imprecise information." *Social Choice and Welfare*, 41(2): 427–452. 3, 4, 6

Gajdos, T., T. Hayashi, J. M. Tallon, and J. C. Vergnaud. 2008. "Attitude toward imprecise information." *Journal of Economic Theory*, 140(1): 27–65. 2, 3

Genest, Christian, and James V. Zidek. 1986. "Combining probability distributions: A critique and an annotated bibliography." *Statistical Science*, 1(1): 114–135. Publisher: Institute of Mathematical Statistics. 6

Ghirardato, Paolo, Fabio Maccheroni, and Massimo Marinacci. 2004. "Differentiating ambiguity and ambiguity attitude." *Journal of Economic Theory*, 118(2): 133–173. 01001. 20

Gilboa, Itzhak, Dov Samet, and David Schmeidler. 2004. "Utilitarian Aggregation of Beliefs and Tastes." *Journal of Political Economy*, 112(4): 932–938. Publisher: The University of Chicago Press. 3, 6

Gilboa, Itzhak, Larry Samuelson, and David Schmeidler. 2014. "No-Betting-Pareto Dominance." *Econometrica*, 82(4): 1405–1442. _eprint: https://onlinelibrary.wiley.com/doi/pdf/10.3982/ECTA11281. 6

Gollier, Christian. 2011. "Portfolio Choices and Asset Prices: The Comparative Statics of Ambiguity Aversion." *The Review of Economic Studies*, 78(4): 1329–1344. 26

Hansen, Lars Peter, and Thomas J. Sargent. 2001. "Robust Control and Model Uncertainty." *The American Economic Review*, 91(2): 60–66. 6

Harsanyi, John C. 1955. "Cardinal Welfare, Individualistic Ethics, and Interpersonal Comparisons of Utility." *Journal of Political Economy*, 14. 6

Heal, Geoffrey, and Antony Millner. 2018. "Uncertainty and ambiguity in environmental economics: conceptual issues." *Handbook of Environmental Economics*, 30. 6

Hill, Brian. 2012. "Unanimity and the aggregation of multiple prior opinions." 25. 3

Hylland, Aanund, and Richard Zeckhauser. 1979. "The Impossibility of Bayesian Group Decision Making with Separate Aggregation of Beliefs and Values." *Econometrica*, 47(6): 1321–1336. Publisher: [Wiley, Econometric Society]. 6

Jose, Victor Richmond R., Yael Grushka-Cockayne, and Kenneth C. Lichtendahl. 2013. "Trimmed Opinion Pools and the Crowd's Calibration Problem." *Management Science*, 60(2): 463–475. Publisher: INFORMS. 6

Kets, W., and Marciano M. Siniscalchi. 2016. "A generalized notion of probabilistic sophistication that applies to ambiguous beliefs." 3

Kihlstrom, Richard E, and Leonard J Mirman. 1974. "Risk aversion with many commodities." *Journal of Economic Theory*, 8(3): 361–388. 12

Klibanoff, Peter, Massimo Marinacci, and Sujoy Mukerji. 2005. "A Smooth Model of Decision Making under Ambiguity." *Econometrica*, 73(6): 1849–1892. 17

Larrick, Richard P., and Jack B. Soll. 2006. "Intuitions About Combining Opinions: Misappreciation of the Averaging Principle." *Management Science*, 52(1): 111–127. Publisher: INFORMS. 6

Maccheroni, Fabio, Massimo Marinacci, and Aldo Rustichini. 2006. "Ambiguity Aversion, Robustness, and the Variational Representation of Preferences." *Econometrica*, 74(6): 1447–1498. _eprint: https://onlinelibrary.wiley.com/doi/pdf/10.1111/j.1468-0262.2006.00716.x. 17

Machina, Mark J. 2014. "Ambiguity Aversion with Three or More Outcomes." *American Economic Review*, 104(12): 3814–3840. 3, 13, 18

Machina, Mark J., and Marciano Siniscalchi. 2014. "Chapter 13 - Ambiguity and Ambiguity Aversion." In *Handbook of the Economics of Risk and Uncertainty*. Vol. 1 of *Handbook of the Economics of Risk and Uncertainty*, , ed. Mark Machina and Kip Viscusi, 729–807. North-Holland. 6

McConway, K. J. 1981. "Marginalization and Linear Opinion Pools." *Journal of the American Statistical Association*, 76(374): 410–414. Publisher: [American Statistical Association, Taylor & Francis, Ltd.]. 6

Meinshausen, Malte, Nicolai Meinshausen, William Hare, Sarah C. B. Raper, Katja Frieler, Reto Knutti, David J. Frame, and Myles R. Allen. 2009. "Greenhouse-gas emission targets for limiting global warming to 2°C." *Nature*, 458(7242): 1158–1162. 23

Millner, Antony, Simon Dietz, and Geoffrey Heal. 2013. "Scientific Ambiguity and Climate Policy." *Environmental and Resource Economics*, 55(1): 21–46. 6, 26

Mongin, Philippe. 1995. "Consistent Bayesian Aggregation." *Journal of Economic Theory*, 66(2): 313–351. 3, 6

Mongin, Philippe. 1998. "The paradox of the Bayesian experts and state-dependent utility theory." *Journal of Mathematical Economics*, 29(3): 331–361. 3

Mongin, Philippe. 2016. "Spurious unanimity and the Pareto principle." *Economics & Philosophy*, 32(3): 511–532. Publisher: Cambridge University Press. 3, 6, 9

Morris, Peter A. 1977. "Combining Expert Judgments: A Bayesian Approach." *Management Science*, 16. 6

Nascimento, Leandro. 2012. "The ex ante aggregation of opinions under uncertainty: Aggregation of opinions under uncertainty." *Theoretical Economics*, 7(3): 535–570. 3

National Research Council (U.S.), ed. 2009. *Science and decisions: advancing risk assessment.* Washington, D.C:National Academies Press. 30

Nau, Robert F. 2006. "Uncertainty Aversion with Second-Order Utilities and Probabilities." *Management Science*, 52(1): 136–145. 17

Peter, Richard. 2019. "Revisiting precautionary saving under ambiguity." *Economics Letters*, 174: 123–127. 26

Quiggin, John. 1982. "A theory of anticipated utility." *Journal of Economic Behavior & Organization*, 3(4): 323–343. 9

Savage, Leonard J. 1954. *The Foundations of Statistics.* Courier Corporation. 16474 Google-Books-ID: zSv6dBWneMEC. 7

Siniscalchi, Marciano. 2009. "Vector Expected Utility and Attitudes Toward Variation." *Econometrica*, 77(3): 801–855. 17

Skiadas, Costis. 2013. "Scale-invariant uncertainty-averse preferences and source-dependent constant relative risk aversion." *Theoretical Economics*, 8(1): 59–93. _eprint: https://onlinelibrary.wiley.com/doi/pdf/10.3982/TE1004. 3

Stanca, Lorenzo. 2021. "Smooth aggregation of Bayesian experts." *Journal of Economic Theory*, 196: 105308. 3, 6

Stone, M. 1961. "The Opinion Pool." *The Annals of Mathematical Statistics*, 32(4): 1339–1342. Publisher: Institute of Mathematical Statistics. 6

Strzalecki, Tomasz. 2013. "Temporal Resolution of Uncertainty and Recursive Models of Ambiguity Aversion." *Econometrica*, 81(3): 1039–1074. _eprint: https://onlinelibrary.wiley.com/doi/pdf/10.3982/ECTA9619. 13

Wakker, Peter. 1993. "Additive representations on rank-ordered sets: II. The topological approach." *Journal of Mathematical Economics*, 22(1): 1–26. 45, 49

Yaari, Menahem E. 1969. "Some remarks on measures of risk aversion and on their uses." *Journal of Economic Theory*, 1(3): 315–329. 00000. 11

A Theoretical and applied properties

A.1 Proof of Proposition 1

For any expertise $\mathscr{P}$, notice that to any prospect α, we can associate a unique vector $(\sigma_1,\ldots,\sigma_{K_\alpha})$ (defined as in Definition 5) and a unique matrix $\left(q_k^i\right)_{i,k}$ of transformed probabilities ($q_k^i = f\left(p_k^i\right), \forall i,k$) as in the Proposition's statement. Reciprocally, to any such pair $(\sigma_k)_k$, $\left(q_k^i\right)_{i,k}$, we can construct a corresponding prospect α, even if it means choosing a suitable expertise $\mathscr{P} = (P_1,\ldots,P_N)$. Consider either a prospect α or its corresponding pair $(\sigma_k)_k$, $\left(q_k^i\right)_{i,k}$. As the decision-rule is RDU on distribution-consensual prospects, if α is distribution-consensual, we have $U^{P_i}(\alpha) = U^{\mathscr{P}}(\alpha), \forall i$ so that $\sum_{k=1}^{K_\alpha}\sigma_k I\left(q_k^1,\ldots,q_k^N\right) = U^{\mathscr{P}}(\alpha) = \max_i U^{P_i}(\alpha) = \max_{1\leq i\leq N}\sum_{k=1}^{K_\alpha}\sigma_k q_k^i$. Now, notice that α is non distribution-consensual if and only if there are indices $i,j,k.$ such that $q_k^i \neq q_k^j$ and $\sigma_k > 0$. Take any such $(\sigma_k)_k$, $\left(q_k^i\right)_{i,k}$ and consider the outcome x such that $u(x) = \max_i U^{P_i}(\alpha)$, so that $x \succsim^{P_i} \alpha$, $\forall i$. Then, if the decision-rule is disagreement-averse, we have $x \succ^{\mathscr{P}} \alpha$, i.e. $\max_i \sum_{k=1}^{K_\alpha}\sigma_k q_k^i > \sum_{k=1}^{K_\alpha}\sigma_k I\left(q_k^1,\ldots,q_k^N\right)$. Reciprocally, take any non distribution-consensual prospect α and consider any outcome x such that $x \succsim^{P_i} \alpha$, $\forall i$. If the Proposition's condition holds, we have $U^{\mathscr{P}}(\alpha) < \max_i U^{P_i}(\alpha) \leq u(x)$, i.e. $x \succ^{\mathscr{P}} \alpha$, so the decision-rule is disagreement averse.

A.2 Proof of Proposition 2

For the first condition:

$$I\left(q_k^1,\ldots,q_k^N\right) \leq \sum_{i=1}^{N}\lambda_i q_k^i \Rightarrow \sum_{k=1}^{K}\sigma_k I\left(q_k^1,\ldots,q_k^N\right) \leq \sum_{i=1}^{N}\lambda_i \sum_{k=1}^{K}\sigma_k q_k^i \leq \max_{1\leq i\leq N}\sum_{k=1}^{K}\sigma_k q_k^i$$

For the second condition: Set $\sigma_{K+1} = 1 - \sum_{k=1}^{K}\sigma_k$ and $q_{K+1}^i = 0$ for all $1 \leq i \leq N$. Then, using successively that I is concave and increasing:

$$\sum_{k=1}^{K+1}\sigma_k I\left(q_k^1,\ldots,q_k^N\right) \leq I\left(\sum_{k=1}^{K+1}\sigma_k q_k^1,\ldots,\sum_{k=1}^{K+1}\sigma_k q_k^N\right) \leq \max_{1\leq i\leq N}\sum_{k=1}^{K}\sigma_k q_k^i$$

In both cases the first inequality is strict when one has $q_k^i \neq q_k^j$ and $\sigma_k > 0$ for some indices i, j, k.

A.3 Proof of Propositions 3 and 5

We first prove Proposition 3 and then explain how the proof can be adapted to prove Proposition 5.

Proof of Proposition 3 If $\succsim_A$ and $\succsim_B$ are comparable in terms of disagreement aversion, then they coincide on distribution-consensual prospects. Indeed, assume without loss of generality that $\succsim_A$ exhibits greater disagreement aversion than $\succsim_B$ and take any pair of distribution-consensual prospects α, β. Then, by two successive applications of the second condition in Definition 3, $\alpha \sim_A^{\mathscr{P}} \beta \Rightarrow \left(\alpha \succsim_B^{\mathscr{P}} \beta \text{ and } \beta \succsim_B^{\mathscr{P}} \alpha\right)$, i.e. $\alpha \sim_B^{\mathscr{P}} \beta$. Moreover, the first condition gives $\alpha \succ_A^{\mathscr{P}} \beta \Rightarrow \alpha \succ_B^{\mathscr{P}} \beta$ and $\beta \succ_A^{\mathscr{P}} \alpha \Rightarrow \beta \succ_B^{\mathscr{P}} \alpha$. Reciprocally, the contraposition of the last three implications gives $\alpha \nsim_B^{\mathscr{P}} \beta \Rightarrow \alpha \nsim_A^{\mathscr{P}} \beta$, $\beta \succsim_B^{\mathscr{P}} \alpha \Rightarrow \beta \succsim_A^{\mathscr{P}} \alpha$, and $\alpha \succsim_B^{\mathscr{P}} \beta \Rightarrow \alpha \succsim_A^{\mathscr{P}} \beta$. Thus, if $\succsim_A$ and $\succsim_B$ are comparable in terms of disagreement aversion, they share the same RDU representation on distribution-consensual prospects, and we shall denote u and f their common utility index and probability transformation.

$\Rightarrow$ Take any expertise $\mathscr{P}$ and any non-constant vector $\vec{p} = (p_1, \ldots, p_N)$. Let $\vec{q} = (f(p_1), \ldots, f(p_N))$. Define $x = u^{-1}(I_A(\vec{q}))$. Denote by $\left(\vec{X}, \vec{p}\right)$ the prospect α with only extremal outcomes (i.e. $\alpha_k \in \{X^-, X^+\}, \forall k$) and such that $F_\alpha^{P_i}(X^+) = p_i, \forall i$. By comparative disagreement aversion, $\left(\vec{X}, \vec{p}\right) \sim_A^{\mathscr{P}} x \Rightarrow \left(\vec{X}, \vec{p}\right) \succ_B^{\mathscr{P}} x$. By definition, $\left(\vec{X}, \vec{p}\right) \sim_A^{\mathscr{P}} x$ if and only if $U_A\left(\left(\vec{X}, \vec{p}\right)\right) = u(x) = I_A(\vec{q})$, which holds by assumption. Thus, $\left(\vec{X}, \vec{p}\right) \succ_B^{\mathscr{P}} x$, i.e. $I_A(\vec{p}) < I_B(\vec{p})$.

$\Leftarrow$ Take any non distribution-consensual prospect α, distribution-consensual prospect β, and expertise $\mathscr{P}$ such that $\alpha \sim_A^{\mathscr{P}} \beta$. Defining $\left(p_k^i\right)$ the probabilities associated to α, notice that $\vec{p}_k = \left(p_k^1, \ldots, p_k^N\right)$ is non-constant for some k so that $I_A(\vec{p}_k) < I_B(\vec{p}_k)$ for those k. For the remaining k where $\vec{p}_k = (p_k, \ldots, p_k)$ is constant, $I_A(\vec{p}_k) = I_B(\vec{p}_k) = p_k$. As $\succsim_A$ and $\succsim_B$ share the representation functions f and u and by Definition 5 or 6, this implies $U_B^{\mathscr{P}}(\beta) = U_A^{\mathscr{P}}(\beta) = U_A^{\mathscr{P}}(\alpha) < U_B^{\mathscr{P}}(\alpha)$, i.e.

$\alpha \succ_B^{\mathscr{P}} \beta$.

Proof of Proposition 5. The $\Rightarrow$ direction can be proven just as above. For $\Leftarrow$, i.e. to prove the comparative ambiguity aversion, one needs to cover the three cases of Definition 7. To prove the last case, where $\alpha \sim_A \beta$ is non-utility-consensual, one just needs to adjust the $\Leftarrow$ part of the above proof, by taking a non-utility consensual (instead of non-distribution consensual) α. To prove the other cases, it suffices to use that $I_A \leq I_B$ in the representations U_A and U_B.

A.4 Link between the Pareto condition and linear pooling

We give here only the sketch of the proof of Proposition 4, which states that a decision-rule is both utility-averaging and probability-averaging if and only if it is a linear pooling. Indeed, we prove rigorously a stronger result in online Appendix A.5: that a probability-averaging decision-rule is Paretian if and only if it is a linear pooling. That a linear pooling is all at the same time probability-averaging, utility-averaging, and hence Paretian, can be seen from the Definitions. Now, if a decision-rule is both utility-averaging and probability-averaging, it accepts two representations, which are equal up to an increasing bijection. Considering distribution-consensual prospects, and then looking successively at prospects with only one outcome, with only extremal outcomes, and with two outcomes (one of which we vary), it can be shown that the bijection is indeed the identity function, and that the two representations share the same functions u and f. Considering non-distribution-consensual prospects with only extremal outcomes, it follows that the two representations also share the averager I. Considering prospects such that $\sigma_k = \frac{1}{K_\alpha}$, $\forall k$, we thus obtain the functional equation for I:

$$\sum_{k=1}^{K_\alpha} \frac{1}{K_\alpha} I\left(f\left(p_k^1\right), \ldots, f\left(p_k^N\right)\right) = I\left(\sum_{k=1}^{K_\alpha} \frac{1}{K_\alpha} f\left(p_k^1\right), \ldots, \sum_{k=1}^{K_\alpha} \frac{1}{K_\alpha} f\left(p_k^N\right)\right),$$

which holds on all $\left(p_k^i\right)$ such that $p_1^i \geq \ldots \geq p_{K_\alpha}^i$, $\forall i$. Without the latter restriction on the domain, this would exactly be Jensen's functional equation, whose solution is known to be affine (e.g. Aczél and Oser 2006). To handle the domain restriction, one can notice that the solution applies locally to any neighborhood in the interior of the domain, and use the connectedness of the domain to show that the affine function is the same on all these neighborhoods. Given that $I(0,\ldots,0)=0$, the averager I is here linear.

A.5 Link between the Pareto condition and linear pooling

Proposition 10. *Let $\succcurlyeq$ be a probability-averaging decision-rule. $\succcurlyeq$ is Paretian if and only if there are positive weights $\vec{\lambda}=(\lambda_1,\ldots,\lambda_N)$ summing to 1 such that $I(\vec{p})=\sum_{i=1}^{N}\lambda_i p_i, \forall \vec{p}$.*

Proof. Denote by (u,f,I) the representation of $\succcurlyeq$.

$\Leftarrow$ Suppose that there exist weights $\vec{\lambda}$ as in the Proposition's statement, so that $U^{\mathscr{P}}=\sum_{i=1}^{N}\lambda_i U^{P_i}$. Take any prospects α,β and expertise $\mathscr{P}$ such that $\beta \succcurlyeq^{P_i} \alpha, \forall i$. By Definition 5, $U^{P_i}(\beta)\geq U^{P_i}(\alpha), \forall i$ and by assumption, $U^{\mathscr{P}}(\beta)=\sum_{i=1}^{N}\lambda_i U^{P_i}(\beta)\geq \sum_{i=1}^{N}\lambda_i U^{P_i}(\alpha)=U^{\mathscr{P}}(\alpha)$, so that $\beta \succcurlyeq^{\mathscr{P}} \alpha$.

$\Rightarrow$[15] The remainder of the proof consists of four steps. First, we construct a utility-consensual three-outcome prospect with a particular structure (namely, that the cumulative probabilities of the different experts for a given outcome are confined in an interval that does not overlap with the intervals of the other outcomes). We also construct very close prospects by holding the probabilities attached to one outcome and varying by infinitesimal but opposite amounts the probabilities attached to the two other outcomes by a given expert. Second, as these prospects are utility-consensual (for the same utility level), and by the Pareto condition, we equate their representation formulas and obtain a property for the function I (equation 11 below). This property is akin to the equality of some partial derivative at two points, i.e. vectors; we then transform it so that the vectors involved differ only along one variable. Third, we use the Lemma 2 below to show that this local property amounts to linearity along one variable (equation 12). Last, we assemble the linearity along each variable to prove the linearity of I.

1. Take $\vec{p},\vec{b}\in(0,1)^N$ such that $\vec{p}\gg\vec{b}$ (i.e. $\forall i, p_i>b_i$). Let us construct $y>0$, weights $\sigma_1,\sigma_2,\sigma_3$ and a finite sequence $\mathbf{q}=(\vec{q}_k)_{k\leq 3}=\left(\vec{p},\vec{b},\vec{q}_3\right)$ that is strictly decreasing (i.e. such that $\vec{q}_1\gg\vec{q}_2\gg\vec{q}_3$,), and such that $\sum_{k=1}^{3}\sigma_k f\left(q_{i,k}\right)=y, \forall i$.

[15] We are grateful to Rémi Peyre who gave us the idea of the proof.

Define $z=\frac{1}{2}\min_i\{b_i\}$, $\bar{z}=(z,\dots,z)$ and $\delta=\frac{z}{2}\left(\max_i\{f(p_i)+f(b_i)\}\right)^{-1}$ so that $\vec{b}\gg\bar{z}\gg\delta\left(f(\vec{p})+f\left(\vec{b}\right)\right)$.[16] Define $\vec{q}_3=\bar{z}-\delta\left(f(\vec{p})+f\left(\vec{b}\right)\right)$, $\sigma_1=\sigma_2=\frac{1}{\delta^{-1}+2}$, $\sigma_3=\frac{\delta^{-1}}{\delta^{-1}+2}$, and $y=\frac{\delta^{-1}}{\delta^{-1}+2}z$. It is clear that $\vec{b}\gg\vec{q}_3$ and by construction, $\sum_{k=1}^{3}\sigma_k f\left(q_{i,k}\right)=y,\forall i$. For $k\in\{1,\dots,N\}$, denote $\vec{\mathbb{1}}_k\in\{0,1\}^N$ the vector with ith component $\mathbb{1}_{i=k}$. For $\varepsilon>0$ and $k\in\{1,\dots,N\}$, define the sequence $\mathbf{q}^{\varepsilon,k}=\left(\vec{q}_k^{\,\varepsilon,k}\right)_{k\leq 3}$ by $\vec{q}_1^{\,\varepsilon,k}=f^{-1}\left(f(\vec{p})-\vec{\mathbb{1}}_k\varepsilon\right)$, $\vec{q}_2^{\,\varepsilon,k}=f^{-1}\left(f\left(\vec{b}\right)+\vec{\mathbb{1}}_k\varepsilon\right)$, and $\vec{q}_3^{\,\varepsilon,k}=\vec{q}_3$. We have $\sum_{k=1}^{3}\sigma_k f\left(q_{i,k}^{\varepsilon,k}\right)=y$ and for $\varepsilon>0$ sufficiently small, $\left(\vec{q}_k^{\,\varepsilon,k}\right)_k$ is strictly decreasing. For $k\in\{1,2,3\}$, define $x_k=u^{-1}(\sigma_k)$. Denote $\vec{x}=(x_k)_{k\in\{1,2,3\}}$. Take an expertise $\mathscr{P}=(P_1,\dots,P_N)$ such that for all $\varepsilon\geq 0$ and all $k\in\{1,\dots,N\}$, there exists a prospect with outcomes $\vec{x}$ and probabilities $\mathbf{q}^{\varepsilon,k}$, and denote by $\left(\vec{x},\mathbf{q}^{\varepsilon,k}\right)$ any such prospect (or simply $(\vec{x},\mathbf{q})$ for $\varepsilon=0$). Remark 5 in Appendix B.3 shows that these objects are well defined. By construction and Definition 5, $U^{P_i}((\vec{x},\mathbf{q}))=\sum_{k=1}^{3}\sigma_k f\left(q_{i,k}\right)=y,\forall i$ so by the Pareto condition, $U^{\mathscr{P}}((\vec{x},\mathbf{q}))=\sum_{k=1}^{3}\sigma_k f\left(\widetilde{I}(\vec{q}_k)\right)=y$. Similarly, $U^{P_i}\left(\left(\vec{x},\mathbf{q}^{\varepsilon,k}\right)\right)=\sum_{k=1}^{3}\sigma_k f\left(q_{i,k}^{\varepsilon,k}\right)=y,\forall i,k$ so that $U^{\mathscr{P}}\left(\left(\vec{x},\mathbf{q}^{\varepsilon,k}\right)\right)=\sum_{k=1}^{3}\sigma_k f\left(\widetilde{I}\left(\vec{q}_k^{\,\varepsilon,k}\right)\right)=y,\forall k$. Fix $k\in\{1,\dots,N\}$.

2. From $U^{\mathscr{P}}\left(\left(\vec{x},\mathbf{q}^{\varepsilon,k}\right)\right)=U^{\mathscr{P}}((\vec{x},\mathbf{q}))$, we get

$$I\left(f(\vec{p})-\vec{\mathbb{1}}_k\varepsilon\right)+I\left(f\left(\vec{b}\right)+\vec{\mathbb{1}}_k\varepsilon\right)=I(f(\vec{p}))+I\left(f\left(\vec{b}\right)\right). \tag{11}$$

Define $\vec{p}_{\vec{b}}^{\,k}:=f(\vec{p})+\vec{\mathbb{1}}_k(f(b_k)-f(p_k))$, $\vec{b}_{\vec{p}}^{\,k}:=f\left(\vec{b}\right)+\vec{\mathbb{1}}_k(f(p_k)-f(b_k))$, and take any vector $\vec{r}$ such that $\vec{1}\gg\vec{r}\gg\vec{p}$. Although equation (11) is written for $\left(\vec{b},\vec{p}\right)$, it gives a relation valid for any pair $(\vec{x},\vec{y})$ with $\vec{y}\gg\vec{x}$. Summing equation 11 with the same relation for $\left(\vec{p}_{\vec{b}}^{\,k},\vec{r}\right)$ and subtracting the relation for $\left(\vec{b},\vec{r}\right)$, we obtain $I\left(f(\vec{p})-\vec{\mathbb{1}}_k\varepsilon\right)+I\left(f\left(\vec{p}_{\vec{b}}^{\,k}\right)+\vec{\mathbb{1}}_k\varepsilon\right)=I(f(\vec{p}))+I\left(f\left(\vec{p}_{\vec{b}}^{\,k}\right)\right)$.

3. Using the latter relation, the Lemma 2 below applies to $J_k^{\vec{p}}:q\mapsto I\left(f(\vec{p})+\vec{\mathbb{1}}_k(q-f(p_k))\right)$, and we obtain the following lin-

[16] For any vector $\vec{p}$, we denote abusively $f(\vec{p}):=(f(p_1),\dots,f(p_N))$.

earity result: there exists $\lambda_k^{\vec{p}} > 0$ such that for all $\varepsilon \in [0, f(p_k)]$, $I\left(f(\vec{p}) - \vec{\mathbb{1}}_k \varepsilon\right) = I(f(\vec{p})) - \lambda_k^{\vec{p}} \varepsilon$. Similarly, the following relation holds: $I\left(f\left(\vec{b}_{\vec{p}}^k\right) - \vec{\mathbb{1}}_k \varepsilon\right) + I\left(f\left(\vec{b}\right) + \vec{\mathbb{1}}_k \varepsilon\right) = I\left(f\left(\vec{b}_{\vec{p}}^k\right)\right) + I\left(f\left(\vec{b}\right)\right)$ and allows to apply Lemma 2 to $J_k^{\vec{b}} : q \mapsto I\left(f\left(\vec{b}_{\vec{p}}^k\right) + \vec{\mathbb{1}}_k (q - f(p_k))\right)$ to show that for all $\varepsilon \in [-f(b_k), 1 - f(b_k)]$,

$$I\left(f\left(\vec{b}\right) + \vec{\mathbb{1}}_k \varepsilon\right) = I\left(f\left(\vec{b}\right)\right) + \lambda_k^{\vec{b}} \varepsilon. \tag{12}$$

Furthermore, injecting these linearity results in equation 11 shows that $\lambda_k^{\vec{b}} = \lambda_k^{\vec{p}}$, which suffices to show that $\lambda_k^{\vec{b}}$ does not depend on $\vec{b}$.

4. As the linearity result 12 holds for all $f\left(\vec{b}\right) \in (0,1)^N$, it also holds for all $f\left(\vec{b}\right) \in [0,1]^N$ by continuity of I. Take any $\vec{q} \in [0,1]^N$. For $k \in \{1, \ldots, N\}$, apply this linearity result successively to $\vec{p}_k = \vec{q} - \sum_{i=1}^k \vec{\mathbb{1}}_i q_i$ and $\varepsilon_k = q_k$. We get $I(\vec{q}) = I\left(\vec{q} - \vec{\mathbb{1}}_1 q_1\right) + \lambda_1 q_1 = I\left(\vec{q} - \sum_{i=1}^k \vec{\mathbb{1}}_i q_i\right) + \sum_{i=1}^k \lambda_i q_i = \sum_{i=1}^N \lambda_i q_i$. ■

Lemma 2. *Let $m > 0$ and let $J : [0,m] \to [0,1]$ a continuous and strictly increasing function such that for all $p > b \in (0,1)$, there exists $\eta > 0$ such that for all $\varepsilon \in [0,\eta]$, $J(p-\varepsilon) + J(b+\varepsilon) = J(p) + J(b)$. Then there exists $\lambda > 0$ such that for all $q \in [0,m]$, $J(q) = J(0) + \lambda q = J(m) - \lambda(m-q)$.*

Proof. Define $\lambda = \frac{1}{m}(J(m) - J(0))$. By definition, $J(0) + \lambda q = J(m) - \lambda(m-q)$. Suppose by the absurd that there exists $q \in [0,m]$ such that $J(q) \neq J(0) + \lambda q$. It is clear that $q \in (0,m)$. Without loss of generality, assume $J(q) < J(0) + \lambda q$. Define $p = \inf\{x > q \,|\, J(x) = J(0) + \lambda x\}$ and $b = \sup\{x < q \,|\, J(x) = J(0) + \lambda x\}$. By continuity of J, $J(p) = J(0) + \lambda p$, $J(b) = J(0) + \lambda b$, and there exists $\varepsilon_p, \varepsilon_b > 0$ such that $\forall x \in [p - \varepsilon_p, p), J(x) < J(0) + \lambda x$ and $\forall x \in (b, b + \varepsilon_b], J(x) < J(0) + \lambda x$. By hypothesis, there exists $\eta > 0$ such that for all $\varepsilon \in [0,\eta]$, $J(p-\varepsilon) + J(b+\varepsilon) = J(p) + J(b)$. Define $\delta = \min\left\{\varepsilon_p, \varepsilon_b, \eta\right\}$. By construction, we have $J(p-\delta) < J(0) + \lambda(p-\delta)$ and $J(b+\delta) < J(0) + \lambda(b+\delta)$. Summing the last two inequalities, we obtain $J(p-\delta) + J(b+\delta) < 2J(0) + \lambda(p+b)$. We also have $J(p-\delta) + J(b+\delta) = J(p) + J(b)$. From the above, this implies $J(p-\delta) + J(b+\delta) = 2J(0) + \lambda(p+b)$, which contradicts the previous inequality. ■

A.6 Proof of Proposition 6

Take any expertise $\mathscr{P}$, prospect α with at least one non-extremal outcome, and distribution-consensual prospect β. As $\succsim_{PA}$ and $\succsim_{UA}$ share the functions u and f, they coincide on distribution-consensual prospects, and we can denote $U^{\mathscr{P}}(\beta) := U^{\mathscr{P}}_{UA}(\beta) = U^{\mathscr{P}}_{PA}(\beta)$. Also, if α is distribution-consensual, $\alpha \sim^{\mathscr{P}}_{PA} \beta \Rightarrow \alpha \sim^{\mathscr{P}}_{UA} \beta$. For the remainder of the proof, we consider the case where α is non distribution-consensual, so that $\left(p^1_k, \ldots, p^N_k\right)$ is non-constant for some k such that $\sigma_k \neq 0$. Set $\sigma_{K_\alpha+1} = 1 - \sum_{k=1}^{K_\alpha} \sigma_k$ and $p^i_{K_\alpha+1} = 0$ for all $1 \leq i \leq N$. The strict concavity inequality yields: $U^{\mathscr{P}}_{PA}(\alpha) = \sum_{k=1}^{K_\alpha+1} \sigma_k I\left(f\left(p^1_k\right), \ldots, f\left(p^N_k\right)\right) < I\left(\sum_{k=1}^{K_\alpha+1} \sigma_k f\left(p^1_k\right), \ldots, \sum_{k=1}^{K_\alpha+1} \sigma_k f\left(p^N_k\right)\right) = U^{\mathscr{P}}_{UA}(\alpha)$. Thus, $\alpha \succsim^{\mathscr{P}}_{PA} \beta \Rightarrow U^{\mathscr{P}}_{UA}(\alpha) > U^{\mathscr{P}}_{PA}(\alpha) \geq U^{\mathscr{P}}(\beta) \Rightarrow \alpha \succ^{\mathscr{P}}_{UA} \beta$. To conclude, notice that we have covered all cases needed to prove that $\succsim_{PA}$ exhibits greater ambiguity aversion and greater disagreement aversion than $\succsim_{UA}$.

A.7 Proof of Proposition 7

The first-order condition of (8) is:

$$\underbrace{\sum_{k=2}^{K} \left(\partial_1 u(a,\omega_k) - \partial_1 u(a,\omega_{k-1})\right) I_\tau\left(f\left(p^1_k\right), \ldots, f\left(p^N_k\right)\right) + \partial_1 u(a,\omega_1)}_{U'_\tau(a)} = 0, \quad (13)$$

given that $I_\tau\left(f\left(p^1_1\right), \ldots, f\left(p^N_1\right)\right) = 1$. Since decision-maker A is more disagreement-averse than B, we have $I_A\left(f\left(p^1_k\right), \ldots, f\left(p^N_k\right)\right) \leq I_B\left(f\left(p^1_k\right), \ldots, f\left(p^N_k\right)\right)$ for all k with strict inequality for some k since the group of experts disagrees. If $\partial_1 u(a,\omega)$ strictly increases with ω, we have $U'_A(a) < U'_B(a)$ and $a^*_A < a^*_B$. If $\partial_1 u(a,\omega)$ strictly decreases with ω, we have $U'_A(a) > U'_B(a)$ and $a^*_A > a^*_B$.

A.8 Proof of Proposition 8

With a smooth "KMM" form (5), the first-order condition of (9) is:

$$\underbrace{\sum_{i=1}^{N}\lambda_i\psi'_\tau\left(\sum_{k=1}^{K}\sigma_k f\left(p_k^i\right)\right)\cdot\left(\sum_{k=2}^{K}\left(\partial_1 u(a,\omega_k)-\partial_1 u(a,\omega_{k-1})\right)f\left(p_k^i\right)+\partial_1 u(a,\omega_1)\right)}_{\frac{U'_\tau(a)}{\left(\psi_\tau^{-1}\right)'\left(\sum_{i=1}^{N}\lambda_i\psi_\tau\left(\sum_{k=1}^{K}\sigma_k f\left(p_k^i\right)\right)\right)}}=0, \tag{14}$$

which rewrites:

$$\sum_{i=1}^{N}\underbrace{\frac{\lambda_i\psi'_\tau\left(\sum_{k=1}^{K}\sigma_k f\left(p_k^i\right)\right)}{\sum_{l=1}^{N}\lambda_l\psi'_\tau\left(\sum_{k=1}^{K}\sigma_k f\left(p_k^l\right)\right)}}_{\tilde\lambda_i(\psi_\tau,a)}\cdot\underbrace{\left(\sum_{k=2}^{K}\left(\partial_1 u(a,\omega_k)-\partial_1 u(a,\omega_{k-1})\right)f\left(p_k^i\right)+\partial_1 u(a,\omega_1)\right)}_{\rho_i(a)}=0. \tag{15}$$

Since the experts are ordered in the sense of first-order stochastic dominance and $\partial_2 u(a,\omega)>0$, we have $\sum_{k=1}^{K}\sigma_k f\left(p_k^1\right)\leq\sum_{k=1}^{K}\sigma_k f\left(p_k^2\right)\leq\ldots\leq\sum_{k=1}^{K}\sigma_k f\left(p_k^N\right)$, with at least one strict inequality given that the group of experts disagrees. In (15), we can view $\tilde\lambda_i(\psi_\tau,a)$ as a distribution function where i would be the random variable. By hypothesis, we have $\psi_A=h\circ\psi_B$ where h is a strictly increasing and strictly concave function. The likelihood ratio of $\tilde\lambda_i(\psi_A,a)$ and $\tilde\lambda_i(\psi_B,a)$ then writes:

$$\frac{\tilde\lambda_i(\psi_A,a)}{\tilde\lambda_i(\psi_B,a)}=h'\left(\psi_B\left(\sum_{k=1}^{K}\sigma_k f\left(p_k^i\right)\right)\right)\cdot\frac{\sum_{l=1}^{N}\lambda_l\psi'_B\left(\sum_{k=1}^{K}\sigma_k f\left(p_k^l\right)\right)}{\sum_{l=1}^{N}\lambda_l\psi'_A\left(\sum_{k=1}^{K}\sigma_k f\left(p_k^l\right)\right)}. \tag{16}$$

Since we have $\sum_{k=1}^{K}\sigma_k f\left(p_k^i\right)$ increasing with i with at least one strict inequality, $\psi'_B>0$, $\psi'_A>0$ and $h''<0$, the likelihood ratio (16) is decreasing with i with at least one strict inequality. Thus, the distribution $\tilde\lambda_i(\psi_B,a)$ strictly dominates the distribution $\tilde\lambda_i(\psi_A,a)$ in the sense of monotone likelihood ratio, which implies that the former strictly first-order stochastically dominates the latter. If $\partial_1 u(a,\omega)$ strictly increases (decreases) with ω, we have $\rho_1(a)\leq\rho_2(a)\leq\ldots\leq\rho_N(a)$ ($\rho_1(a)\geq\rho_2(a)\geq\ldots\geq\rho_N(a)$) with at least one strict inequality, since the experts are ordered in the sense of first-order stochastic dominance and the group of experts disagrees. As a consequence, since $\tilde\lambda_i(\psi_B,a)$ strictly first-order stochastically dominates $\tilde{\lambda}_i(\psi_A,a)$, we get, for a given a, $\sum_{i=1}^N \tilde{\lambda}_i(\psi_A,a)\rho_i(a) < \sum_{i=1}^N \tilde{\lambda}_i(\psi_B,a)\rho_i(a)$ $(\sum_{i=1}^N \tilde{\lambda}_i(\psi_A,a)\rho_i(a) > \sum_{i=1}^N \tilde{\lambda}_i(\psi_B,a)\rho_i(a))$. Thus, with $a = a_B^*$, we have $\sum_{i=1}^N \tilde{\lambda}_i(\psi_A,a_B^*)\rho_i(a_B^*) < 0$ $(\sum_{i=1}^N \tilde{\lambda}_i(\psi_A,a_B^*)\rho_i(a_B^*) > 0)$, and $U_A'(a_B^*) < 0$ $(U_A'(a_B^*) > 0)$. Finally, we have $a_A^* < a_B^*$ $(a_A^* > a_B^*)$.

A.9 Proof of Proposition 9

A direct application of Proposition 7 shows that the optimal saving level $a_{PA,\lambda}^*$ is increasing with λ. On the other hand, we cannot apply Proposition 8 since experts are not ordered in terms of first-order stochastic dominance. Following a reasonning similar to the proof of Proposition 8, we show that the optimal saving level $a_{UA,\lambda}^*$ is decreasing with λ. With a utility-averaging decision-rule and a smooth "KMM" form, the optimal saving level satisfies (15) with $K = 3$ states of nature, $t_1 = 0$, $t_2 = 1$, $t_3 = 3$, $N = 2$ experts, $p_1^1 = 1$, $p_2^1 = 1$, $p_3^1 = 0$, $p_1^2 = 1$, $p_2^2 = 0.4$, $p_3^2 = 0.4$, $a \in [0,0.9]$ and $u(a,t) = 2 - a - \frac{1}{8}(2-a)^2 + t + a - \frac{1}{8}(t+a)^2$. In contrast with the proof of Proposition 8, we have $\sum_{k=1}^K \sigma_k f\left(p_k^1\right) = 2 - a - \frac{1}{8}(2-a)^2 + 1 + a - \frac{1}{8}(1+a)^2$ strictly larger than $\sum_{k=1}^K \sigma_k f\left(p_k^2\right) = 2 - a - \frac{1}{8}(2-a)^2 + \frac{12}{10} + a - \frac{6}{80}a^2 - \frac{4}{80}(3+a)^2$ since $\sum_{k=1}^K \sigma_k f\left(p_k^1\right) - \sum_{k=1}^K \sigma_k f\left(p_k^2\right) = \frac{1}{8} + \frac{1}{20}a > 0$ for $a \geq 0$. Considering ψ_C more concave than ψ_D, this implies that the likelihood ratio of $\tilde{\lambda}_i(\psi_C,a)$ and $\tilde{\lambda}_i(\psi_D,a)$ is strictly increasing with i, and the distribution $\tilde{\lambda}_i(\psi_D,a)$ is strictly first-order stochastically dominated by the distribution $\tilde{\lambda}_i(\psi_C,a)$. Moreover, we have $\partial_{1,2}u(a,t) = -\frac{1}{4} < 0$ and $\rho_1(a) = -\frac{1}{2}a + \frac{1}{4} > \rho_2(a) = -\frac{1}{2}a + \frac{1}{5}$. Thus, for a given a, $\sum_{i=1}^N \tilde{\lambda}_i(\psi_C,a)\rho_i(a) < \sum_{i=1}^N \tilde{\lambda}_i(\psi_D,a)\rho_i(a)$. This implies $U_C'(a_D^*) < 0$ and $a_C^* < a_D^*$. Finally, this shows that with a KMM kind of averager of the form (10), the optimal saving level $a_{UA,\lambda}^*$ is decreasing with the ambiguity aversion parameter λ.

B Axiomatic construction

In this section, we show how the class of probability-averaging decision-rules can be obtained from a set of axioms. This part heavily relies on Bommier (2017), and thus indirectly on Chateauneuf (1999) and Wakker (1993)

B.1 Axioms

We denote by $\mathscr{L}$ the set of decumulative distribution functions. Given an expertise $\mathscr{P}$, we define the *assessment* of a prospect α as $\mathbf{F}_\alpha^{\mathscr{P}} := \left(F_\alpha^{P^1}, \ldots, F_\alpha^{P^N}\right) \in \mathscr{L}^N$. We shall assume that assessments are the "common language" of expertises, in the sense that the decision-rule is indifferent between two prospects with the same assessment. More precisely, a prospect assessed by a certain expertise is equivalent to a *certainty-equivalent* sure prospect that only depends on its assessment. In other words, disagreement on the probabilities of the states of the world matters only to the extent that it translates into disagreement on the distribution of outcomes.

Axiom 1. Common language of expertises. *For all expertises $\mathscr{P}, \mathscr{P}'$ and all prospects α, α' such that $\mathbf{F}_\alpha^{\mathscr{P}} = \mathbf{F}_{\alpha'}^{\mathscr{P}'}$, there exists $x \in X$ such that $\alpha \sim^{\mathscr{P}} x$ and $\alpha' \sim^{\mathscr{P}'} x$.*

The following axiom is innocuous and mainly states basic continuity conditions.

Axiom 2. Well-behaved weak order. *Take any expertise $\mathscr{P}$. The weak order on prospects, $\succcurlyeq^{\mathscr{P}}$, is non-trivial (i.e. $\exists \alpha, \beta \in X^\Omega, \alpha \succ^{\mathscr{P}} \beta$) and continuous: for all prospect α, the sets $\left\{\beta \in X^\Omega \,|\, \alpha \succcurlyeq^{\mathscr{P}} \beta\right\}$ and $\left\{\beta \in X^\Omega \,|\, \beta \succcurlyeq^{\mathscr{P}} \alpha\right\}$ are closed subsets of X^Ω. Furthermore, for any $\alpha, \beta \in X^\Omega$ and $x_A, x_B \in X$, $x_A \sim^{\mathscr{P}} \alpha \succcurlyeq^{\mathscr{P}} \beta \sim^{\mathscr{P}} x_B \implies x_A \geq x_B$. Finally, the mapping $\mathscr{P} \mapsto \succcurlyeq^{\mathscr{P}}$ is continuous, in the following sense. For all sequence $(\mathscr{P}_i)_{i \in \mathbb{N}}$ that converges to an expertise $\mathscr{P}$ and all sequence $(\alpha_i)_{i \in \mathbb{N}}$ that converges to a prospect α, we have that for all $\beta, \gamma \in X^\Omega$, $\beta \succcurlyeq^{\mathscr{P}_i} \alpha_i \succcurlyeq^{\mathscr{P}_i} \gamma$, $\forall i \in \mathbb{N} \Rightarrow \beta \succcurlyeq^{\mathscr{P}} \alpha \succcurlyeq^{\mathscr{P}} \gamma$.*

The specificity of our axiomatization lies in our monotonicity axiom. This axiom is weaker the Pareto condition. It only requires that a prospect α is preferred to another β if α dominates β with respect to first-order stochastic dominance, i.e. if for all experts the decumulative distribution function of α is never below that of β.

Axiom 3. Monotonicity with respect to first-order stochastic dominance. *For all expertise $\mathscr{P} = (P_1, \ldots, P_N)$ and all prospects α, β, if $F_\alpha^{P_i} \geq F_\beta^{P_i}$ for all $i \in \{1, \ldots, N\}$, then $\alpha \succcurlyeq^{\mathscr{P}} \beta$, and if in addition $F_\alpha^{P_i} \neq F_\beta^{P_i}$ for some i, then $\alpha \succ^{\mathscr{P}} \beta$.*

The next axiom specifies that a common outcome of two prospects at some state of the world can be changed without impacting the comparison between those prospects, as long as the change does not affect the rankings of each prospect's outcomes.

Definition 8 (Comonotonic substitution)**.** *Let* $\alpha \in X^{\Omega}$ *and* $\omega \in \Omega$*. Denote* $x = \alpha(\omega)$*. Let* $z \in X$*. If* z *is such that there does* not *exist an outcome* $y \in \{\alpha_1, \ldots, \alpha_{K_\alpha}\}$ *with* $x < y < z$ *or* $z < y < x$*, we call* z *a* comonotonic substitute of x for α *and denote by* $\alpha^{\omega}_{x \to z}$ *the prospect which is derived from* α *by substituting* z *for* x *at* ω*, i.e. the prospect defined by*

$$\alpha^{\omega}_{x \to z}(s) = \begin{cases} z & \text{if } s = \omega \\ \alpha(s) & \text{otherwise.} \end{cases}$$

Axiom 4. Comonotonic sure-thing principle. *Take any expertise* $\mathscr{P}$*. Let* $\alpha, \beta \in X^{\Omega}$ *and* $\omega \in \Omega$ *such that* $\alpha(\omega) = \beta(\omega) = x$*. If* z *is a comonotonic substitute of* x *both for* α *and* β *then* $\alpha \succcurlyeq^{\mathscr{P}} \beta \Leftrightarrow \alpha^{\omega}_{x \to z} \succcurlyeq^{\mathscr{P}} \beta^{\omega}_{x \to z}$*.*

Our setting allows for disagreement preference, i.e. disagreement aversion or disagreement loving. We add an axiom to restrict the type of disagreement preference that the decision-rule can exhibit; namely, we require that when experts disagree over the probability of an outcome, the way disagreement is resolved does not depend on the outcome but just on the probabilities. This axiom is introduced to obtain a simpler formula for our representation result, but it is not required to derive our results on disagreement aversion (see Bommier (2017) for a representation result without this axiom in a close setting). Although one cannot characterize disagreement aversion on two-outcome prospects with common outcomes (as agreement on the certainty-equivalent would imply agreement on the probabilities), the axiom is best stated as a property on such prospects.

Definition 9 (Double substitution on binomial bets)**.** *Let* $x < y \in X$*. Let* $\alpha \in X^{\Omega}$ *such that* $\alpha_k \in \{x, y\}, \forall k \in \{1, \ldots, K_\alpha\}$*. We define* $\overline{\alpha}$ *by:*

$$\overline{\underline{\alpha}}(\omega) = \begin{cases} X^- & \textit{for } \alpha(\omega) = x \\ X^+ & \textit{for } \alpha(\omega) = y. \end{cases}$$

Axiom 5. Level-independent disagreement preference. *Take any expertise $\mathscr{P}$. For all $x < y \in X$ and all $\alpha, \beta \in X^\Omega$ such that $\alpha_k, \beta_k \in \{x, y\}, \forall k$, we have $\alpha \succcurlyeq^{\mathscr{P}} \beta \Leftrightarrow \overline{\underline{\alpha}} \succcurlyeq^{\mathscr{P}} \overline{\underline{\beta}}$.*

Our last axiom requires that the decision-rule reduces to rank-dependent expected utility (RDU) on distribution-consensual prospects. In other words, when experts all agree on the distribution of outcomes entailed by the prospect, the utility can be computed as the Choquet integral known as RDU. Equivalently, this axiom can be replaced by an alternative axiom: comonotonic mixture independence (see Chateauneuf 1999).

Axiom 6. RDU on distribution-consensual prospects. *The decision-rule $\succcurlyeq$ is RDU on distribution-consensual prospects (see Definition 1).*

B.2 Representation result

Proposition 11. *A decision-rule $\succcurlyeq$ fulfills axioms 1 to 6 if and only if it is probability-averaging. Furthermore, the representation is unique up to a positive affine transformation for the function u.*

B.3 Proof of Proposition 11

Remark 3. From Axioms 1 and 2, we can derive a natural weak order on $\mathscr{L}^N$, denoted by $\succcurlyeq$, by comparing each pair of assessments through a pair of sure prospects they are equivalent to. Formally, let $\mathbf{A}, \mathbf{B} \in \mathscr{L}^N$. Let $\alpha, \beta \in X^\Omega$ and $\mathscr{P}, \mathscr{Q}$ such that $\mathbf{F}_\alpha^{\mathscr{P}} = \mathbf{A}$ and $\mathbf{F}_\beta^{\mathscr{Q}} = \mathbf{B}$ (it is easy to verify that such $\alpha, \beta, \mathscr{P}, \mathscr{Q}$ always exist). From Axiom 1, we know that there exists $x \in X$ such that $\alpha \sim^{\mathscr{P}} x$ and from Axiom 2, this x must be unique. We can thus identify $\mathbf{A}$ with x. Similarly, we identify $\mathbf{B}$ with a certain $y \in X$ such that $\beta \sim^{\mathscr{Q}} y$. We can then set $\mathbf{A} \succcurlyeq \mathbf{B} \Leftrightarrow x \geq y$.

Let us show that $\succcurlyeq$ is continuous, in the sense that for all sequences $(\mathbf{A}_i)_{i\in\mathbb{N}}$ converging to $\mathbf{A}$, for all outcomes $y, z \in X$, $y \succcurlyeq \mathbf{A}_i \succcurlyeq z, \forall i \in \mathbb{N} \Rightarrow y \succcurlyeq \mathbf{A} \succcurlyeq z$.

Take any such sequence $(\mathbf{A}_i)$ and take converging sequences of prospects $\alpha_i \to \alpha$ and expertises $\mathscr{P}_i \to \mathscr{P}$ such that $\mathbf{F}_{\alpha_i}^{\mathscr{P}_i} = \mathbf{A}_i$ and $\mathbf{F}_{\alpha}^{\mathscr{P}} = \mathbf{A}$.[17] Let $\varepsilon > 0$. By Axiom 2, as $\mathscr{P}_i \to \mathscr{P}$ and $\alpha_i \to \alpha$, we have that for all $y,z \in X$, $y \succcurlyeq^{\mathscr{P}_i} \alpha_i \succcurlyeq^{\mathscr{P}_i} z, \forall i \in \mathbb{N} \Rightarrow y \succcurlyeq^{\mathscr{P}} \alpha \succcurlyeq^{\mathscr{P}} z$, i.e. $y \succcurlyeq \mathbf{A}_i \succcurlyeq z, \forall i \in \mathbb{N} \Rightarrow y \succcurlyeq \mathbf{A} \succcurlyeq z$.

By a similar argument, Definition 5 can be restated in terms of assessments instead of prospects, and formula (4) becomes:

$$U(\mathbf{A}) = \sum_{k=1}^{K_\alpha} \sigma_k \phi(\mathbf{A}(\alpha_k)) \tag{17}$$

where $\phi := f \circ \widetilde{I}$.

Definition 10. *We denote* $\mathscr{C}_{\uparrow}^{n} = \{(x_1,\ldots,x_n) \in X^n \,|\, x_1 \leq \ldots \leq x_n\}$ *the set of non-decreasing sequences of outcomes with* n *elements. The set of outcomes of* $\mathbf{A}$ *is defined by* $X_{\mathbf{A}} := \{x \in X \,|\, \nexists y < x, \mathbf{A}(y) = \mathbf{A}(x)\}$. *Let us also denote* $\mathscr{L}_n^N = \left\{\mathbf{A} \,|\, \exists (x_1,\ldots,x_n) \in \mathscr{C}_{\uparrow}^{n}, X_{\mathbf{A}} \subseteq \{x_k\}_{k\in\{1,\ldots,n\}}\right\}$ *the subset of assessments with no more than* n *outcomes.*

Remark 4. $\mathscr{C}_{\uparrow}^{n}$ is equivalent to $\mathscr{C}_{r}^{n}$ in Wakker (1993), $\mathscr{C}_{\uparrow}^{n}$ in Chateauneuf (1999) and $X_{\uparrow}^{n}$ in Bommier (2017).

Remark 5. Take any expertise $\mathscr{P}$. Consider a prospect α (or an assessment $\mathbf{A}$), we denote $\vec{x} = (x_1,\ldots,x_n) \in \mathscr{C}_{\uparrow}^{n}$ its outcomes organized in strictly increasing order. There exists $\mathbf{p} = (p_{i,k}) \in [0,1]^{nN}$ such that $\mathbf{F}_{\alpha}^{\mathscr{P}}(x_k) = (p_{k,1},\ldots,p_{k,N})$ (or $\mathbf{A}(x_k) = (p_{k,1},\ldots,p_{k,N})$) for $k \in \{1,\ldots,n\}$. Besides, such pair $(\vec{x},\mathbf{p})$ is unique. Conversely, each pair $(\vec{x},\mathbf{p})$ defines a unique assessment $\mathbf{A}$ defined by $\mathbf{A}(y) = (p_{k,1},\ldots,p_{k,N})$ if $x_{k-1} < y \leq x_k$ for $k \in \{2,\ldots,n\}$, $\mathbf{A}(y) = \vec{0}$ if $x_n < y$, and $\mathbf{A}(y) = \vec{1}$ if $y \leq x_1$. Consequently, we identify any pair $(\vec{x},\mathbf{p})$ with its assessment $\mathbf{A}_{\vec{x},\mathbf{p}}$ and we identify an assessment $\mathbf{A}$ with any of its corresponding pair $(\vec{x}_{\mathbf{A}},\mathbf{p}_{\mathbf{A}})$.

[17] To see that such objects exist, take $\alpha_i, \mathscr{P}_i$ such that $\mathbf{F}_{\alpha_i}^{\mathscr{P}_i} = \mathbf{A}_i$. Take a subsequence $\alpha_{\psi(i)}$ of α_i that converges, define α its limit, then take a subsequence $\mathscr{P}_{\varphi(\psi(i))}$ that converges, define $\mathscr{P}$ its limit. It remains to show that $\mathbf{F}_{\alpha}^{\mathscr{P}} = \mathbf{A}$. As $\mathbf{F}_{\alpha_i}^{\mathscr{P}_i} \to \mathbf{F}_{\alpha}^{\mathscr{P}}$ and $\mathbf{A}_i \to \mathbf{A}$, for all $\varepsilon > 0$, there exists n such that $\forall i > n, \left|\mathbf{F}_{\alpha}^{\mathscr{P}} - \mathbf{F}_{\alpha_i}^{\mathscr{P}_i}\right| + |\mathbf{A}_i - \mathbf{A}| \leq \varepsilon$. We conclude by observing that $\forall i \in \mathbb{N}, \left|\mathbf{F}_{\alpha}^{\mathscr{P}} - \mathbf{F}_{\alpha_i}^{\mathscr{P}_i}\right| + \cancel{\left|\mathbf{F}_{\alpha_i}^{\mathscr{P}_i} - \mathbf{A}_i\right|} + |\mathbf{A}_i - \mathbf{A}| \geq \left|\mathbf{F}_{\alpha}^{\mathscr{P}} - \mathbf{A}\right| \geq 0$.

By a slight abuse of notation, when $\vec{x} \in \mathscr{C}_{\uparrow}^{2}$, we often denotes $\left(\vec{x}, \begin{pmatrix} \vec{1} \\ \vec{p} \end{pmatrix}\right)$ by $(\vec{x}, \vec{p})$.

Proof. To derive the representation for assessments, let us start by fixing the number of outcomes n, and let us treat separately the cases $n = 2$ and $n \geq 3$.

1. Definition of I; n=1. Let us denote by $\vec{X} = (X^-, X^+)$ the pair of extremal outcomes. Take U, f and u as in the RDU Axiom (6). Without loss of generality, let us assume

$$u\left(X^-\right) = 0 \text{ and } u\left(X^+\right) = 1.$$

We denote by $\overline{p}$ the constant vector $(p, \ldots, p)$. We then have $U\left(\vec{X}, \overline{p}\right) = f(p)$ for any $p \in [0,1]$. Take any $x \in X$ and define $p = f^{-1}(u(x)) \in [0,1]$. By RDU on distribution-consensual prospects (Axiom 6), $u(x) = f(p)$ implies that p is the unique number such that $x \sim \left(\vec{X}, \overline{p}\right)$. Similarly, take any $\vec{p}$. By Axiom (1), there is $x \in X$ such that $\left(\vec{X}, \vec{p}\right) \sim x$. Hence, there is a unique $p \in [0,1]$ such that $\left(\vec{X}, \vec{p}\right) \sim x \sim \left(\vec{X}, \overline{p}\right)$. Let us define $\widetilde{I}(\vec{p}) = p$, and $I : \vec{p} \mapsto f \circ \widetilde{I} \circ \left(f^{-1}(p_1), \ldots, f^{-1}(p_N)\right)$. The continuity and monotonicity of $\widetilde{I}$ (and hence of I) derive from the continuity of $\succcurlyeq$[18] and from Axiom (3), respectively. Notice that[19]

$$U\left(\vec{X}, \vec{p}\right) = f\left(\widetilde{I}(\vec{p})\right) = f(p) = U\left(\vec{X}, \overline{p}\right) = u(x) = U(x). \tag{18}$$

2. The case n=2. Now, take any $(\vec{y}, \vec{p}) \in \mathscr{L}_2^N$ with $\vec{y} = (y_0, y_1)$, $y_0 < y_1$ and let

[18] Take $\varepsilon > 0$. Denote η the positive number such that $x - \eta \sim \left(\vec{X}, \overline{p - \varepsilon}\right)$. Take any sequence $(\vec{p}_k)_k$ converging to $\vec{p}$, with $\left(\vec{X}, \vec{p}_k\right) \sim \left(\vec{X}, \overline{p_k}\right)$ for all k. There exists $n \in \mathbb{N}$ such that for all $k > n$, $\left(\vec{X}, \vec{p}_k\right) \succ x - \eta$. Indeed, suppose by the absurd that there is a subsequence $\vec{p}_{\phi(i)}$ such that for all k, $x - \eta \succcurlyeq \left(\vec{X}, \vec{p}_{\phi(i)}\right)$. Then, by Axiom (2), as $\left(\vec{X}, \vec{p}_{\phi(i)}\right)$ converges towards $\left(\vec{X}, \vec{p}\right)$, $x - \eta \succcurlyeq \left(\vec{X}, \overline{p}\right)$. This contradicts $\left(\vec{X}, \overline{p}\right) \sim x \succ x - \eta$. Thus, there exists $n \in \mathbb{N}$ such that for all $k > n$, $p - \varepsilon \leq \widetilde{I}(\vec{p}_k)$. This shows that $\widetilde{I}$ is lower semi-continuous. By a similar reasoning, it can be shown that $\widetilde{I}$ is upper semi-continuous, hence $\widetilde{I}$ is continuous.

[19] This implies the following interpretation: the utility brought by a prospect equals the chance of the maximal outcome in an equivalent bet between the extremal outcomes.

$(p,y) \in [0,1] \times X$ be the unique pair such that $(\vec{y},\vec{p}) \sim y \sim (\vec{y},\overline{p})$. From RDU on distribution-consensual prospects (Axiom 6), $y \sim (\vec{y},\overline{p})$ implies $U(y) = U(\vec{y},\overline{p})$.

Now, level-independent disagreement preference (Axiom 5) implies that

$$(\vec{y},\vec{p}) \sim (\vec{y},\overline{p}) \Leftrightarrow \left(\vec{X},\vec{p}\right) \sim \left(\vec{X},\overline{p}\right) \Leftrightarrow \widetilde{I}(\vec{p}) = p \Leftrightarrow U(\vec{y},\vec{p}) \sim U(\vec{y},\overline{p}) \tag{19}$$

which in turn implies $U(y) = U(\vec{y},\vec{p})$. Finally, for any $\mathbf{A},\mathbf{B} \in \mathscr{L}_2^N$ with $\mathbf{A} \succcurlyeq \mathbf{B}$, there exist outcomes $x_A \succcurlyeq x_B$ such that $x_A \sim \mathbf{A} \succcurlyeq \mathbf{B} \sim x_B$ and by the previous result, $U(\mathbf{A}) = u(x_A) \geq u(x_B) = U(\mathbf{B})$, which achieves the proof for $n \leq 2$. Note that as U is a representation of $\succcurlyeq$ on $\mathscr{L}_2^N$, it is unique up to a strictly increasing transformation, and as $U(\cdot,\vec{p})$ is additive and continuous, this transformation must be affine.

3. The general case. Let us show by induction on $n \geq 2$ that $\mathbf{A} \mapsto U(\vec{x}_\mathbf{A},\mathbf{p_A})$ represents $\succcurlyeq$ on $\mathscr{L}_n^N$, where U is as in Definition 5 and $u, f, \widetilde{I}, \phi = f \circ \widetilde{I}$ are the functions previously defined. The initial case $n = 2$ has just been proven. Now, suppose that U represents $\succcurlyeq$ on $\mathscr{L}_n^N$ and, noting that $\mathscr{L}_n^N \subset \mathscr{L}_{n+1}^N$, let us show that this holds also true on $\mathscr{L}_{n+1}^N \backslash \mathscr{L}_n^N$. Let $\mathbf{A} \in \mathscr{L}_{n+1}^N \backslash \mathscr{L}_n^N$; $\vec{x}_\mathbf{A} = (x_1,\ldots,x_{n+1})$ is such that $x_1 < \ldots < x_{n+1}$. By continuity, there exist $\widehat{x_n}, \widehat{x} \in X$ such that $x_n \leq \widehat{x_n} \leq x_{n+1}$

$$((x_1,\ldots,x_{n-1},x_n,x_{n+1}),\mathbf{p_A}) \sim \widehat{x} \sim ((x_1,\ldots,x_{n-1},\widehat{x_n},\widehat{x_n}),\mathbf{p_A}). \tag{20}$$

Let us denote by $\mathbf{A}_{\widehat{\mathbf{n+1}}}$ the assessment of the right-hand side; notice that $\mathbf{A}_{\widehat{\mathbf{n+1}}} \in \mathscr{L}_n^N$. By induction, we have $U\left(\mathbf{A}_{\widehat{\mathbf{n+1}}}\right) = U(\widehat{x})$. By the comonotonic sure-thing principle (Axiom 4), we can replace x_{n-1} by x_n in equation (20) and get:

$$((x_1,\ldots,x_n,x_n,x_{n+1}),\mathbf{p_A}) \sim ((x_1,\ldots,x_n,\widehat{x_n},\widehat{x_n}),\mathbf{p_A}). \tag{21}$$

The two sides of equation (21) are elements of $\mathscr{L}_n^N$, denoted respectively $\mathbf{A_n}$ and $\mathbf{A}_{\widehat{\mathbf{n}}}$, and whose associated probability matrices differ only by their last two rows. By induction, we have: $U(\mathbf{A_n}) = U(\mathbf{A}_{\widehat{\mathbf{n}}})$, which yields after a cancellation of identical terms in the sum (equation 17):

$$
\begin{aligned}
u(x_n)\left(\phi\left(\mathbf{A_n}(x_n)\right)-\phi\left(\mathbf{A_n}(x_{n+1})\right)\right)+u(x_{n+1})\,\phi\left(\mathbf{A_n}(x_{n+1})\right)= \\
u(x_n)\left(\phi\left(\mathbf{A}_{\widehat{\mathbf{n}}}(x_n)\right)-\phi\left(\mathbf{A}_{\widehat{\mathbf{n}}}(\widehat{x_n})\right)\right)+u(\widehat{x_n})\,\phi\left(\mathbf{A}_{\widehat{\mathbf{n}}}(\widehat{x_n})\right).
\end{aligned}
\tag{22}
$$

Re-arranging the terms in (22) and expressing the probability vectors as rows of $\mathbf{p_A}$ (which are of the form $\mathbf{A}(x_k)$):

$$
\begin{aligned}
u(\widehat{x_n})\,\phi\left(\mathbf{A}_{\widehat{\mathbf{n}}}(\widehat{x_n})\right) &= u(x_n)\left(\phi\left(\mathbf{A_n}(x_n)\right)-\phi\left(\mathbf{A_n}(x_{n+1})\right)\right)-u(x_n)\left(\phi\left(\mathbf{A}_{\widehat{\mathbf{n}}}(x_n)\right)\right. \\
&\quad \left.-\phi\left(\mathbf{A}_{\widehat{\mathbf{n}}}(\widehat{x_n})\right)\right)+u(x_{n+1})\,\phi\left(\mathbf{A_n}(x_{n+1})\right) \\
&= u(x_n)\left(\phi\left(\mathbf{A_n}(x_n)\right)-\phi\left(\mathbf{A_n}(x_{n+1})\right)+\phi\left(\mathbf{A}_{\widehat{\mathbf{n}}}(\widehat{x_n})\right)-\phi\left(\mathbf{A}_{\widehat{\mathbf{n}}}(x_n)\right)\right) \\
&\quad +u(x_{n+1})\,\phi\left(\mathbf{A_n}(x_{n+1})\right) \\
&= u(x_n)\left(\cancel{\phi\left(\mathbf{A}(x_{n-1})\right)}-\phi\left(\mathbf{A}(x_{n+1})\right)+\phi\left(\mathbf{A}(x_n)\right)\cancel{-\phi\left(\mathbf{A}(x_{n-1})\right)}\right) \\
&\quad +u(x_{n+1})\,\phi\left(\mathbf{A}(x_{n+1})\right)
\end{aligned}
\tag{23}
$$

Denoting $S=\sum_{k=1}^{n-2}u(x_k)\left(\phi\left(\mathbf{A}(x_k)\right)-\phi\left(\mathbf{A}(x_{k+1})\right)\right)$ the first terms in the summation $U(\mathbf{A})$, we have:

$$
\begin{aligned}
U(\mathbf{A})-S &= u(x_{n-1})\left(\phi\left(\mathbf{A}(x_{n-1})\right)-\phi\left(\mathbf{A}(x_n)\right)\right)+u(x_n)\left(\phi\left(\mathbf{A}(x_n)\right)-\phi\left(\mathbf{A}(x_{n+1})\right)\right) \\
&\quad +u(x_{n+1})\,\phi\left(\mathbf{A}(x_{n+1})\right) \\
&\overset{\text{eq. (23)}}{=} u(x_{n-1})\left(\phi\left(\mathbf{A}(x_{n-1})\right)-\phi\left(\mathbf{A}(x_n)\right)\right)+u(\widehat{x_n})\,\phi\left(\mathbf{A}_{\widehat{\mathbf{n}}}(\widehat{x_n})\right) \\
U(\mathbf{A})-S &= u(x_{n-1})\left(\phi\left(\mathbf{A}_{\widehat{\mathbf{n+1}}}(x_{n-1})\right)-\phi\left(\mathbf{A}_{\widehat{\mathbf{n+1}}}(\widehat{x_n})\right)\right)+u(\widehat{x_n})\,\phi\left(\mathbf{A}_{\widehat{\mathbf{n+1}}}(\widehat{x_n})\right)
\end{aligned}
\tag{24}
$$

where we expressed the probability vectors using $\mathbf{A}_{\widehat{\mathbf{n+1}}}$ to get the last line.

Observing that $U(\mathbf{A})-S$ equals the last two terms in the summation $U\left(\mathbf{A}_{\widehat{\mathbf{n+1}}}\right)$ and noting that $U\left(\mathbf{A}_{\widehat{\mathbf{n+1}}}\right)$ already coincides with $U(\mathbf{A})$ on the first terms S, the first last equation implies $U(\mathbf{A})=U\left(\mathbf{A}_{\widehat{\mathbf{n+1}}}\right)$. Thus, we have $U(\mathbf{A})=U(\widehat{x})$, which shows that U represents $\succcurlyeq$ on $\mathscr{L}_{n+1}^{N}$.

The reciprocal (that any probability-averaging decision-rule satisfies axiom 1 to 6) is easy, so we will just provide the sketch of the proof. Take any prospect α, and expertise $\mathscr{P}$, define $x = u^{-1}\left(U^{\mathscr{P}}(\alpha)\right)$; it is immediate to see that $\alpha \sim^{\mathscr{P}} x$ and that x depends only on $\mathbf{F}_{\alpha}^{\mathscr{P}}$, proving Axiom 1. Axioms 2 and 3 ensues directly from the continuity and monotonicity properties of the functions composing U. Axiom 4 derives from the monotonicity of u, f and I, while Axiom 6 is immediate when one considers the restriction of $\succcurlyeq^{\mathscr{P}}$ to $\mathscr{L}^{\mathscr{P}}$. Level-independent disagreement preference stems from the fact that for any prospect α whose representation is of the form $((x,y),\vec{p})$ with $x < y \in X$, $U^{\mathscr{P}}(\alpha) = (u(y) - u(x))\,\phi(\vec{p})$. ■