Objective Imprecise Probabilistic Information, Second Order Beliefs and Ambiguity Aversion: an Axiomatization

Raphaël Giraud*

CRESE - University of Franche-Comté

raphael.giraud@univ-fcomte.fr

First version: July 2005
This version: July 2006

Abstract

We axiomatize a model of decision under objective ambiguity or imprecise risk. The decision maker forms a subjective (non necessarily additive) belief about the likelihood of probability distributions and computes the average expected utility of a given act with respect to this second order belief. We show that ambiguity aversion like the one revealed by the Ellsberg paradox requires that second order beliefs be nonadditive. Some special cases of the model are examined and different forms of ambiguity aversion are characterized.

Keywords: Imprecise probabilistic information, second order beliefs, non-additive probabilities, ambiguity aversion, Ellsberg paradox, Choquet integral.

Journal of Economic Literature: D81.

1 Introduction

1.1 Motivation

Ambiguity and ambiguity aversion have become the center of attention in the last twenty years or so in decision theory. This interest has grown from the challenge leveled by the Ellsberg paradox (Ellsberg 1961) against Savage's Subjective Expected Utility Model (Savage 1954). The main feature of the Ellsberg paradox is ambiguity aversion: people tend to choose less ambiguous bets over more ambiguous ones, even when under certain conditions the ambiguous ones may be more favorable than the unambiguous ones.

*I thank audiences of the Centre d'Economie de la Sorbonne (University Paris I), and of the ISIPTA 05, FUR XII and RUD 2006 conferences.

Now, ambiguity aversion may be interpreted as a form of pessimism: when favorable and unfavorable scenarios (probability distributions on the states of the world) are compatible with the information available to the decision maker, an ambiguity averse one will always deem unfavorable scenarios more likely than favorable ones. However, interpreting ambiguity aversion in this way requires the existence of second order beliefs, i.e. beliefs over probability distributions. In this paper, we wish to investigate the conditions under which such second order beliefs exist and how they interact with information.

1.2 Goal of the paper

The goal of this paper is to axiomatize a decision model under imprecisely known probabilistic information using second order beliefs in a way compatible with the Ellsberg paradox. The approach we take to that effect is similar to the approaches of Wang (2001, revised 2003), Gajdos, Tallon, and Vergnaud (2004), Gajdos, Hayashi, Tallon, and Vergnaud (2006), Nehring (2002) and Nehring (2001): we assume that decision makers have preferences on pairs $(f, \mathscr{P})$, where f is an act à la Savage from the set S of states of nature to the set X of outcomes, and $\mathscr{P}$ is a set of probability distributions on S. Consider for instance two Ellsberg urns: the first contains 90 red, blue and yellow balls, among which exactly 30 are red. The other contains 90 red, blue and yellow balls, among which exactly 30 are red and at least 10 are blue. Would you rather bet on red in the first urn or bet on blue in the second urn? This is the kind of question we assume the decision maker is able to answer. A more mundane example would be the following: when deciding to design a policy, a political decision maker can ask experts. If the policy maker is used to work with a given expert, his or her confidence in the estimates of the expert will be high, so that he or she will not consider the information provided by the expert as ambiguous. If, on the other hand, he or she is not very much acquainted with the expert, the information he or she will deliver will be deemed ambiguous. Now, if the policy maker has to choose between a policy recommended by the known expert and a policy recommended by the less known expert, which one should he or she choose?

If second order beliefs are additive and if reduction of compound lotteries is allowed, then after the reduction is performed, the decision maker holds additive first order belief; this is incompatible with the Ellsberg paradox. Therefore, there are two possible ways of dealing with second order beliefs in a way compatible with ambiguity aversion:

- Abandon the possibility of reducing of compound lotteries. This is the route followed by Klibanoff, Marinacci, and Mukerji (2005), Nau (2006) and Seo (2006) (see also Rustichini (1992)), as clearly explained in the latter. The criterion axiomatized in these papers is the following, called Smooth Ambiguity Model by Klibanoff, Marinacci, and Mukerji and Second Order Subjective Expected Utility (SOSEU) by Seo:

$$SOSEU(f, \mathscr{P}) = \int_{\mathscr{P}} \varphi \left(\int_S u \circ f \, \mathrm{d}P \right) \, \mathrm{d}\pi^{\mathscr{P}}(P),$$

  where $\pi^{\mathscr{P}}$ is a probability measure.

- Relax additivity of second order beliefs. This is the route we shall follow. Indeed, the model we propose is the following: the decision maker forms a subjective prior on the multiple scenarios that are compatible with the available probabilistic information, $\mathscr{P}$. This prior is a capacity $\nu^{\mathscr{P}}$, i.e. a monotonic and normalized (non-necessarily additive) set function. Then, he or she computes the average expected utility of a given act f, by the formula:

$$V(f, \mathscr{P}) = \int_{\mathscr{P}} \left(\int_S u \circ f \, \mathrm{d}P \right) \, \mathrm{d}\nu^{\mathscr{P}}(P),$$

  where the first integral is a Choquet integral w.r.t. the capacity $\nu^{\mathscr{P}}$.

Just as Yaari (1987)'s dual theory of choice under risk is dual to expected utility in the sense that it transforms probabilities rather than outcomes, our model is dual to SOSEU in the sense that our φ is linear and our $\pi^{\mathscr{P}}$ is non-additive. For this reason we propose to call this model Second Order Dual Expected Utility (SODEU).

In choosing this alternative route to axiomatizing the natural idea that when a decision maker faces an imprecise probabilistic information, he or she forms some prior on the scenarios compatible with this information, we were inspired by a remark to be found in Chateauneuf, Cohen, and Jaffray (2006)[1]. These authors show how the Ellsberg paradox for the three-color urn, usually spelled out in terms of Savage acts, can be recast in the Anscombe-Aumann framework using the composition of the urn as state space so that it becomes clear that it does not only violate Savage's Sure Thing Principle, but also (Anscombe and Aumann 1963)'s Independence Axiom. The principle of the proof of this point is simple: consider the set $S := \{R, B, Y\}$ of states corresponding to the color of the ball and the set of outcomes $X = \{0, 100\}$. For a given number k of black balls in the urn, consider P_k the distribution on S

[1]It seems however that the original idea was Schmeidler's (personal communication by J.-Y. Jaffray). In a sense it is a kind of folk result, but in this book chapter it is formally proved.

given by $P_k = (\frac{1}{3}, \frac{k}{90}, \frac{60-k}{90})$. Consider as well the set $\mathscr{P} = \{P_k \mid k = 1, \ldots, 60\}$. Now to each $f : S \to X$, associate the second-order act $F^f : \mathscr{P} \to \Delta(X)$, where $\Delta(X)$ is the set of lotteries on X, by letting $F^f(P_k) = P_k^f$, where $P_k^f(x) = P_k(f^{-1}(x))$. If E is an event in S, then let $f_E(s) = 100$ if $s \in E$ and $f_E(s) = 0$ otherwise. Table I gives the correspondence between the acts involved in the Ellsberg paradox and the second order acts that are associated to them.

act	**consequence on P_k**
F^R ($\approx f_R$)	$(0, \frac{60}{90}\ ;\ 100, \frac{30}{90})$
F^B ($\approx f_B$)	$(0, \frac{90-k}{90}\ ;\ 100, \frac{k}{90})$
$F^{\{R,Y\}}$ ($\approx f_{\{R,Y\}}$)	$(0, \frac{k}{90}\ ;\ 100, \frac{90-k}{90})$
$F^{\{B,Y\}}$ ($\approx f_{\{B,Y\}}$)	$(0, \frac{30}{90}\ ;\ 100, \frac{60}{90})$

Table I: The Ellsberg in the Anscombe-Aumann setup

Consider now the act f_Y. Its corresponding second-order act is defined by $F^Y(P_k) = (0, \frac{30+k}{90}\ ;\ 100, \frac{60-k}{90})$ and we have

$$\tfrac{1}{2}F^R + \tfrac{1}{2}F^Y = \tfrac{1}{2}F^{\{R,Y\}} + \tfrac{1}{2}\delta_0$$

on the one hand and

$$\tfrac{1}{2}F^B + \tfrac{1}{2}F^Y = \tfrac{1}{2}F^{\{B,Y\}} + \tfrac{1}{2}\delta_0$$

on the other, where δ_0 is the constant second-order act taking the Dirac at 0 as its value.

Assuming that $F^f \succsim F^g$ if and only if $f \succsim g$, then modal preferences contradict the independence axiom because of the previous relations.

This suggests that, in order to account for the Ellsberg paradox, the prior on the scenarios cannot be additive. This should pave the way to the characterization of various behaviors under uncertainty thanks to the vast literature on the Choquet integral now available. We provide some examples in the sequel.

1.3 Related literature

We already discussed the relationship of our work with the literature on second-order beliefs, so we concentrate here on other aspects of the related literature.

The idea of generalizing the notion of risk, i.e. known probability distribution, to some notion of imprecise risk, i.e. imprecisely known probability distribution, was introduced by (Jaffray 1989), who proposed to replace lotteries by belief functions, that roughly speaking correspond to sets of lotteries. He obtains a representation in the spirit of the Arrow-Hurwicz criterion (Arrow and Hurwicz 1972). This idea was recently revived by several authors who proposed to model imprecise risk directly by sets of lotteries (Olszewski 2002, revised 2006, Ahn 2003, revised 2005, Stinchcombe 2003): Olszewski also axiomatizes an α-maxmin rule in the spirit of the Arrow-Hurwicz criterion, while Ahn proposes what can be interpreted in the framework of sets of lotteries as a conditional version of the SOSEU model, although Ahn's theorem works only for sets of lotteries having very specific topological properties; Stinchcombe studies various properties of continuous and independent preferences on the set of lotteries.

As we will show in the sequel, our approach differs from the former approaches in that we consider pairs $(f, \mathscr{P})$ where f is an act from some state space S to some outcome space X and $\mathscr{P}$ is a set of probability distributions on S, but we do not assume that the decision maker is indifferent between two acts that induce the same set of lotteries on X together with $\mathscr{P}$. The idea is that the source of the set of lotteries matters here. The approach we take originates in the paper by Wang (2001, revised 2003) that takes as primitives triples $(f, \mathscr{P}, P^*)$ where P^* is a reference prior. Wang main result is to provide axiomatic foundations for a general version of the minimum relative entropy principle of Anderson, Hansen, and Sargent (1999), but he also characterizes a maxmin rule whereby the decision maker maximizes the minimum expected utility with respect to the set of objective priors. This rule is characterized by an axiom of aversion to uncertainty, that is more an axiom of strong aversion to objective ambiguity, saying that the decision maker always prefers situations with a more precise objective information (i.e. a smaller set of priors). In the same setting, Gajdos, Tallon, and Vergnaud (2004) weaken the notion of aversion to objective ambiguity, what they call aversion to imprecision, and obtain a generalized maxmin rule whereby the decision maximizes the minimum expected utility over a *subset* of the set of initial priors. One special case of this model is a form of perturbed expected utility, given by the formula

$$V(f, \mathscr{P}, P^*) = (1-\alpha)\int u \circ f \mathrm{d}P^* + \alpha \min_{P \in \mathscr{P}} \int u \circ f \mathrm{d}P,$$

a model that was first proposed in Ellsberg (1961) in order to rationalize the behavior described in this paper, and that also appears in Tapking (2004) in the context of the study of updating rules for capacities. Gajdos, Hayashi, Tallon, and Vergnaud (2006) generalize the work done in Gajdos, Tallon, and Vergnaud (2004) by dispensing with the reference prior, now part of the representation, and providing foundations for the interpretation of the parameter α as revealing aversion to imprecision. Kopylov (2006) provides alternative foundations for the perturbed expected utility formula, characterizing it by an axiom that can be interpreted as a combination of ambiguity aversion and loss aversion.

? studies a related framework where along with the usual preference relation the decision maker is endowed with a (potentially incomplete) comparative likelihood relation, that is assumed to represent the decision maker's beliefs. The object of investigation is the compatibility of betting preferences as derived from the preference relation with the decision maker's beliefs represented by the comparative likelihood relation. Aspects of this compatibility will also be studied in this paper.

Finally, Amarante (2006) also studies the functional form that we call here SODEU. More specifically, he shows that in the Anscombe-Aumann setup, any invariant biseparable preference relation can be represented by a Choquet integral over a set of priors.

1.4 Organization of the Paper

The paper is organized as follows: section 2 introduces the set up and the axioms, in section 3 the main theorem is stated and interpreted. Section 4 presents important special cases of the functional. Section 5 studies ambiguity aversion in this setup. Section 6 contains concluding remarks. Proofs are to be found in the appendix.

2 The Model

2.1 Set Up and Basic Definitions

The set of states of nature is here denoted S and endowed with a σ-algebra Σ. The set of outcomes is a measurable space $(X, \mathscr{B})$ where $\mathscr{B}$ contains the singletons. We denote by $\mathscr{F}$ the set of simple Savage acts, i.e. the set of finite-valued measurable functions f from S to X. Let $pc(\Sigma)$ be the set of all probability charges (finitely additive and normalized set functions) on Σ. Following Nehring (2002), we shall say that a set $\mathscr{P} \subseteq pc(\Sigma)$ is *convex-ranged* if for all $\alpha \in (0, 1)$, for all $A \in \Sigma$, there exists $B \in \Sigma$, $B \subseteq A$ such that $P(B) = \alpha P(A)$ for all $P \in \mathscr{P}$. The Lyapunov convexity theorem implies that if $\mathscr{P}$ is the convex hull of a finite number of countably additive non-atomic probability measures, then $\mathscr{P}$ is convex-ranged. Nehring (2002) provides other examples. We denote by $\mathfrak{P}$ the set of all non-empty convex-ranged subsets $\mathscr{P}$ of $pc(\Sigma)$.

Following **?**, the objects of choice in our setting will be pairs $(f, \mathscr{P})$ in $\mathscr{F} \times \mathfrak{P}$. A pair $(f, \mathscr{P})$ of this sort corresponds to a situation where the objectively given information relevant to act f is consistent with an imprecise probabilistic representation given by $\mathscr{P}$. We assume here that preferences are expressed over the set $\mathscr{F} \times \mathfrak{P}$ and are represented by the relation $\succsim$.

2.2 Axioms

We assume the following standard axiom:

Axiom 1 (Weak Order) *$\succsim$ is transitive and complete.*

Comparisons between two acts accompanied by different imprecise information $(f, \mathscr{P})$ and $(g, \mathscr{P}')$ may seem awkward, but it is in fact very natural: for instance, when a businessperson is about to sign a contract in a country, the probability of the contract being enforced depends in particular on the legal system. The less information the decision maker has about the country, the more imprecise his or her estimate of this probability based on this information. When one has to choose between investments in different countries, one has therefore to compare similar decisions in precisely and imprecisely known legal contexts, i.e. under different ambiguous pieces of information.

For $f, g \in \mathscr{F}$ and $A \in \Sigma$, the *A-graft of f with g*, denoted by fAg, is the act such that $fAg(s) = f(s)$ if $s \in A$ and $fAg(s) = g(s)$ if $s \notin A$.

If $f \in \mathscr{F}$, denote Σ_f the algebra generated by f. If $F \subseteq \mathscr{F}$, denote Σ_F the algebra generated by $\cup_{f \in F} \Sigma_f$. Let $\mathscr{P} \in \mathfrak{P}$. Say that $A \in \Sigma$ is *independent of F w.r.t. $\mathscr{P}$*, denoted $A \perp_{\mathscr{P}} F$, if the algebra generated by A is independent of Σ_F for each $P \in \mathscr{P}$.

We can now state the continuity axiom we shall use, that roughly says that given information $\mathscr{P}$, there are always events of sufficiently small measure w.r.t. to each prior in $\mathscr{P}$ for preferences not to be affected by a modification of acts on one of such events.

Axiom 2 (Information-Contingent Continuity) *For all $\mathscr{P} \in \mathfrak{P}$, for all $f, g, h \in \mathscr{F}$, if $(f, \mathscr{P}) \succ (g, \mathscr{P}) \succ (h, \mathscr{P})$ then there exists $A, B \in \Sigma$, $\alpha, \beta \in (0, 1)$ such that:*

(i) $A \perp_{\mathscr{P}} \{f, h\}$, $B \perp_{\mathscr{P}} \{f, h\}$;

(ii) $P(A) = \alpha$, $P(B) = \beta$, *for all* $P \in \mathscr{P}$*;*

(iii) $(fAh, \mathscr{P}) \succ (g, \mathscr{P}) \succ (fBh, \mathscr{P})$,

In order to state the next axioms, we shall need the following definition. For any $f \in \mathscr{F}$ and $P \in \mathscr{P}(S)$, we let P^f denote the probability measure induced by f on X, i.e. for all $B \in \mathscr{B}$,

$$P^f(B) = P(f^{-1}(B)).$$

As f is finite-valued, P^f has finite support. We let $\Delta(X)$ be the set of all finitely-supported probability measures or *lotteries* on X. Then $P^f \in \Delta(X)$.

For any set $\mathscr{P} \in \mathfrak{P}$, any $\pi \in \Delta(X)$, let

$$K(\pi, \mathscr{P}) := \{k \in \mathscr{F} \mid P^k = \pi, \forall P \in \mathscr{P}\}.$$

For $\mathscr{P} \in \mathfrak{P}$, let $\succsim_{\mathscr{P}}$ be the binary relation defined for all $\pi, \pi' \in \Delta(X)$ by:

$$\pi \succsim_{\mathscr{P}} \pi' \quad \Longleftrightarrow \quad \exists k \in K(\pi, \mathscr{P}), \exists k' \in K(\pi', \mathscr{P}), (k, \mathscr{P}) \succsim (k', \mathscr{P}).$$

In order for this relation to be well-defined, we must impose the following axiom:

Axiom 3 (No Framing Effect) *For all* $\pi \in \Delta(X)$*, for all* $\mathscr{P} \in \mathfrak{P}$*, for all* $k, k' \in K(\pi, \mathscr{P})$*,* $(k, \mathscr{P}) \sim (k', \mathscr{P})$*.*

Intuitively, acts in $K(\pi, \mathscr{P})$ differ only by the permutation of outcomes on events of equal probability for all priors in $\mathscr{P}$, a manipulation that amounts to relabeling these events. From a (strict) normative point of view, such a relabeling should not affect the decision maker's preference, because this would amount to a framing effect. This is the intuition that motivates axiom 3.

In the usual setting of decision under risk, objects of choice are assumed to be lotteries. In fact, this setting corresponds to a situation where the decision maker has precise information (i.e. $\mathscr{P}$ is a singleton), and axiom 3 holds. When information is imprecise, one approach taken in the literature is to define preferences directly on sets of lotteries (Olszewski 2002, revised 2006, Ahn 2003, revised 2005, Stinchcombe 2003). This approach requires a strengthening of axiom 3 to appear as a special case of the approach followed here, namely an axiom requiring that acts inducing the same *set* of lotteries from a set of priors $\mathscr{P}$ be indifferent, while axiom 3 requires this only if the induced set of lotteries is a singleton.

The introduction of the relation $\succsim_{\mathscr{P}}$ allows us to state the following definition:

Definition 1 *Let* $f, g \in \mathscr{F}$ *and* $\mathscr{P} \in \mathfrak{P}$*. We shall say that* f *and* g *are* $\mathscr{P}$*-*

comonotonic if, for all $P, Q \in \mathscr{P}$,

$$P^f \succ_{\mathscr{P}} Q^f \quad \Longrightarrow \quad P^g \succsim_{\mathscr{P}} Q^g.$$

The intuition behind this definition is the following. Given an information set $\mathscr{P}$ and an act f, one can associate to each $P \in \mathscr{P}$ a lottery P^f. Moreover, one can associate to f an ordering $\succsim^f$ on $\mathscr{P}$ defined by:

$$P \succsim^f Q \quad \Longleftrightarrow \quad P^f \succsim_{\mathscr{P}} Q^f.$$

This ordering answers the following question: "If I were given the choice between an act that, conditional on my information, unambiguously induces the lottery P^f, and one that unambiguously induces the lottery Q^f, which one would I choose?" If I choose P^f, this means that, given the choice of act f, I deem the scenario corresponding to P to be more favorable than the scenario corresponding to Q. For instance, in the three color Ellsberg urn with 30 red ball and 60 black or yellow balls, given the act f_B corresponding to betting on black, the decision maker would deem the scenario corresponding to the following proportions of red, black and yellow balls: $(\frac{1}{3}, \frac{2}{3}, 0)$ more favorable than the scenario $(\frac{1}{3}, 0, \frac{2}{3})$, so that:

$$\left(\frac{1}{3}, \frac{2}{3}, 0\right) \succ^{f_B} \left(\frac{1}{3}, 0, \frac{2}{3}\right).$$

This ordering therefore ranks scenarios according to how relatively favorable they are given act f. Now, two acts are $\mathscr{P}$-comonotonic as defined if, roughly speaking, they order the scenarios in the same way regarding how favorable they are. If they order scenarios in the same way, they do not provide any hedging opportunity against each other, not in the usual sense of compensating bad states of nature for one act with good states for the other, but of compensating bad scenarios with good scenarios.

Comonotonicity is a notion that has been introduced in the literature on decision under uncertainty for states of nature, not for probabilistic scenarios. Recall that, given an information set $\mathscr{P}$, two acts f and g are comonotonic if, for all $s, s' \in S$,

$$(f(s), \mathscr{P}) \succ (f(s'), \mathscr{P}) \Longrightarrow (g(s), \mathscr{P}) \succsim (g(s'), \mathscr{P}).$$

One may wonder whether there is some relationship between the two notions. The answer is no. To see that, we shall give two examples.

Example 1 (Comonotonicity does not imply $\mathscr{P}$-comonotonicity) *Let* $S = X =$

$[0,1]$, $\mathscr{P} = \{P, Q\}$ *where* P *is the Lebesgue measure and* Q *has the following density with respect to* P*:*

$$q(s) = \begin{cases} 4s & \text{if } s \in [0, \frac{1}{2}], \\ 4(1-s) & \text{if } s \in [\frac{1}{2}, 1]. \end{cases}$$

Consider the following preference relation over act-information pairs:

$$(f, \mathscr{P}) \succsim (g, \mathscr{P}) \Longleftrightarrow \int f \, \mathrm{d}(\frac{1}{2}P + \frac{1}{2}Q) \geq \int g \, \mathrm{d}(\frac{1}{2}P + \frac{1}{2}Q),$$

and the following acts:

$$f(s) = \begin{cases} 0 & \text{if } s \in [0, \frac{1}{4}], \\ 1 & \text{if } s \in (\frac{1}{4}, 1], \end{cases} \quad \text{and} \quad g(s) = \begin{cases} 0 & \text{if } s \in [0, \frac{3}{4}], \\ 1 & \text{if } s \in (\frac{3}{4}, 1]. \end{cases}$$

Then f *and* g *are comonotonic (they are both non-decreasing). Now*

$$P^f(1) = \frac{3}{4} \quad \text{and} \quad Q^f(1) = \frac{7}{8},$$

so that $Q \succ^f P$*, and*

$$P^g(1) = \frac{1}{4} \quad \text{and} \quad Q^g(1) = \frac{1}{8},$$

so that $P \succ^g Q$*:* f *and* g *are not* $\mathscr{P}$*-comonotonic.*

Example 2 ($\mathscr{P}$-comonotonicity does not imply comonotonicity) *In the same setting as in the previous example, consider the following acts:*

$$f(s) = \begin{cases} 0 & \text{if } s \in [0, \frac{1}{2}], \\ 1 & \text{if } s \in (\frac{1}{2}, 1], \end{cases} \quad \text{and} \quad g(s) = \begin{cases} 1 & \text{if } s \in [0, \frac{1}{2}], \\ 0 & \text{if } s \in (\frac{1}{2}, 1]. \end{cases}$$

Then f *and* g *are not comonotonic, but they are* $\mathscr{P}$*-comonotonic, as* P *and* Q *have symmetric densities around* $\frac{1}{2}$*.*

Now, if two acts f and g are $\mathscr{P}$-comonotonic, i.e. do not provide any hedge against each other with respect to the potential scenarios, and if f is preferred to g given information $\mathscr{P}$, this means, roughly speaking, that "on average" f performs better than g with respect to information $\mathscr{P}$, for instance if it is definitely better with respect to good scenarios, though it might not dominate g with respect to bad scenarios. If h is $\mathscr{P}$-comonotonic with both f and g, mixing it – in the sense of grafting – with both of them will result in two acts that bear the same relation as f and g with respect to potential scenarios. Therefore, their preference ranking should be the same as that of the original acts, provided the probabilities on the algebra generated by f, g and h given by the scenarios in $\mathscr{P}$ are not affected by conditioning on the event used to perform the grafting operation, that is, provided the event is independent of f, g and h w.r.t. to all the scenarios. This normatively appealing behavior is what the next axiom requires:

Axiom 4 (Information-Contingent Comonotonic Independence) *For all $\mathscr{P} \in \mathfrak{P}$, for all $f, g, h \in \mathscr{F}$ pairwise $\mathscr{P}$-comonotonic, for all $A \in \Sigma$ such that $A \perp_{\mathscr{P}} \{f, g, h\}$,*

$$(f, \mathscr{P}) \succsim (g, \mathscr{P}) \Longleftrightarrow (fAh, \mathscr{P}) \succsim (gAh, \mathscr{P}).$$

Based on the same interpretive line, it seems normatively compelling that act f be preferred to act g given information $\mathscr{P}$ whenever in each scenario of $\mathscr{P}$, act f induces a more desirable lottery than act g:

Axiom 5 (Information-Contingent Dominance) *For all $f, g \in \mathscr{F}$, for all $\mathscr{P} \in \mathfrak{P}$,*

$$(\forall P \in \mathscr{P},\ P^f \succsim_{\mathscr{P}} P^g) \quad \Longrightarrow \quad (f, \mathscr{P}) \succsim (g, \mathscr{P}).$$

The next axiom requires only that the problem be non-trivial.

Axiom 6 (Non-Degeneracy) *For all $\mathscr{P} \in \mathfrak{P}$, there exist $f, g \in \mathscr{F}$ such that $(f, \mathscr{P}) \succ (g, \mathscr{P})$.*

Constant acts in $\mathscr{F}$ correspond to actions that are not state-contingent. Uncertainty is therefore irrelevant to them, and so is, of course, information about this uncertainty. This is the meaning of the next axiom. In a sense, this axiom also implies that the objectively given information does not affect the decision-maker's confidence in the accuracy of the description of uncertainty by the list of states in S.

Axiom 7 (Preferences under Certainty) *For all $\mathscr{P}, \mathscr{P}' \in \mathfrak{P}$, for all $x \in X$, $(x, \mathscr{P}) \sim (x, \mathscr{P}')$.*

The last two axioms are essentially technical and require richness in the setting. Axiom 8 would automatically hold whenever X is a connected topological space and for each $\mathscr{P}$ the preference over $\mathscr{F}$ given $\mathscr{P}$ is continuous:

Axiom 8 (Certainty Equivalent) *For all $\mathscr{P} \in \mathfrak{P}$, for all $f \in \mathscr{F}$, there exists $x \in X$ such that $(x, \mathscr{P}) \sim (f, \mathscr{P})$.*

The next axiom says that any rule that transforms each scenario compatible with objective information into a specific lottery on the space of outcomes can be implemented in a subjectively equivalent way by choosing an act in the set of available acts.

Axiom 9 (Denseness of the Set of Acts) *For all $\mathscr{P} \in \mathfrak{P}$, for all function*

$$F : \mathscr{P} \to \Delta(X),$$

if there exist $\bar{\pi}$, $\underline{\pi} \in \Delta(X)$ such that

$$\bar{\pi} \succsim_{\mathscr{P}} F(P) \succsim_{\mathscr{P}} \underline{\pi}, \quad \forall P \in \mathscr{P},$$

then there exists $f \in \mathscr{F}$ such that, for all $P \in \mathscr{P}$,

$$P^f \sim_{\mathscr{P}} F(P).$$

It is easy to show that the previous axioms imply that, for each $P \in \mathscr{P}$, there exists $f^P \in \mathscr{F}$ such that

$$P^{f^P} \sim_{\mathscr{P}} F(P).$$

The effect of this axiom is therefore to strengthen this result to get a uniform f.

3 The Representation Theorem

3.1 Statement

In order to introduce the representation theorem, we recall the following definition:

Definition 2 *Let $(\Omega, \mathscr{A})$ be a measurable space. A* capacity *on $(\Omega, \mathscr{A})$ is a function $\nu : \mathscr{A} \to \mathbb{R}$ such that:*

(i) $\nu(\varnothing) = 0$ *and* $\nu(\Omega) = 1$;

(ii) For all $A, B \in \mathscr{A}$, $A \subseteq B \quad \Longrightarrow \quad \nu(A) \leq \nu(B)$.

Let $\varphi : \Omega \to \mathbb{R}$. We say that φ is $\mathscr{A}$-measurable if, for all $t \in \mathbb{R}$,

$$(\varphi > t) := \{\omega \in \Omega \mid \varphi(\omega) \geq t\} \in \mathscr{A}.$$

Then, for all $\mathscr{A}$-measurable functions φ, the Choquet integral *of φ with respect to ν is defined by:*

$$\int_\Omega \varphi \, d\nu := \int_{-\infty}^{0} \nu(\varphi \geq t) - 1 \, \mathrm{d}t + \int_0^{+\infty} \nu(\varphi \geq t) \, \mathrm{d}t.$$

If φ is finite-valued, there exist families $(A_i)_{i=1,\dots,n}$, $A_i \in \mathscr{A}$ and $(x_i)_{i=1,\dots,n}$ with $x_i \in \mathbb{R}$ such that $x_1 \leq x_2 \leq \dots \leq x_n$ and φ takes the value x_i on the set A_i. Then, the Choquet integral of φ with respect to ν is given by:

$$\int_{\Omega} \varphi \; d\nu := \sum_{i=1}^{n} x_i[\nu(\cup_{j=i}^{n} A_j) - \nu(\cup_{j=i+1}^{n} A_j)]. \tag{1}$$

We can now state the main representation theorem:

Theorem 1
If $\succsim$ satisfies axioms 1 through 9, then, there exist a non-constant function $u : X \to \mathbb{R}$ and, for all $\mathscr{P} \in \mathfrak{P}$, a capacity $\nu^{\mathscr{P}}$ on $\mathscr{P}$ such that:

$$(f, \mathscr{P}) \succsim (g, \mathscr{P}') \quad \Longleftrightarrow \quad V(f, \mathscr{P}) \geq V(g, \mathscr{P}'),$$

where

$$V(f, \mathscr{P}) = \int_{\mathscr{P}} \left(\int_S u \circ f \; dP \right) \mathrm{d}\nu^{\mathscr{P}}(P) \tag{2}$$

Moreover, u is defined up to an affine increasing transformation and for each $\mathscr{P}$, $\nu^{\mathscr{P}}$ is unique.

Remark 1 *All axioms except axioms 8 (certainty equivalent) and 9 (denseness), are also necessary. The proof is omitted.*

The proof strategy is very intuitive. It mainly consists in three steps. First, to each act-information pair $(f, \mathscr{P}) \in \mathscr{F} \times \mathfrak{P}$, associate the Anscombe-Aumann act

$$\begin{array}{rccc} F^f : & \mathscr{P} & \longrightarrow & \Delta(X) \\ & P & \longmapsto & P^f. \end{array}$$

Then, define the relation $\succsim_{AA}$ on Anscombe-Aumann acts generated by act-information pairs letting

$$F^f \succsim_{AA} F^g \quad \Longleftrightarrow \quad (f, \mathscr{P}) \succsim (g, \mathscr{P})$$

and show that it can be extended to a binary relation on the set of all bounded Anscombe-Aumann acts. Finally, show that the extended relation $\succsim'_{AA}$ satisfies Schmeidler (89)'s axioms.

The theorem provides a very natural (as if) description of the decision-maker's behavior under objective ambiguity: given some imprecise information objectively describable by a (convex-ranged) set of probability distributions, the decision maker forms a prior regarding the relative likelihood of each of the scenarios associated with each probability distribution. This prior is not necessarily additive (and must not be, indeed, if the decision maker exhibits ambiguity aversion, as we shall see below). He or she then computes the average (in the sense of Choquet) expected utility of the acts considered and chooses the act with higher average expected utility. This decision procedure is consistent with an intuitive account of the Ellsberg paradox whereby ambiguity aversion is explained by the fact that the decision maker deems the unfavorable scenarios as more likely than the favorable ones.

3.2 An Example

Consider a machine that is out of order[2]. The decision maker has two possibilities: having the machine repaired or buying a new one. Having the machine repaired costs c while buying a new one costs p. The revenue from using the machine is b. Two states of the world are possible: either the machine works after having been repaired (state s), or it does not (state s'). Denote by f the act corresponding to having the machine repaired and g the act corresponding to buying a new one. We have $f(s) = b - c$, $f(s') = -c$ and $g(s) = g(s') = b - p$ as the net profit from the new machine is independent from the fact that the older one works after being repaired. We assume $b > p > c > 0$. In order to repair the machine, an electronic component is needed. This component can be of three different types A, B or C. The probability for the machine to be successfully repaired is p_1 if the component is of type A, p_2 if it is of type B, p_3 if it is of type C, with $p_1 < p_2 < p_3$. Moreover the information known about the average composition of a batch from which the component is taken is that the proportion of components of a given type is at most α, with $\frac{1}{2} \geq \alpha \geq \frac{1}{3}$. This can be summarized by a set

$$\Pi = \{\pi \in \Delta(\mathscr{P}) \mid \pi_i \leq \alpha, \forall i = 1, 2, 3\},$$

where $\mathscr{P} = \{p_1, p_2, p_3\}$, $\Delta(\mathscr{P})$ is the set of probability distributions over $\mathscr{P}$ and π_i is the probability of p_i. The condition $\pi_i \leq \alpha$ for all i is readily seen to imply that $1 - 2\alpha \leq \pi_i$ for all i. Jaffray (1989) shows that this set of probabilities can be represented by its lower envelope[3], the capacity ν_* such that $\nu_*(p_i) = 1 - 2\alpha$ for $i = 1, 2, 3$ and $\nu_*(\{p_i, p_j\}) = 1 - \alpha$ for $i, j = 1, 2, 3$, or by its upper envelope[4] ν^* defined by $\nu^*(p_i) = \alpha$ for $i = 1, 2, 3$ and $\nu^*(\{p_i, p_j\}) = 2\alpha$ for $i, j = 1, 2, 3$. Assuming that $u(b - c) = 2$, $u(b - p) = 1$ and $u(-c) = 0$, using the functional axiomatized in

[2]For the sake of simplicity, in this example, we focus on the use of the decision rule axiomatized and do not try to meet all the technical requirements of the theorem.

[3]In the sense that $\Pi = \{\pi \in \Delta(\mathscr{P}) \mid \pi \geq \nu_*\}$.

[4]In the sense that $\Pi = \{\pi \in \Delta(\mathscr{P}) \mid \pi \leq \nu^*\}$.

the theorem first with ν_* and second with ν^* yields the following values for f:

$$V_*(f) = 2(\alpha p_1 + \alpha p_2 + (1-2\alpha)p_3),$$

$$V^*(f) = 2((1-2\alpha)p_1 + \alpha p_2 + \alpha p_3)$$

and $V_*(g) = V^*(g) = 1$. Therefore, when he or she uses ν_* the decision maker must have the machine repaired if and only if

$$\alpha p_1 + \alpha p_2 + (1-2\alpha)p_3 \geq \frac{1}{2}$$

and when he or she uses ν^*:

$$(1-2\alpha)p_1 + \alpha p_2 + \alpha p_3 \geq \frac{1}{2}.$$

In words, in both cases some weighted average of the probabilities of success must exceed 1/2. The level of α can be seen as a measure of the imprecision of information concerning the proportion of components of a given type: the higher α, the higher the imprecision. Now imprecision of information can be seen alternatively as leaving room for a high probability of ending with a good component or with a bad one. Therefore, according to whether one sees the glass half-full of half-empty, imprecision can be seen as good or bad. The case of ν_* corresponds to the "half-empty" point of view: the higher α, the more demanding the rule is, as this gives more weight to the bad cases, requiring the lowest probability of success to be still rather good. On the contrary, the use of ν^* corresponds to the "half-full" point of view, as when imprecision increases it becomes a less strict decision rule, only asking for the highest probability of success to be high.

4 Special Cases

The representation given here is quite flexible (although not being completely unrestricted as the axioms show). It can indeed encompass a variety of models. We will next provide some examples.

4.1 Information-Based Choquet Expected Utility

The theorem yields a characterization of the decision rule used by the decision maker's which involves second-order beliefs, i.e. beliefs over probabilistic scenarios. However, it says nothing about his or her beliefs about states of the world, and in particular about the relationship between these beliefs and the objective information. The following proposition addresses this issue. It is a straightforward consequence of the properties of the Choquet integral: positive homogeneity, comonotonic additivity and monotonicity. We omit details.

Proposition 1

Let $\succsim$ satisfy all the conditions of the theorem. Then, for each $\mathscr{P}$, there exists a unique capacity $\rho^{\mathscr{P}} : \Sigma \to [0,1]$ such that, for all $A \in \Sigma$, for all $x, y \in X$ such that $(x, \mathscr{P}) \succ (y, \mathscr{P})$

$$V(xAy, \mathscr{P}) = \rho^{\mathscr{P}}(A)u(x) + (1 - \rho^{\mathscr{P}}(A))u(y). \tag{3}$$

Morevover, $\rho^{\mathscr{P}}$ is defined for all $A \in \Sigma$ by

$$\rho^{\mathscr{P}}(A) = \int_{\mathscr{P}} P(A)\ d\nu^{\mathscr{P}}(P) \tag{4}$$

and satisfies the following properties:

(i) For all $A, B \in \Sigma$,

$$(\forall P \in \mathscr{P}, P(A) \geq P(B)) \implies \rho^{\mathscr{P}}(A) \geq \rho^{\mathscr{P}}(B).$$

(ii) For all $A, B \in \Sigma$, such that $A \cap B = \varnothing$, if, for all $P, Q \in \mathscr{P}$, $P(A) > Q(A) \implies P(B) \geq Q(B)$, then $\rho^{\mathscr{P}}(A \cup B) = \rho^{\mathscr{P}}(A) + \rho^{\mathscr{P}}(B)$.

This proposition shows first that the preferences axiomatized in this paper belong to the biseparable class studied by Ghirardato and Marinacci (2001). Following their terminology, the capacity $\rho^{\mathscr{P}}$ may be interpreted as the decision maker's willingness to bet, i.e. the number of euros he or she is willing to pay for a bet yielding one euro if event A obtains and nothing otherwise. If one is willing to define the fact that A is deemed more likely than B if betting on A is preferred to betting on B, then $\rho^{\mathscr{P}}$ can be said to represent beliefs given information $\mathscr{P}$. However, as pointed out by Nehring (1994), in the context of ambiguity, this definition is somewhat arbitrary: one could also define belief by the fact that betting on the complement of B is preferred to betting on the complement of A, and, in the context of ambiguous information these notions would not be equivalent. Indeed, the second notion would be numerically represented by $\rho^{\mathscr{P}}$'s dual capacity $\bar{\rho}^{\mathscr{P}}$ defined by $\bar{\rho}^{\mathscr{P}}(A) = 1 - \rho^{\mathscr{P}}(A^c)$, which does not yield the same ordering on Σ.

This being said, it is noteworthy that willingness to bet is here defined from the available information as an aggregation of this information that satisfies a unanimity property: if in all probabilistic scenarios A is more likely than B, i.e. if A is unambiguously more likely than B, then the decision maker will be more willing to bet on A than to bet on B. This is property (i), a rationality property of subjective beliefs with respect to objective information. Property (ii) says, in turn, that if the scenarios in which disjoint events A and B are not very likely to obtain are the same, then the willingness to bet on the join of these events is the sum of the willingness to bet on each of them. This reflects the fact that in some sense there is no interaction between them, which would appear as an additional term in the sum.

An important consequence of this proposition is that, if $\nu^{\mathscr{P}}$ is additive, then so is $\rho^{\mathscr{P}}$. But this is incompatible with the Ellsberg paradox, as it is well known. Therefore, in order to be descriptively accurate and to account for ambiguity aversion, $\nu^{\mathscr{P}}$ must not be additive.

Given that $V(., \mathscr{P})$ is a Choquet expected utility when restricted to binary acts, one may wonder under what conditions it has this functional form on all acts. It turns out that it suffices to impose a dominance axiom to obtain this result. Define, for $A, B \in \Sigma$,

$$A \trianglerighteq^{\mathscr{P}} B \quad \Longleftrightarrow \quad (xAy, \mathscr{P}) \succsim (xBy, \mathscr{P}),$$

for some $x, y \in X$ with $(x, \mathscr{P}) \succ (y, \mathscr{P})$. This defines a comparative likelihood relation on events. The previous proposition shows that when all the axioms hold it is well defined and numerically represented by $\rho^{\mathscr{P}}$. Let, for $f \in \mathscr{F}$ and $x \in X$,

$$\{f \succsim x\}_{\mathscr{P}} := \{s \in S \mid (f(s), \mathscr{P}) \succsim (x, \mathscr{P})\}.$$

Consider the following axiom

Axiom 10 (Cumulative dominance (Sarin and Wakker 1992))
For all $f, g \in \mathscr{F}$, $(\forall x \in X, \{f \succsim x\}_{\mathscr{P}} \trianglerighteq^{\mathscr{P}} \{g \succsim x\}_{\mathscr{P}}) \quad \Longrightarrow \quad (f, \mathscr{P}) \succsim (g, \mathscr{P})$.

This axiom has the flavor of the first order stochastic dominance axiom, although it is stated here without any reference to a probability distribution. It says that if the decision maker believes that act f will yield a higher outcome than a given one with higher likelihood than g will, given information $\mathscr{P}$, then f is preferable to g.

Proposition 2
Let $\succsim$ be a preference relation represented by a SODEU functional. Then, $\succsim$ satisfies axiom 10 if and only if for all $(f, \mathscr{P}) \in \mathscr{F} \times \mathfrak{P}$,

$$V(f, \mathscr{P}) = \int_S u \circ f d\rho^{\mathscr{P}}.$$

The import of this proposition is twofold. First, it allows to understand how non-additive probabilities and Choquet expected utility arise from the perception of ambiguity by the decision maker. He or she aggregates the probabilistic information at his or her disposal through a weighted average, but in a way that is event-dependent: the weight that is attributed to a prior depends on the fact that it gives a relatively high or low probability to the event considered. This mechanism is however consistent with the probabilistic information in the sense that it respects the fact that some event has a uniformly higher probability than another. Second, it shows that Choquet expected utility is a special case of a more general model that obtains when some fairly natural rationality condition is satisfied. This reinforces the idea that the increased descriptive power of this model relative to expected utility does not come at the cost of abandoning too much of the normative status the latter was able to achieve.

4.2 Restricted Maxmin Expected Utility

Gajdos, Hayashi, Tallon, and Vergnaud (2006) axiomatize a version of maxmin expected utility where the decision maker maximizes the minimum expected utility over a subset of the objective set of priors. Specifically, let $\mathfrak{P}_c$ be the set of closed convex hulls of finite sets of countably additive non-atomic probability measures. By the Lyapunov theorem, $\mathfrak{P}_c \subseteq \mathfrak{P}$ and, moreover, $\mathfrak{P}_c$ is a mixture set when we define the following mixture operation:

$$\lambda\mathscr{P} + (1-\lambda)\mathscr{P}' = \{\lambda P + (1-\lambda)P' \mid P \in \mathscr{P}, P' \in \mathscr{P}'\}.$$

Adapting the notations to our setup, Gajdos, Hayashi, Tallon, and Vergnaud's axioms deliver a function $\Phi : \mathfrak{P} \to \mathfrak{P}_C$, where $\mathfrak{P}_C$ is the set of closed[5] convex hulls of elements of $\mathfrak{P}$, such that

(i) for all $\mathscr{P} \in \mathfrak{P}$, $\Phi(\mathscr{P}) \subseteq \overline{\text{co}}(\mathscr{P})$;

(ii) for all $\lambda \in [0,1]$, for all $\mathscr{P}, \mathscr{P}' \in \mathfrak{P}_c$,

$$\Phi(\lambda\mathscr{P} + (1-\lambda)\mathscr{P}') = \lambda\Phi(\mathscr{P}) + (1-\lambda)\Phi(\mathscr{P}')$$

and such that the decision maker maximizes the minimum expected utility over priors in $\Phi(\mathscr{P})$.

[5] In the weak* topology, i.e. the product topoloy.

We will now show that this functional form can be obtained in our setting by specifying the family of capacities $\nu^{\mathscr{P}}$ in an appropriate way. In order to introduce the restrictions needed to obtain this representation, we recall the following definition: a capacity ν defined over some measurable space $(\Omega, \mathscr{A})$ is *convex* if for all $A, B \in \mathscr{A}$,

$$\nu(A \cup B) + \nu(A \cap B) \geq \nu(A) + \nu(B).$$

We can now state the next proposition:

Proposition 3
Let $\succsim$ be a preference relation represented by the SODEU functional V such that $u(X)$ is connected and for all $\mathscr{P} \in \mathfrak{P}$, $\nu^{\mathscr{P}}$ is convex. Then, there exists a unique function $\Phi : \mathfrak{P} \to \mathfrak{P}_C$, such that for all $\mathscr{P} \in \mathfrak{P}$, $\Phi(\mathscr{P}) \subseteq \overline{\text{co}}(\mathscr{P})$ and

$$V(f, \mathscr{P}) = \min_{P \in \Phi(\mathscr{P})} \int u \circ f \mathrm{d}P. \tag{5}$$

Moreover, if, in addition, we have that, for all $\mathscr{P}, \mathscr{P}' \in \mathfrak{P}_c$, for all $\lambda \in [0, 1]$, for all $f \in \mathscr{F}$,

$$V(f, \lambda\mathscr{P} + (1-\lambda)\mathscr{P}') = \lambda V(f, \mathscr{P}) + (1-\lambda)V(f, \mathscr{P}'),$$

then

$$\Phi(\lambda\mathscr{P} + (1-\lambda)\mathscr{P}') = \lambda\Phi(\mathscr{P}) + (1-\lambda)\Phi(\mathscr{P}').$$

To what axioms does the previous restriction correspond in our setting? The answer is quite easy to figure out, given the axioms in Gajdos, Hayashi, Tallon, and Vergnaud (2006). Consider the following axioms:

Axiom 11 (Uncertainty Aversion) *For all $\mathscr{P} \in \mathfrak{P}$, for all $f, g \in \mathscr{F}$, for all $A \perp_{\mathscr{P}} \{f, g\}$,*

$$(f, \mathscr{P}) \sim (g, \mathscr{P}) \Longrightarrow (fAg, \mathscr{P}) \succsim (f, \mathscr{P}) \sim (g, \mathscr{P}).$$

Axiom 12 (Information Independence) *For all $\mathscr{P}, \mathscr{P}', \mathscr{P}'' \in \mathfrak{P}_c$, for all $f \in \mathscr{F}$, for all $\lambda \in [0, 1]$,*

$$(f, \mathscr{P}) \succsim (f, \mathscr{P}') \Longleftrightarrow (f, \lambda\mathscr{P} + (1-\lambda)\mathscr{P}'') \succsim (f, \lambda\mathscr{P}' + (1-\lambda)\mathscr{P}'').$$

By (Schmeidler 1989)'s argument, adding axiom 11 to the other axioms of the representation theorem implies immediately that $\nu^{\mathscr{P}}$ is convex for all $\mathscr{P} \in \mathfrak{P}$. We skip the details.

In turn, axiom 12 characterizes mixture linearity with respect to information sets, as shown by the next proposition:

Proposition 4
Let $\succsim$ be a preference relation represented by the SODEU functional V. Then $\succsim$ satisfies axiom 12 if and only if for all $f \in \mathscr{F}$, $\mathscr{P}, \mathscr{P}' \in \mathfrak{P}_c$,

$$V(f, \lambda\mathscr{P} + (1-\lambda)\mathscr{P}') = \lambda V(f, \mathscr{P}) + (1-\lambda)V(f, \mathscr{P}').$$

For completeness we repeat the simple proof of Gajdos, Hayashi, Tallon, and Vergnaud (2006). See appendix.

4.3 The security level and potential outcome effects

Gajdos, Hayashi, Tallon, and Vergnaud (2006), Kopylov (2006) have axiomatized a special case of the restricted maxmin of the following form:

$$V^S(f, \mathscr{P}) = (1-\varepsilon)\int u \circ f \mathrm{d}c(\mathscr{P}) + \varepsilon \min_{P \in \mathscr{P}} \int u \circ f \mathrm{d}P,$$

where $c(\mathscr{P}) \in pc(\Sigma)$ and $\varepsilon \in [0,1]$. This functional form, that we shall call the security level functional, has many possible interpretations, more or less formally grounded. Ellsberg, who introduced it, interprets the coefficient $(1-\varepsilon)$ as degree of confidence in the probability estimate $c(\mathscr{P})$. The idea is that the decision maker forms prior $c(\mathscr{P})$ based on information $\mathscr{P}$, computes the expected utility with respect to it, and then corrects it by a factor $\varepsilon(\min_{P \in \mathscr{P}} \int u \circ f \mathrm{d}P - \int u \circ f \mathrm{d}c(\mathscr{P}))$ that takes into accounts mistakes by computing the difference between the estimated expected utility and the prudent maxmin rule.

Another interpretation, the one found in Gajdos, Hayashi, Tallon, and Vergnaud (2006), is that ε measures the aversion towards imprecision of the decision maker. This interpretation if grounded on a comparative definition of imprecision aversion and a definition of relative imprecision premium measuring the willingness to pay of the decision maker for a more precise information.

Finally, an interpretation that is not mentioned in any paper to the best of our knowledge, but that can be inferred from Kopylov's axiomatization is that the decision maker cares for the security level of a given act, i.e. he or she favors acts that guarantee a higher utility outcome, to an extent measured by coefficient ε, so that his or her choice depends on both the expected utility of the act and its security level. The idea to introduce this security factor in the study of decision under risk is usually attributed to Lopes (1987). It was axiomatized by Jaffray (1988) and Gilboa (1988) in that context.

Lopes not only pointed out the importance of the security factor in the decision making process under risk. She also drew attention on a symmetric factor: the potential factor of a given act, i.e. the maximal outcome that this act can yield. Cohen (1992) has proposed a non continuous model that incorporates both the security factor and the potential factor, along with expected utility, in the decision making process under risk. In the context of imprecise risk we are considering here, it is also possible to incorporate both factors in a decision model, by generalizing the V^S functional to a functional V^{SP} defined as follows:

$$V^{SP}(f,\mathscr{P}) = (1-\varepsilon)\int u\circ f\mathrm{d}c(\mathscr{P}) + \varepsilon(\gamma \min_{P\in\mathscr{P}}\int u\circ f\mathrm{d}P + (1-\gamma)\max_{P\in\mathscr{P}}\int u\circ f\mathrm{d}P),$$

with $\gamma\in[0,1]$.

This functional form can be retrieved by choosing for $\nu^{\mathscr{P}}$ a capacity of the following type. Let $\upsilon_\varnothing$ be the capacity defined by $\upsilon_\varnothing(\mathcal{P})=1$ for all non-empty $\mathcal{P}\subseteq\mathscr{P}$ and $\upsilon_{\mathscr{P}}$ be defined by $\upsilon_{\mathscr{P}}(\mathcal{P})=0$ for all $\mathcal{P}\subsetneq\mathscr{P}$. Let $m(\mathscr{P})$ be a probability charge on $2^{\mathscr{P}}$. Let $\nu^{\mathscr{P}}$ be defined by:

$$\nu^{\mathscr{P}} = (1-\varepsilon)m(\mathscr{P}) + \varepsilon(\gamma\upsilon_{\mathscr{P}} + (1-\gamma)\upsilon_\varnothing).$$

This kind of capacity is introduced and studied in Chateauneuf, Eichberger, and Grant (2006), where they are dubbed *neo-additive capacities*. Now, if we define $c(\mathscr{P})$ by

$$c(\mathscr{P})(A) = \int_{\mathscr{P}} P(A)\mathrm{d}m(\mathscr{P})(P),$$

and if V is the SODEU functional associated with $\nu^{\mathscr{P}}$, then $V=V^{SP}$.

The axiomatic foundations of this functional form in the present context can be deduced from the work in Chateauneuf, Eichberger, and Grant (2006). This requires some definitions and a bit of notation.

Definition 3 *Let $(f,\mathscr{P})\in\mathscr{F}\times\mathfrak{P}$ and $\pi\in\Delta(X)$. π is in the indifference set of the infimum of $(f,\mathscr{P})$, denoted* $\mathrm{Inf}(f,\mathscr{P})$, *if*

(i) $(\underline{f},\mathscr{P})\sim(f,\mathscr{P})$, *where $\underline{f}$ is such that*

$$\begin{aligned} P^{\underline{f}} &\sim_{\mathscr{P}} P^f \qquad \forall P\notin\underline{\mathcal{P}} \\ P^{\underline{f}} &\sim_{\mathscr{P}} \pi \qquad \forall P\in\underline{\mathcal{P}} \end{aligned}$$

and $\underline{\mathcal{P}} = \{P\in\mathscr{P}\mid \pi\succ_{\mathscr{P}} P^f\}$, *(ii)* $(\underline{g}, \mathscr{P}) \succ (f, \mathscr{P})$*, where, for all* $\pi' \succ_{\mathscr{P}} \pi$*,* $\underline{g}$ *is such that*

$$\begin{aligned} P^{\underline{g}} &\sim_{\mathscr{P}} P^f && \forall P \notin \underline{\mathcal{Q}} \\ P^{\underline{g}} &\sim_{\mathscr{P}} \pi' && \forall P \in \underline{\mathcal{Q}} \end{aligned}$$

and $\underline{\mathcal{Q}} = \{P \in \mathscr{P} \mid \pi' \succsim_{\mathscr{P}} P^f\}$.

The indifference set of the supremum of $(f, \mathscr{P})$*,* $\mathrm{Sup}(f, \mathscr{P})$*, is defined similarly by reversing all inequalities.*

When there is no risk of confusion, by a slight abuse of notation we denote by $\mathrm{Inf}(f, \mathscr{P})$ and $\mathrm{Sup}(f, \mathscr{P})$ arbitrary elements of these sets. These sets give the preference-based definition of extreme distributions induced by a given act.If $\mathscr{P}^f$ is finite, then the indifference class of its smallest elements w.r.t. $\succsim_{\mathscr{P}}$ is $\mathrm{Inf}(f, \mathscr{P})$.

To each act-information pair, we associate the following sets of acts: $\underline{\mathscr{F}}(f, \mathscr{P})$, which is the set of acts h such that there exists a scenario where both f and h induce a lottery that is worse than their respective infimum indifference sets:

$$\underline{\mathscr{F}}(f, \mathscr{P}) := \{h \in \mathscr{F} \mid \exists P \in \mathscr{P}, P^f \precsim_{\mathscr{P}} \mathrm{Inf}(f, \mathscr{P}) \text{ and } P^h \precsim_{\mathscr{P}} \mathrm{Inf}(h, \mathscr{P})\},$$

and $\overline{\mathscr{F}}(f, \mathscr{P})$ which is the set of acts h such that there exists a scenario where both f and h induce a lottery that is better than their respective supremum indifference sets:

$$\overline{\mathscr{F}}(f, \mathscr{P}) := \{h \in \mathscr{F} \mid \exists P \in \mathscr{P}, P^f \succsim_{\mathscr{P}} \mathrm{Sup}(f, \mathscr{P}) \text{ and } P^h \succsim_{\mathscr{P}} \mathrm{Sup}(h, \mathscr{P})\}$$

We introduce the following axiom: let $\mathfrak{P}_4 \subset \mathfrak{P}$ be the set of information sets containing at least 4 distinct priors.

Axiom 13 (Extreme Scenarios Sensitivity) *For any* $f, g, h \in \mathscr{F}$*, for any* $\mathscr{P} \in \mathfrak{P}_4$ *such that* $(f, \mathscr{P}) \sim (g, \mathscr{P})$ *and* $h \in \underline{\mathscr{F}}(g, \mathscr{P}) \cap \overline{\mathscr{F}}(g, \mathscr{P})$*, for any* $A \perp_{\mathscr{P}} \{f, g, h\}$*,*

1. *If* $h \in \underline{\mathscr{F}}(f, \mathscr{P})$ *then* $(gAh, \mathscr{P}) \succsim (fAh, \mathscr{P})$*,*

2. *If* $h \in \overline{\mathscr{F}}(f, \mathscr{P})$ *then* $(fAh, \mathscr{P}) \succsim (gAh, \mathscr{P})$*.*

The first part of this axiom says that if f and h induce their unfavorable distributions in the same scenarios, then hedging g with h can be no worse than hedging f with h. The second part of this axiom has a similar interpretation.

This axiom is essentially necessary and sufficient to characterize the functional V^{SP} as a special case of the SODEU functional, as shown by the following proposition.

Proposition 5

Let $\succsim$ be represented by an SODEU functional V such that for all $\mathscr{P} \in \mathfrak{P}$, $\nu^{\mathscr{P}}(\mathcal{P}) = 0 \implies \mathcal{P} = \varnothing$, for all $\mathcal{P} \subseteq \mathscr{P}$. Assume that $\succsim$ satisfies axiom 9 and that there exists $\bar{\pi}, \underline{\pi} \in \Delta(X)$ such that $\bar{\pi} \succ_{\mathscr{P}} \underline{\pi}$ and for all $\pi \in \Delta(X)$, $\bar{\pi} \succsim_{\mathscr{P}} \pi \succsim_{\mathscr{P}} \underline{\pi}$. Then $V(f, \mathscr{P}) = V^{SP}(f, \mathscr{P})$ for all $f \in \mathscr{F}$ and for all $\mathscr{P} \in \mathfrak{P}_4$ if and only if $\succsim$ satisfies axiom 13.

4.4 Weighted Worst Scenarios

A very extreme special case of the functional V^{SP} introduced above obtains by taking $\varepsilon = \gamma = 1$. It is the well-known maxmin rule axiomatized in a fully subjective setting by Gilboa and Schmeidler (1989). In the objective setting, an axiomatization was given by Wang (2001, revised 2003). The idea of this axiomatization is that decision makers usually prefer having precise rather than imprecise information. Therefore, Wang defines what he calls uncertainty aversion and what, following the remarks in Gajdos, Hayashi, Tallon, and Vergnaud (2006), we call strong aversion to imprecision:

Definition 4 *A decision maker is strongly averse to imprecision if, for all $\mathscr{P}, \mathscr{P}' \in \mathfrak{P}$,*

$$\mathscr{P} \subseteq \mathscr{P}' \implies (f, \mathscr{P}) \succsim (f, \mathscr{P}'), \ \forall f \in \mathscr{F}.$$

As a consequence of this definition, we have the following result, which is a version in our setting of similar results in the literature (see, e.g., Wang (2001, revised 2003, Theorem 3.1)):

Proposition 6

A decision maker satisfying the axioms of the representation theorem is strongly averse to imprecision if and only if for all sets $\mathscr{P} \in \mathfrak{P}$ and all $f \in \mathscr{F}$,

$$V(f, \mathscr{P}) = \inf_{P \in \mathscr{P}} \int_S u \circ f \mathrm{d}P.$$

This decision rule is rather extreme. To see that consider the following example:

Example 3 *Let $S = [0, 1]$, $X = \mathbb{R}$, $u(x) = x$ for all $x \in X$ and $f = \mathbb{1}_{[0, \frac{1}{3}]}$. Consider the set $\mathscr{P} = \{P_1, P_2, P_3\}$ where P_1 is the Lebesgue measure on $[0, 1]$, P_2 defined by $P_2(A) = P_1(A \mid [\frac{2}{3}, 1])$ and $P_3 = \varepsilon P_1 + (1 - \varepsilon) P_2$, $\varepsilon \in [0, 1]$. Then if the decision maker has strong aversion to imprecision, then $(f, \{P_3\}) \succ (f, \mathscr{P})$ for all $\varepsilon > 0$. But if ε is very small, this mean preferring a precise but very unfavorable information set to an imprecise one but that contains at least one not so unfavorable scenario:*

P_1.

Now the SODEU functional contains as a special case a functional that generalizes the maxmin rule in the following sense: instead of considering only the overall worst scenario, the decision maker considers also worst scenarios on some subsets of the information set, and attributes some weight to these less severe worst scenarios. The intuition behind this decision rule would be that the decision maker does not consider himself or herself unlucky enough to only consider the worst possible scenario. He or she considers that she might as well be lucky and have to deal only with a less unfavorable set of scenarios than the one he or she is facing. Another way of understanding this rule is that the decision maker might have some doubt as to the objective nature of the set of priors he or she is facing, and this doubt may be represented by a probability distribution on the subsets of this set. The idea would be that the objectively given set is but the upper bound on how imprecise the information is but that it can be less imprecise with some probability.

To avoid technical complications and focus on the intuition, we will present the rule in the case where the information set is finite. Denote $\mathfrak{P}_0$ the set of finite information sets. The functional $V : \mathscr{F} \times \mathfrak{P}_0$ is a *Weighted Worst Scenarios (WWS) functional* if for all $\mathscr{P} \in \mathfrak{P}_0$ there exists a probability distribution $\alpha^{\mathscr{P}}$ on $2^{\mathscr{P}} \setminus \{\varnothing\}$ such that for all $f \in \mathscr{F}$,

$$V(f, \mathscr{P}) = \sum_{\varnothing \neq \mathcal{P} \subseteq \mathscr{P}} \alpha^{\mathscr{P}}(\mathcal{P}) \min_{P \in \mathcal{P}} \int u \circ f \mathrm{d}P.$$

This functional is indeed a special case of the SODEU functional: it suffices to take for $\nu^{\mathscr{P}}$ a belief function and take for $\alpha^{\mathscr{P}}$ its Moebius transform. Recall that a capacity ν defined on some measurable space $(\Omega, \mathscr{A})$ is a belief function if and only if for all $n \geq 1$, for all $A_1, \ldots, A_n \in \mathscr{A}$,

$$\nu(\cup_{i=1}^n A_i) \geq \sum_{\varnothing \neq I \subseteq \{1,\ldots,n\}} (-1)^{|I|+1} \nu(\cap_{j \in I} A_j).$$

We refer to Gilboa and Schmeidler (1995) for details and for how to extend this functional form to infinite sets $\mathscr{P}$.

The WWS functional provides an answer to the paradox raised in example 3. Indeed, simple calculations show that

$$(f, \{P_3\}) \succ (f, \mathscr{P}) \Longleftrightarrow \varepsilon > \frac{\alpha^{\mathscr{P}}(P_1)}{1 - \alpha^{\mathscr{P}}(P_3) - \alpha^{\mathscr{P}}(\{P_1, P_3\})}.$$

Therefore, the decision maker will prefer the precise scenario P_3 only if this scenario is not too unfavorable, where the notion of "too unfavorable" is a subjective notion related to $\alpha^{\mathscr{P}}$.

Notice that the WWS functional can be rewritten in a form that makes it easily comparable to the security level functional V^S. If $\mathscr{P} = \{P_1, \ldots, P_n\}$, letting

$$\varepsilon = 1 - \sum_{i=1}^{n} \alpha^{\mathscr{P}}(P_i), \quad c(\mathscr{P}) = \sum_{i=1}^{n} \frac{\alpha^{\mathscr{P}}(P_i)}{1-\varepsilon} P_i \quad \text{and} \quad \beta^{\mathscr{P}} = \frac{1}{\varepsilon}\alpha^{\mathscr{P}},$$

we have:

$$V^{WWS}(f, \mathscr{P}) = (1-\varepsilon)\int u \circ f \mathrm{d}c(\mathscr{P}) + \varepsilon(\sum_{\substack{\mathcal{P}\subseteq\mathscr{P},\\ |\mathcal{P}|>1}} \beta^{\mathscr{P}}(\mathcal{P}) \min_{P\in\mathcal{P}} \int u \circ f \mathrm{d}P).$$

This formula shows that the WWS functional is a generalization of the V^S functional of a different type than the V^{SP} functional. We leave its axiomatization for further research.

5 Ambiguity Attitude

One of the main assets of a setting with an objective but ambiguous probabilistic information is that one can clearly distinguish perceived ambiguity from objective ambiguity. Ambiguity attitudes can affect perceived ambiguity, as shown by the first representation theorem in Gajdos, Hayashi, Tallon, and Vergnaud (2006), so that the interpretation of a fully subjective set of priors like the one that appears in the axiomatization of Gilboa and Schmeidler (1989) is not completely clear. Gajdos, Hayashi, Tallon, and Vergnaud have shown in a precise way how ambiguity attitude affects perceived ambiguity: the less ambiguity averse the decision maker is, the smaller number of priors he or she takes into account. At the limit, ambiguity neutral decision makers collapse the set of priors to a unique prior.

In this section, we intend to pursue the analysis of ambiguity attitude in the context of the SODEU functional.

5.1 Ambiguity Aversion

What is however the natural definition of ambiguity aversion in our setting? In order to answer this question, we introduce, for all $\mathscr{P} \in \mathfrak{P}$ the notion of $\mathscr{P}$-unambiguous acts

Definition 5 *For all $\mathscr{P} \in \mathfrak{P}$, $f \in \mathscr{F}$ is a $\mathscr{P}$-unambiguous act if $P^f = Q^f, \forall P, Q \in \mathscr{P}$.*

Notice that a $\mathscr{P}$-unambiguous act is $\mathscr{P}$-comonotonic to any act in $\mathscr{F}$. Therefore, $\mathscr{P}$-unambiguous acts are *crisp* in the sense of Ghirardato, Maccheroni, and Marinacci (2001), i.e. hedging with respect to them does not affect preference. Let $\mathscr{F}^{\mathscr{P}-ua}$ be the set of $\mathscr{P}$-unambiguous acts. We can now give the following definition of comparative ambiguity aversion given information $\mathscr{P}$.

Definition 6 *Let $\succsim_1$ and $\succsim_2$ be the preference relations of two decision makers. Then decision maker 1 is more ambiguity averse than decision maker 2 given information $\mathscr{P}$ if and only if, for all $k \in \mathscr{F}^{\mathscr{P}-ua}$, for all $f \in \mathscr{F}$:*

$$(f, \mathscr{P}) \succsim_1 (k, \mathscr{P}) \quad \Longrightarrow \quad (f, \mathscr{P}) \succsim_2 (k, \mathscr{P})$$

and

$$(f, \mathscr{P}) \succ_1 (k, \mathscr{P}) \quad \Longrightarrow \quad (f, \mathscr{P}) \succ_2 (k, \mathscr{P})$$

This definition of ambiguity aversion is similar to the ones in Epstein (1999) and Ghirardato and Marinacci (2002), and it is most natural. The following proposition is an immediate consequence of the results of the latter.

Proposition 7
Let $\succsim_1$ and $\succsim_2$ be the preference relations of two decision makers satisfying the axioms of the representation theorem. Then decision maker 1 is more ambiguity averse than decision maker 2 given information $\mathscr{P}$ if and only if $\nu_2^{\mathscr{P}} \geq \nu_1^{\mathscr{P}}$.

This shows, as in Ghirardato and Marinacci (2002), that a decision maker that is less ambiguity averse than another one is always more willing to bet, always more confident, always less pessimistic.

5.2 Imprecision Aversion

Now, we have analyzed the behavior of a decision-maker given a fixed information. How will changing information affect his behavior? It is rather natural to consider that an ambiguity averse decision maker, when about to make a bet, will most of the time prefer having precise information on the probabilities rather than imprecise one, whatever the nature of the bet. This idea leads to the following definition of *imprecision aversion* (see Gajdos, Hayashi, Tallon, and Vergnaud (2006)).

Definition 7 *Let $\succsim_1$ and $\succsim_2$ be the preference relations of two decision makers.*

Then decision maker 1 is more imprecision averse than decision maker 2 given *information $\mathscr{P}$ if and only if, for all $A \in \Sigma$, $x, y \in X$ such that $(x, \mathscr{P}) \succ_i (y, \mathscr{P})$, $i = 1, 2$, and $P \in co(\mathscr{P})$,*

$$(xAy, \mathscr{P}) \succsim_1 (xAy, \{P\}) \quad \Longrightarrow \quad (xAy, \mathscr{P}) \succsim_2 (xAy, \{P\}).$$

It turns out that this notion of imprecision aversion is intimately linked to the willingness to bet of the agent.

Proposition 8
Let $\succsim_1$ and $\succsim_2$ be the preference relations of two decision makers with an SODEU representation. Then, decision maker 1 is more imprecision averse than decision maker 2 given information $\mathscr{P}$ if and only if $\rho_2^{\mathscr{P}} \geq \rho_1^{\mathscr{P}}$.

It is a straightforward corollary of both propositions that in the SODEU model ambiguity aversion implies imprecision aversion.

6 Conclusion

We have axiomatized in what seems to us a rather simple way a model of decision making under ambiguous objective information (imprecise risk) where the decision maker maximizes the (Choquet) average expected utility of a given act with respect to some second order belief. We characterized some special cases, among which the case where this functional form reduces to Choquet Expected Utility with respect to a capacity that is consistent with information in the sense that it gives higher likelihood to A than to B whenever for each prior A is more likely than B. This provides foundations for the intuition according to which decision makers facing imprecise risk aggregate information into one single likelihood measure in a way that is nevertheless compatible with ambiguity aversion. We show how different forms of ambiguity aversion can be characterized in our model in terms of the different capacities that can be defined in the SODEU model, and how they are related to each other.

There are some open problems that we plan to address in future research. Some are of a technical nature, some of a conceptual one, some of a practical one.

As for technical issues, our axiomatization of second order beliefs has in our opinion the advantage on some other axiomatizations of not taking as primitives some second-order devices like second-order acts in Klibanoff, Marinacci, and Mukerji (2005) or lotteries over acts in Seo (2006). However this comes at the cost of an extra non-necessary axiom, the denseness axiom, and of adding more structure on the admissible sets of priors, namely range-convexity, and therefore ruling out complete ignorance as modeled by taking the set of all possible priors. The question of how to keep the same advantage without incurring that cost is open for future research.

As for conceptual issue, we see at least three interesting possible developments. First, some time should be devoted in future research to the links between the second order capacities associated with different information sets. This question is intimately linked to the question of updating information and conditioning capacities, which, as it is well known, does not have a canonical solution. Second, as we pointed out, the SODEU model is dual to the SOSEU model of (Klibanoff, Marinacci, and Mukerji 2005) and (Seo 2006). Therefore, it would be interesting to axiomatize the nesting model, in order to make their relationship clear in terms of behavioral foundations. Third, our model, like many others, is not consistent with the Allais paradox, in the sense that it reduces to expected utility for precise risk. Therefore, it would be desirable to extend it to cover this case.

As for practical issues, applications should be developed to at least to different domains. First, in principal agent problems, it is often assumed that the principal and the agent have a common prior. Our model might be a way of studying the case where they have a common *imprecise* prior. Second, in location models in industrial organization, firms do know in a precise way the distribution of customers on the product space. Our model could be used to study the case of an imprecisely known distribution.

A Proofs

Theorem 1. Before proving the theorem, we prove a very useful lemma that is of independent interest and that will be used repeatedly in the proof.

Lemma 1 *Let $\mathscr{P} \in \mathfrak{P}$, $F \subseteq \mathscr{F}$ finite and $\alpha \in (0,1)$. Then, there exists $A \in \Sigma$ such that $P(A) = \alpha$ for all $P \in \mathscr{P}$ and $A \perp_{\mathscr{P}} F$.*

Proof. We shall prove the proposition for three acts f, g, h only. Generalization to any finite F is straightforward but involves cumbersome notations.

Let f, g, h be simple acts. Therefore, there exist

- finite sequences $\{x_i\}_{1\leq i\leq n}$, $\{y_j\}_{1\leq j\leq m}$ and $\{z_k\}_{1\leq k\leq p}$ of elements of X,
- finite measurable partitions $\{F_i\}_{1\leq i\leq n}$, $\{G_j\}_{1\leq j\leq m}$ and $\{H_k\}_{1\leq k\leq p}$ of S

such that $f(F_i) = \{x_i\}$, $g(G_j) = \{y_j\}$, $h(H_k) = \{z_k\}$.

By range-convexity of $\mathscr{P}$, for all i, j, k, there exists $A_{ijk} \subseteq F_i \cap G_j \cap H_k$ such that

$$P(A_{ijk}) = \alpha P(F_i \cap G_j \cap H_k), \quad \forall P \in \mathscr{P}.$$

Let $A := \bigcup_{i,j,k} A_{ijk}$.

Clearly, because $\{F_i\}_{1 \le i \le n}$, $\{G_j\}_{1 \le j \le m}$ and $\{H_k\}_{1 \le k \le p}$ are partitions, $P(A) = \alpha$ for all $P \in \mathscr{P}$. Now let us show that it is independent from f (the proof for g and h are identical).

Take $x \in X$. If $x \notin f(S)$, then $P(f^{-1}(x) \cap A) = 0 = 0 \times P(A) = P^f(x)P(A)$. If $x \in f(S)$, there exists i_0 such that $x = x_{i_0}$. Therefore, for all $P \in \mathscr{P}$,

$$\begin{aligned} P(f^{-1}(x) \cap A) &= P(F_{i_0} \cap A) = P\left(F_{i_0} \cap \bigcup_{i,j,k} A_{ijk}\right) \\ &= P\left(\bigcup_{j,k} A_{i_0jk}\right) \qquad \text{as } \{F_i\}_{1 \le i \le n} \text{ is a partition} \\ &= \sum_{j,k} P(A_{i_0jk}) \qquad \text{as } \{G_j\}_{1 \le j \le m} \text{ and } \{H_k\}_{1 \le k \le p} \text{ are partitions} \\ &= \alpha \sum_{j,k} P(F_{i_0} \cap G_j \cap H_k) = \alpha P\left(F_{i_0} \cap \bigcup_{j,k} (G_j \cap H_k)\right) = \alpha P(F_{i_0}) \\ &= P(A)P(f^{-1}(x)). \end{aligned}$$

□

The proof of the theorem proceeds in several steps.

Step 1. Fix $\mathscr{P} \in \mathfrak{P}$. Given $f \in \mathscr{F}$, one can canonically associate a function:

$$\begin{aligned} F^f_{\mathscr{P}} :& \mathscr{P} \to \Delta(X) \\ & P \mapsto P^f. \end{aligned}$$

This function will be call the *Anscombe-Aumann or AA-act generated by f under information $\mathscr{P}$.*

Let $\mathscr{A}(\mathscr{F}, \mathscr{P}) := \{F \in \Delta(X)^{\mathscr{P}} \mid \exists f \in \mathscr{F}, F = F^f_{\mathscr{P}}\}$ be the set of all AA-act generated by $\mathscr{F}$ under information $\mathscr{P}$. For convenience and when no confusion might arise, we shall drop the reference to $\mathscr{P}$ and write only F^f and $\mathscr{A}(\mathscr{F})$.

Step 2. Consider relation $\succsim_{\mathscr{P}}$. The following lemma will imply that it is reflexive.

Lemma 2 *For all $\mathscr{P} \in \mathfrak{P}$, for all $\pi \in \Delta(X)$, there exists $k \in \mathscr{F}$ such that, $P^k = \pi$ for all $P \in \mathscr{P}$.*

Proof. Let $E = \{x_1, ..., x_n\}$ be the support of π. The proof will proceed by induction on the size of E.

If $n = 1$, π is a degenerate measure with atom x_1. Therefore, as x_1 generates $\pi = \delta_{x_1}$ for all $P \in \mathscr{P}$, we can take $k = x_1$.

Now assume the lemma is true for $n \geq 1$ and show it therefore holds for $n + 1$. Take $x_1 \in E$. Because $n + 1 \geq 2$, we have $0 < \pi(x_1) < 1$. Define π^{x_1} by:

$$\begin{cases} \pi^{x_1}(x) = \frac{\pi(x)}{1-\pi(x_1)} & \text{if } x \neq x_1 \\ \pi^{x_1}(x_1) = 0. \end{cases}$$

The size of the support of π^{x_1} is now n. We can therefore apply the induction hypothesis to find an act k_1 such that $P^{k_1} = \pi^{x_1}$ for all $P \in \mathscr{P}$. Now, by range-convexity of $\mathscr{P}$, it is possible to find a set $A_1 \in \Sigma$ such that $P^{x_1 A_1 k_1} = \pi(x_1)\delta_{x_1} + (1 - \pi(x_1))\pi^{x_1}$ for all $P \in \mathscr{P}$, where $x_1 A_1 k_1$ is the act yielding x_1 on A_1 and equal to k_1 elsewhere. Setting $k = x_1 A_1 k_1$ thus completes the proof, as $\pi(x_1)\delta_{x_1} + (1 - \pi(x_1))\pi^{x_1} = \pi$. □

As a consequence, it implies, together with axiom 5, that if $F^f = F^g$, then $(f, \mathscr{P}) \sim (g, \mathscr{P})$. Therefore, one can define a preference relation $\succsim_{AA}^{\mathscr{P}}$ on $\mathscr{A}(\mathscr{F})$ by setting:

$$F^f \succsim_{AA}^{\mathscr{P}} F^g \iff (f, \mathscr{P}) \succsim (g, \mathscr{P}).$$

Here again, we shall drop $\mathscr{P}$ when no confusion might arise.

As a consequence of lemma 1, by range-convexity of $\mathscr{P}$, for any $f, g \in \mathscr{F}$ and any $\alpha \in [0, 1]$ there exists $A \in \Sigma$ such that, for all $P \in \mathscr{P}$,

$$P^{fAg} = \alpha P^f + (1 - \alpha)P^g,$$

i.e. $F^{fAg} = \alpha F^f + (1 - \alpha)F^g$. Therefore, the set $\mathscr{A}(\mathscr{F})$ is convex. We wish to show that preference $\succsim_{AA}$ on $\mathscr{A}(\mathscr{F})$ satisfies all the axioms of the Choquet Expected Utility model of Schmeidler (1989). Clearly $\succsim_{AA}$ is a weak order because $\succsim$ is by axiom 1. We shall enumerate the other axioms as claims.

Claim 1 (Continuity) *For all $F^f, F^g, F^h \in \mathscr{A}(\mathscr{F})$, if $F^f \succ_{AA} F^g \succ_{AA} F^h$, then there exists $\alpha, \beta \in]0, 1[$ such that:*

$$\alpha F^f + (1 - \alpha)F^h \succ_{AA} F^g \quad \text{and} \quad F^g \succ_{AA} \beta F^f + (1 - \beta)F^h.$$

Proof. This follows automatically from axiom 2. □

The following remark will be useful:

Remark 2 *Lemma 2 implies that all constant AA-acts belong to $\mathscr{A}(\mathscr{F})$. Identifying constant AA-acts in $\mathscr{A}(\mathscr{F})$ and elements of $\Delta(X)$, one can therefore consider the restriction to $\Delta(X)$ of the relation $\succsim_{AA}$. We have, for $\pi, \pi' \in \Delta(X)$, $k \in K(\pi, \mathscr{P})$, $k' \in K(\pi', \mathscr{P})$*

$$\pi \succsim_{AA} \pi' \iff F^k \succsim_{AA} F^{k'} \iff (k, \mathscr{P}) \succsim (k', \mathscr{P}) \iff \pi \succsim_{\mathscr{P}} \pi'.$$

Now, two AA-acts F^f and F^g are comonotonic if and only if, for all $P, Q \in \mathscr{P}$:

$$F^f(P) \succ_{AA} F^f(Q) \implies F^g(P) \succsim_{AA} F^g(Q).$$

This formula is equivalent to: $P^f \succ_{AA} Q^f \implies P^g \succsim_{AA} Q^g$,

i.e., by the previous remark, to: $P^f \succ_{\mathscr{P}} Q^f \implies P^g \succsim_{\mathscr{P}} Q^g$.

Hence, two AA-acts F^f and F^g are comonotonic if and only f, g are $\mathscr{P}$-comonotonic. This allows us to state the following claim:

Claim 2 (Comonotonic Independence) *For all $F^f, F^g, F^h \in \mathscr{A}(\mathscr{F})$ pairwise comonotonic and for all $\alpha \in [0,1]$:*

$$F^f \succsim_{AA} F^g \iff \alpha F^f + (1-\alpha)F^h \succsim_{AA} \alpha F^g + (1-\alpha)F^h.$$

Proof. By lemma 1, there exists $A \in \Sigma$, $A \perp_{\mathscr{P}} \{f, g, h\}$, such that $F^{fAh} = \alpha F^f + (1-\alpha)F^h$ and $F^{gAh} = \alpha F^g + (1-\alpha)F^h$. Therefore, because F^f, F^g, F^h pairwise comonotonic implies by the previous remarks f, g, h pairwise $\mathscr{P}$-comonotonic,

$$\begin{aligned} \alpha F^f + (1-\alpha)F^h \succsim_{AA} \alpha F^g + (1-\alpha)F^h &\iff (fAh, \mathscr{P}) \succsim (gAh, \mathscr{P}) \\ &\iff (f, \mathscr{P}) \succsim (g, \mathscr{P}) \quad \text{by axiom 4;} \\ &\iff F^f \succsim F^g. \end{aligned}$$ □

The next claim follows directly from the previous remark and axiom 5:

Claim 3 (Dominance) *For all $F^f, F^g \in \mathscr{A}(\mathscr{F})$, if $F^f(P) \succsim_{AA} F^g(P)$ for all $P \in \mathscr{P}$, then $F^f \succsim_{AA} F^g$.*

The final claim of this step of the proof follows from axiom 6:

Claim 4 (Non-Triviality) *There exist* $F^f, F^g \in \mathscr{A}(\mathscr{F})$ *s.t.* $F^f \succ_{AA} F^g$.

Step 3. We will now proceed to construct the objects of the theorem. Because $\succsim_{AA}$ is a weak order, because of claims 1, 2 and 4 and because all constant AA-acts are pairwise comonotonic and belong to $\mathscr{A}(\mathscr{F})$, restricting $\succsim_{AA}$ to constant acts allows to show, as a consequence of the Mixture-Space Theorem, that there exists an affine non-constant function $U_{\mathscr{P}} : \Delta(X) \to \mathbb{R}$, unique up to an increasing affine transformation, such that, for all $\pi, \pi' \in \Delta(X)$:

$$\pi \succsim_{AA}^{\mathscr{P}} \pi' \iff U_{\mathscr{P}}(\pi) \geq U_{\mathscr{P}}(\pi').$$

We now wish to apply the Choquet Expected Utility theorem of Schmeidler (1989). However, this theorem is *a priori* valid only for the whole set of bounded Anscombe-Aumann acts, i.e. the set of all measurable functions F from $\mathscr{P}$ to $\Delta(X)$ such that there exists $\bar{\pi}, \underline{\pi} \in \Delta(X)$ with

$$\bar{\pi} \succsim_{AA} F(P) \succsim \underline{\pi}, \quad \forall P \in \mathscr{P}.$$

We shall denote this set $\mathscr{A}(\mathscr{P})$, or, simply $\mathscr{A}$.

Claim 5 $\mathscr{A}(\mathscr{F}) \subseteq \mathscr{A}$.

Proof. Let $f \in \mathscr{F}$ and $\{x_1, \ldots, x_n\} = f(S)$. Assume w.l.o.g. that

$$(x_1, \mathscr{P}) \succsim (x_2, \mathscr{P}) \succsim \ldots \succsim (x_n, \mathscr{P}).$$

Then, for all i: $\quad U_{\mathscr{P}}(\delta_{x_1}) \geq U_{\mathscr{P}}(\delta_{x_i}) \geq U_{\mathscr{P}}(\delta_{x_n})$.

Now, for all $P \in \mathscr{P}$: $\quad U_{\mathscr{P}}(P^f) = \sum_{i=1}^{n} P^f(x_i) U_{\mathscr{P}}(\delta_{x_i})$,

therefore, for all $P \in \mathscr{P}$: $\quad U_{\mathscr{P}}(\delta_{x_1}) \geq U_{\mathscr{P}}(P^f) \geq U_{\mathscr{P}}(\delta_{x_n})$,

i.e. $F^f \in \mathscr{A}$ □

We need therefore to extend the relation $\succsim_{AA}$ from $\mathscr{A}(\mathscr{F})$ to $\mathscr{A}$ in such a way that the extension still satisfies the axioms of the theorem. We define therefore the relation $\succsim'_{AA}$ on $\mathscr{A}$ by:

$$F \succsim'_{AA} G \iff F^f \succsim_{AA} F^g,$$

where f,g are such that $P^f \sim_{\mathscr{P}} F(P)$ and $P^g \sim_{\mathscr{P}} G(P)$ for all $P \in \mathscr{P}$. This relation is well-defined because of axioms 1 and 5, and, thanks to axiom 9, it extends $\succsim_{AA}$ on the whole $\mathscr{A}$. In order to verify that it preserves the properties of the original relation, it suffices to show that the mixture operation is well-behaved. Take $\alpha \in [0,1]$, $F,G \in \mathscr{A}$. Then, there exist $f,g \in \mathscr{F}$ s.t. $P^f \sim_{\mathscr{P}} F(P)$ and $P^g \sim_{\mathscr{P}} G(P)$ for all $P \in \mathscr{P}$ and $A \in \Sigma$ such that $A \perp_{\mathscr{P}} \{f,g\}$ and $P(A) = \alpha$ for all $P \in \mathscr{P}$. Then, because $\succsim_{AA}$ satisfies independence on $\Delta(X)$, we have, for all $P \in \mathscr{P}$:

$$(\alpha F + (1-\alpha)G)(P) = \alpha F(P) + (1-\alpha)G(P) \sim_{AA} \alpha P^f + (1-\alpha)P^g.$$

Therefore: $$P^{fAg}(P) \sim_{AA} \alpha F(P) + (1-\alpha)G(P) \quad \forall P \in \mathscr{P}.$$

This implies that

$$\alpha F + (1-\alpha)H \succsim'_{AA} \alpha G + (1-\alpha)H \quad \Longleftrightarrow \quad (fAh, \mathscr{P}) \succsim (gAh, \mathscr{P})$$

for appropriate $f,g,h \in \mathscr{F}$ and $A \in \Sigma$.

We are now in the position to apply Schmeidler's result. The latter implies that there exists a unique capacity $\nu^{\mathscr{P}} : 2^{\mathscr{P}} \to [0,1]$ such that:

$$F \succsim'_{AA} G \quad \Longleftrightarrow \quad \int_{\mathscr{P}} U_{\mathscr{P}}(F(P)) \, \mathrm{d}\nu^{\mathscr{P}}(P) \geq \int_{\mathscr{P}} U_{\mathscr{P}}(G(P)) \, \mathrm{d}\nu^{\mathscr{P}}(P).$$

In particular, restricting this result to $\mathscr{A}(\mathscr{F})$, we have, for any $f,g \in \mathscr{F}$,

$$(f,\mathscr{P}) \succsim (g,\mathscr{P}) \quad \Longleftrightarrow \quad F^f \succsim^{\mathscr{P}}_{AA} F^g \quad \Longleftrightarrow \quad V(f,\mathscr{P}) \geq V(g,\mathscr{P}),$$

where $$V(f,\mathscr{P}) = \int_{\mathscr{P}} U_{\mathscr{P}}(P^f) \, \mathrm{d}\nu^{\mathscr{P}}(P).$$

Step 4. Let $u_{\mathscr{P}} : X \to \mathbb{R}$ be defined by $u_{\mathscr{P}}(x) = U_{\mathscr{P}}(\delta_x)$, where δ_x denote the degenerate measure with support $\{x\}$. By axiom 7, for all $x,y \in X$, for all $\mathscr{P}, \mathscr{P}' \in \mathfrak{P}$: $(x,\mathscr{P}) \succsim (y,\mathscr{P}) \iff (x,\mathscr{P}') \succsim (y,\mathscr{P}')$.

But this is equivalent to: $\delta_x \succsim_{\mathscr{P}} \delta_y \iff \delta_x \succsim_{\mathscr{P}'} \delta_y$,

and to $\delta_x \succsim^{\mathscr{P}}_{AA} \delta_y \iff \delta_x \succsim^{\mathscr{P}'}_{AA} \delta_y$.

Therefore, $u_{\mathscr{P}}$ and $u_{\mathscr{P}'}$ represent the same ordering on X, so we can normalize them so that $u_{\mathscr{P}} = u$ for all $\mathscr{P}$. Let $(f,\mathscr{P})$ and $(g,\mathscr{P}')$ be two act-information pairs. Let $x_f \in X$ be the $\mathscr{P}$-certainty equivalent of f and $x_g \in X$ be the $\mathscr{P}'$-certainty equivalent of g, i.e. $(f,\mathscr{P}) \sim (x_f,\mathscr{P})$ and

$(g,\mathscr{P}') \sim (x_g,\mathscr{P}')$. They exist by axiom 8. We have:

$$\begin{aligned}(f,\mathscr{P}) \succsim (g,\mathscr{P}') &\iff (x_f,\mathscr{P}) \succsim (x_g,\mathscr{P}') \\ &\iff (x_f,\mathscr{P}) \succsim (x_g,\mathscr{P}) \\ &\iff u(x_f) \geq u(x_g).\end{aligned}$$

But, on the other hand, $u(x_f) = V(f,\mathscr{P}) = \int_{\mathscr{P}} U_{\mathscr{P}}(P^f)\ \mathrm{d}\nu^{\mathscr{P}}(P)$ i.e.

$$u(x_f) = \int_{\mathscr{P}} \int_X u\ \mathrm{d}P^f\ \mathrm{d}\nu^{\mathscr{P}}(P) = \int_{\mathscr{P}} \int_S u \circ f\ \mathrm{d}P\ \mathrm{d}\nu^{\mathscr{P}}(P)$$

because $U_{\mathscr{P}}$ is affine and P^f has finite support. As the same holds for g, this completes the existence proof.

Step 5. $\nu^{\mathscr{P}}$ is unique as part of the representation of $\succsim'_{AA}$, but this does not automatically guarantee its uniqueness as part of the representation of $\succsim$. We will now prove it is nevertheless the case. To see that, take $\nu^{\mathscr{P}}$ and $\nu_1^{\mathscr{P}}$ two capacities representing $\succsim$. Let $B(\mathscr{F},\mathscr{P}) := \{U_{\mathscr{P}} \circ F^f \mid f \in \mathscr{F}\}$.

As usual, we drop the reference to $\mathscr{P}$ and write $B(\mathscr{F})$ and U. Define two functionals I and I_1 on $B(\mathscr{F})$ by

$$I(U \circ F^f) = \int U \circ F^f \mathrm{d}\nu^{\mathscr{P}} \quad \text{and} \quad I_1(U \circ F^f) = \int U \circ F^f \mathrm{d}\nu_1^{\mathscr{P}}.$$

Then, because both functionals represent the same ordering on $B(\mathscr{F})$, there exists an increasing function $\sigma : I(B(\mathscr{F})) \to \mathbb{R}$ such that $I_1 = \sigma \circ I$.

Take now $\mathcal{P} \subseteq \mathscr{P}$ and $\pi, \pi' \in \Delta(X)$ such that $\pi \succ_{AA} \pi'$. Normalize U so that $U(\pi) = 1$ and $U(\pi') = 0$. Consider now the second-order act $F_{\mathcal{P}} = \pi\mathcal{P}\pi'$, with obvious notation, and $f_{\mathcal{P}}$ the corresponding act delivered by axiom 9. Then $\nu^{\mathscr{P}}(\mathcal{P}) = \int_{\mathscr{P}} U_{\mathscr{P}}(F_{\mathcal{P}}(P))\mathrm{d}\nu^{\mathscr{P}}(P) = I(U \circ F^{f_{\mathcal{P}}})$ and

$$\nu_1^{\mathscr{P}}(\mathcal{P}) = \int_{\mathscr{P}} U_{\mathscr{P}}(F_{\mathcal{P}}(P))\mathrm{d}\nu_1^{\mathscr{P}}(P) = I_1(U \circ F^{f_{\mathcal{P}}}) = \sigma(I(U \circ F^{f_{\mathcal{P}}})) = \sigma(\nu^{\mathscr{P}}(\mathcal{P})).$$

Therefore, if we could show that σ is affine, then there would exist $a > 0$ and b such that $\nu_1^{\mathscr{P}}(\mathcal{P}) = a\nu^{\mathscr{P}}(\mathcal{P}) + b$ for all $\mathcal{P} \subseteq \mathscr{P}$. But this would imply that $b = 0$ upon taking $\mathcal{P} = \varnothing$ and $a = 1$ upon taking $\mathcal{P} = \mathscr{P}$ and we would be done.

Let us therefore show that σ is affine. Take $t, t' \in I(B(\mathscr{F}))$ and $\alpha \in [0,1]$.

We want to show that $\sigma(\alpha t+(1-\alpha)t') = \alpha\sigma(t)+(1-\alpha)\sigma(t')$. There exist $f, f' \in \mathscr{F}$ such that $t = I(U \circ F^f)$ and $t' = I(U \circ F^{f'})$. As I is continuous as a Choquet integral and as $\Delta(X)$ is convex, there exist π and π' in $\Delta(X)$ such that $U(\pi) = I(U \circ F^f) = t$ and $U(\pi') = I(U \circ F^{f'}) = t'$. Therefore,

$$\begin{aligned}
\sigma(\alpha t + (1-\alpha)t') &= \sigma(\alpha U(\pi) + (1-\alpha)U(\pi')) \\
&= \sigma(\alpha I(U(\pi)) + (1-\alpha)I(U(\pi'))) \\
&= \sigma(I(\alpha U(\pi) + (1-\alpha)U(\pi'))) \\
&= I_1(\alpha U(\pi) + (1-\alpha)U(\pi')) \\
&= \alpha I_1(U(\pi)) + (1-\alpha)I_1(U(\pi')) \\
&= \alpha\sigma(t) + (1-\alpha)\sigma(t'), \qquad \text{i.e. } \sigma \text{ is affine.}
\end{aligned}$$

□

Proposition 2. Necessity being straightforward in view of proposition 1, we only show sufficiency.

Fix an information set $\mathscr{P} \in \mathfrak{P}$.

Let $\Lambda_{\mathscr{P}} = \{B \in \Sigma \mid P(B) = Q(B), \forall P, Q \in \mathscr{P}.\}$. For simplicity, throughout this proof we drop the subscript $\mathscr{P}$ and write ρ for $\rho^{\mathscr{P}}$ and Λ for $\Lambda_{\mathscr{P}}$. We start by a lemma.

Lemma 3 *Let $(A_i)_{i=1,\ldots,n}$ be a partition of S into events. Then there exists a partition $(B_i)_{i=1,\ldots,n}$ consisting of elements of Λ such that, for all $m \leq n$,*

$$\rho(B_1 \cup \ldots \cup B_m) = \rho(A_1 \cup \ldots \cup A_m).$$

Proof. The proof is done by induction on m.

For $m = 1$, by range-convexity of $\mathscr{P}$, there exists $B_1 \in \Sigma$ such that $P(B_1) = \rho(A_1)$ for all $P \in \mathscr{P}$. Therefore, $B_1 \in \Lambda$ and $\rho(B_1) = \rho(A_1)$.

Assume that $B_1, \ldots, B_{m-1}$ have been constructed. Then by range convexity of $\mathscr{P}$, there exists $B_m \subseteq (B_1 \cup \ldots B_{m-1})^c$ such that

$$\begin{aligned}
P(B_m) = &\frac{\rho(A_1 \cup \ldots \cup A_m) - \rho(A_1 \cup \ldots \cup A_{m-1})}{1 - \rho(A_1 \cup \ldots \cup A_{m-1})}(1 - P(B_1 \cup \ldots \cup B_{m-1})) \\
= &\rho(A_1 \cup \ldots \cup A_m) - \rho(A_1 \cup \ldots \cup A_{m-1})
\end{aligned}$$

by definition of $B_1, \ldots, B_{m-1}$. Therefore, $B_m \in \Lambda$, $B_1 \cap \ldots \cap B_m = \varnothing$ and

$$\begin{aligned}\rho(B_1 \cup \ldots \cup B_m) &= P(B_1 \cup \ldots \cup B_m)\\ &= P(B_1 \cup \ldots \cup B_{m-1}) + \rho(A_1 \cup \ldots \cup A_m) - \rho(A_1 \cup \ldots \cup A_{m-1})\\ &= \rho(A_1 \cup \ldots \cup A_m).\end{aligned}$$

□

Consider now an act f. There exists a partition of events $(A_i)_{i=1,\ldots,n}$ and a finite sequence of outcomes $(x_i)_{i=1,\ldots,n}$ such that $(x_1, \mathscr{P}) \succsim (x_2, \mathscr{P}) \succsim \ldots \succsim (x_n, \mathscr{P})$ and $f(A_i) = \{x_i\}$. We write $f = (x_1, A_1, \ldots, x_n, A_n)$. Consider the act $g = (x_1, B_1, \ldots, x_n, B_n)$ with $(B_i)_{i=1,\ldots,n}$ as in the previous lemma. Then, by axiom 10, $(f, \mathscr{P}) \sim (g, \mathscr{P})$. Moreover, $V(g, \mathscr{P}) = \int_S u \circ g \mathrm{d}P$ for some $P \in \mathscr{P}$, therefore we have, as $V(f, \mathscr{P}) = V(g, \mathscr{P})$,

$$\begin{aligned}V(f, \mathscr{P}) = \sum_{m=1}^{n} u(x_m)P(B_m) &= \sum_{m=1}^{n} u(x_m)(\rho(A_1 \cup \ldots \cup A_m) - \rho(A_1 \cup \ldots \cup A_{m-1}))\\ &= \int_S u \circ f \mathrm{d}\rho.\end{aligned}$$

□

Proposition 3. Take $(f, \mathscr{P}) \in \mathscr{F} \times \mathfrak{P}$. Let $pc(\mathscr{P})$ be the set of probability charges on $2^{\mathscr{P}}$. The *core* of $\nu^{\mathscr{P}}$, that we shall denote $M(\mathscr{P})$, is defined by:

$$M(\mathscr{P}) = \{m \in pc(\mathscr{P}) \mid m(\mathcal{P}) \geq \nu^{\mathscr{P}}(\mathcal{P}),\ \forall \mathcal{P} \subseteq \mathscr{P}\}.$$

Because $\nu^{\mathscr{P}}$ is convex, $M(\mathscr{P}) \neq \varnothing$ and we have that:

$$\begin{aligned}V(f, \mathscr{P}) &= \int_{\mathscr{P}} \left(\int_S u \circ f \mathrm{d}P \right) \mathrm{d}\nu^{\mathscr{P}}(P)\\ &= \min_{m \in M(\mathscr{P})} \int_{\mathscr{P}} \left(\int_S u \circ f \mathrm{d}P \right) \mathrm{d}m(P)\\ &= \min_{m \in M(\mathscr{P})} \int_S u \circ f \mathrm{d}\rho^m,\end{aligned}$$

where, for all $A \in \Sigma$, for all $m \in M(\mathscr{P})$: $\qquad \rho^m(A) = \int_{\mathscr{P}} P(A)\mathrm{d}m(P).$

Setting $\Phi(\mathscr{P}) := \{\rho^m \mid m \in M(\mathscr{P})\}$ delivers the sought representation. Clearly $\Phi(\mathscr{P}) \subseteq \overline{\mathrm{co}}(\mathscr{P})$ and it is convex as $M(\mathscr{P})$ is. Let us show that it is weak* closed. Take a sequence (ρ^{m_n}) of elements of $\Phi(\mathscr{P})$ that converges to $\rho \in pc(\Sigma)$. Then, for all $A \in \Sigma$, $\rho^{m_n}(A) \to \rho(A)$, i.e. $\int_{\mathscr{P}} P(A)\mathrm{d}m_n \to \rho(A)$. But, as $M(\mathscr{P})$ is compact, there exists a subsequence (m_{n_k}) that converges to $m \in M(\mathscr{P})$. Therefore, $\int_{\mathscr{P}} P(A)\mathrm{d}m_{n_k} \to \int_{\mathscr{P}} P(A)\mathrm{d}m$. But the mother sequence has to converge to the same limit, so that $\rho(A) = \int_{\mathscr{P}} P(A)\mathrm{d}m$: $\rho \in \Phi(\mathscr{P})$.

Let us now show that $\Phi(\mathscr{P})$ is the unique weak* compact and convex set that represents V in the sense of equation 5. Suppose there is another one $\Psi(\mathscr{P})$ and that there exists $\rho_0 \in \Phi(\mathscr{P}) \setminus \Psi(\mathscr{P})$. Then, by a separation theorem, there exists a simple real-valued function $\varphi : S \to \mathbb{R}$ such that

$$\int_S \varphi \mathrm{d}\rho_0 < \min_{\rho \in \Psi(\mathscr{P})} \int_S \varphi \mathrm{d}\rho.$$

By renormalizing if necessary, because φ is bounded it is w.l.o.g. to assume that $\varphi(S) \subseteq u(X)$, and, because $u(X)$ is an interval, there exists $f \in \mathscr{F}$ such that $\varphi = u \circ f$. Therefore,

$$V(f, \mathscr{P}) = \min_{\rho \in \Phi(\mathscr{P})} \int_S u \circ f \mathrm{d}\rho < \min_{\rho \in \Psi(\mathscr{P})} \int_S u \circ f \mathrm{d}\rho = V(f, \mathscr{P}),$$

a contradiction. Therefore $\Phi(\mathscr{P}) = \Psi(\mathscr{P})$.

Let us now show the mixture linearity property of Φ. Take $\mathscr{P}, \mathscr{P}' \in \mathfrak{P}_c$ and $\lambda \in [0, 1]$. Then, we have that

$$\begin{aligned}
\min_{\rho'' \in \Phi(\lambda\mathscr{P} + (1-\lambda)\mathscr{P}')} \int_S u \circ f \mathrm{d}\rho'' &= V(f, \lambda\mathscr{P} + (1-\lambda)\mathscr{P}') \\
&= \lambda V(f, \mathscr{P}) + (1-\lambda)V(f, \mathscr{P}') \\
&= \lambda \min_{\rho \in \Phi(\mathscr{P})} \int_S u \circ f \mathrm{d}\rho + (1-\lambda) \min_{\rho' \in \Phi(\mathscr{P}')} \int_S u \circ f \mathrm{d}\rho' \\
&= \min_{\rho'' \in \lambda\Phi(\mathscr{P}) + (1-\lambda)\Phi(\mathscr{P}')} \int_S u \circ f \mathrm{d}\rho''.
\end{aligned}$$

By uniqueness, this yields $\Phi(\lambda\mathscr{P} + (1-\lambda)\mathscr{P}') = \lambda\Phi(\mathscr{P}) + (1-\lambda)\Phi(\mathscr{P}')$. $\square$

Proposition 4. Necessity is trivial. Let us prove sufficiency. Since $\mathscr{P}$ and $\mathscr{P}'$ are closed and convex, for each $f \in \mathscr{F}$, there exist $P \in \mathscr{P}$ and $P' \in \mathscr{P}$ such that $(f, \mathscr{P}) \sim (f, \{P\})$ and $(f, \mathscr{P}') \sim (f, \{P'\})$. Two applications of axiom 12 and an application of axiom 1 therefore imply that

$$(f, \lambda\mathscr{P} + (1-\lambda)\mathscr{P}') \sim (f, \lambda\{P\} + (1-\lambda)\{P'\}).$$

Therefore

$$\begin{aligned} V(f, \lambda\mathscr{P} + (1-\lambda)\mathscr{P}') &= V(f, \lambda\{P\} + (1-\lambda)\{P'\}) \\ &= \int u \circ f \mathrm{d}(\lambda P + (1-\lambda)P') \\ &= \lambda \int u \circ f \mathrm{d}P + (1-\lambda) \int u \circ f \mathrm{d}P' \\ &= \lambda V(f, \{P\}) + (1-\lambda) V(f, \{P'\}) \\ &= \lambda V(f, \mathscr{P}) + (1-\lambda) V(f, \mathscr{P}'). \qquad \square \end{aligned}$$

Proposition 5. As a preliminary step, we report a version of the representation theorem in Chateauneuf, Eichberger, and Grant (2006) for the Anscombe and Aumann setup. Let $\mathscr{P} \in \mathfrak{P}_4$ and consider, as in the proof of the main theorem, the set $\mathscr{A}$ of bounded mappings $F : \mathscr{P} \to \Delta(X)$ endowed with the preference relation $\succsim'_{AA}$. Let $F \in \mathscr{A}$ and $\pi \in \Delta(X)$. π is in the indifference set of the infimum of F, denoted $\mathrm{Inf}(F)$, if

(i) $\pi \mathcal{P} F \sim'_{AA} F$, where $\mathcal{P} = \{P \in \mathscr{P} \mid \pi \succ_{\mathscr{P}} F(P)\}$,

(ii) For all $\pi' \succ_{AA} \pi$, $\pi' \mathcal{Q} F \succ F$, where, $\mathcal{Q} = \{P \in \mathscr{P} \mid \pi' \succsim_{\mathscr{P}} F(P)\}$.

The indifference set of the supremum of $(f, \mathscr{P})$, $\mathrm{Sup}(f, \mathscr{P})$, is defined similarly by reversing all inequalities. Let

$$\underline{\mathscr{A}}(F) := \{H \in \mathscr{A} \mid \exists P \in \mathscr{P}, F(P) \precsim_{\mathscr{P}} \mathrm{Inf}(F) \text{ and } H(H) \precsim_{\mathscr{P}} \mathrm{Inf}(H)\},$$

and

$$\overline{\mathscr{A}}(F) := \{H \in \mathscr{H} \mid \exists P \in \mathscr{P}, F(P) \succsim_{\mathscr{P}} \mathrm{Sup}(F) \text{ and } H(P) \succsim_{\mathscr{P}} \mathrm{Sup}(H)\}$$

Consider the following axiom

AA 1 *Let $F, G, H \in \mathscr{A}$ be such that $F \sim'_{AA} G$ and $H \in \underline{\mathscr{A}}(G) \cap \overline{\mathscr{A}}(G)$. Then*

1. *If $H \in \underline{\mathscr{A}}(F)$, then $\alpha G + (1-\alpha)H \succsim'_{AA} \alpha F + (1-\alpha)H$,*
2. *If $H \in \overline{\mathscr{A}}(F)$, then $\alpha F + (1-\alpha)H \succsim'_{AA} \alpha G + (1-\alpha)H$.*

From the proof of the representation theorem in Chateauneuf, Eichberger, and Grant (2006), it can be inferred the following result:

Lemma 4 *Assume $\succsim'_{AA}$ has a CEU representation on $\mathscr{A}$ with capacity $\nu^{\mathscr{P}}$ such that for all $\mathscr{P} \in \mathfrak{P}_4$, $\nu^{\mathscr{P}}(\mathcal{P}) = 0 \implies \mathcal{P} = \varnothing$, for all $\mathcal{P} \subseteq \mathscr{P}$. Then $\nu^{\mathscr{P}}$ is neo-additive if and only if $\succsim'_{AA}$ satisfies axiom AA1.*

In view of this lemma, necessity of the axiom is easily proved, so that the only thing that remains to be shown is that if $\succsim$ satisfies axiom 13, then $\succsim'_{AA}$ as defined in the proof of the representation theorem satisfies axiom AA1. Let therefore $F, G, H \in \mathscr{A}$ be such that $F \sim'_{AA} G$ and $H \in \underline{\mathscr{A}}(G) \cap \overline{\mathscr{A}}(G)$. By axiom 9, there exist $f, g, h \in \mathscr{F}$ such that $P^f \sim_{AA} F(P)$, $P^g \sim_{AA} G(P)$ and $P^h \sim_{AA} H(P)$ for all $P \in \mathscr{P}$. Then $(f, \mathscr{P}) \sim (g, \mathscr{P})$. Assume $H \in \underline{\mathscr{A}}(F)$. Then, $F^h \in \underline{\mathscr{A}}(F)$ by construction, as all that matters for extreme events are indifference classes. But $\underline{\mathscr{A}}(F) = \underline{\mathscr{A}}(F^f)$ and therefore $h \in \underline{\mathscr{F}}(f, \mathscr{P})$ also by construction. Similarly, $H \in \underline{\mathscr{A}}(G) \cap \overline{\mathscr{A}}(G)$ if and only if $h \in \underline{\mathscr{F}}(g, \mathscr{P}) \cap \overline{\mathscr{F}}(g, \mathscr{P})$. Therefore, axiom 13 implies that for all $A \perp_{\mathscr{P}} \{f, g, h\}$, $(gAh, \mathscr{P}) \succsim (fAh, \mathscr{P})$. But, by construction, this is equivalent to $\alpha G + (1-\alpha)H \succsim'_{AA} \alpha F + (1-\alpha)H$. A similar proof can be given for the second part of the axiom. □

Proposition 6. If the decision-maker is averse to imprecision, then, for all $P \in \mathscr{P}$,

$$(f, \{P\}) \succsim (f, \mathscr{P}).$$

Therefore,

$$\int_S u \circ f \mathrm{d}P \geq V(f, \mathscr{P}) \quad \forall P \in \mathscr{P},$$

and thus

$$\inf_{P \in \mathscr{P}} \int_S u \circ f \mathrm{d}P \geq V(f, \mathscr{P}).$$

Conversely,

$$\inf_{P \in \mathscr{P}} \int_S u \circ f \mathrm{d}P \leq \int_S u \circ f \mathrm{d}Q \quad \forall Q \in \mathscr{P}.$$

Thus, integrating both sides with respect to $\nu^{\mathscr{P}}$, we get:

$$\inf_{P \in \mathscr{P}} \int_S u \circ f \mathrm{d}P \leq V(f, \mathscr{P}),$$

as wanted. The converse is trivial. □

Proposition 7. Fix $\mathscr{P} \in \mathfrak{P}$. Notice first that our definition of comparative ambiguity aversion is equivalent to the definition in Ghirardato and Marinacci (2002) when translated to the set of Anscombe-Aumann acts $\mathscr{A}(\mathscr{F})$ introduced in the proof of the main theorem. Therefore, as $\Delta(X) \subset \mathscr{A}(\mathscr{F})$, if decision maker 1 is more ambiguity averse than decision maker 2 given information $\mathscr{P}$, their preference relations coincide on $\Delta(X)$.

In order to apply Ghirardato and Marinacci (2002)'s theorem 17, we need however to show that the associated extensions $\succsim'_{AA,1}$ and $\succsim'_{AA,2}$ bear the same relations in terms of comparative ambiguity attitude as relations $\succsim_{AA,1}$ and $\succsim_{AA,2}$. Clearly, if $\succsim'_{AA,1}$ is more ambiguity averse than $\succsim'_{AA,2}$, then $\succsim_{AA,1}$ is more ambiguity averse than $\succsim_{AA,2}$. To show the converse, take $F \in \mathscr{A}$ and $\pi \in \Delta(X)$ such that $F \succsim'_{AA,1} \pi$. Then, by definition, $F^f \succsim_{AA,1} F^k$ for some f such that $P^f \sim_1 F(P)$ for all $P \in \mathscr{P}$ that exist by axiom 9 and some $k \in K(\pi, \mathscr{P})$ that exist by lemma 2. Then, $(f, \mathscr{P}) \succsim_1 (k, \mathscr{P})$, so that, by assumption, $f \succsim_2 k$, and therefore, $F^f \succsim_{AA,2} F^k$. But, as $\succsim_{AA,1}$ and $\succsim_{AA,2}$ coincide on $\Delta(X)$, this implies that $F \succsim'_{AA,2} \pi$, as wanted. The other implication is proved similarly.

By theorem 17 of Ghirardato and Marinacci (2002), therefore, we know that $\succsim_1$ is more ambiguity averse than $\succsim_2$ if and only if $\nu^2_{\mathscr{P}} \geq \nu^1_{\mathscr{P}}$. □

Proposition 8. Necessity is straightforward. We prove sufficiency. For simplicity we drop the superscript $\mathscr{P}$. Suppose there exists $A \in \Sigma$ such that $\rho_1(A) > \rho_2(A)$. By definition of ρ,

$$\inf_{\mathscr{P}} P(A) \leq \rho(A) \leq \sup_{\mathscr{P}} P(A).$$

Therefore,

$$\inf_{\mathscr{P}} P(A) \leq \rho_2(A) < \rho_1(A) \leq \sup_{\mathscr{P}} P(A).$$

We have to distinguish cases:

Case 1 : $\inf_{\mathscr{P}} P(A) < \rho_2(A) < \rho_1(A) < \sup_{\mathscr{P}} P(A)$. In this case, there exists P_* and P^* in $\mathscr{P}$ such that $\inf\limits_{\mathscr{P}} P(A) \leq P_*(A) < \rho_2(A) < \rho_1(A) < P^*(A) \leq \sup\limits_{\mathscr{P}} P(A)$. Therefore, there exists $\alpha \in (0,1)$ and $P = \alpha P^* + (1-\alpha)P_* \in co(\mathscr{P})$ such that

$$\rho_2(A) < P(A) < \rho_1(A).$$

But this contradicts the definition of imprecision aversion.

Case 2 : $\inf_{\mathscr{P}} P(A) = \rho_2(A) < \rho_1(A) < \sup_{\mathscr{P}} P(A)$. Then there exists $P \in \mathscr{P}$ such that

$$\rho_2(A) \leq P(A) < \rho_1(A),$$

again contradicting the definition of imprecision aversion.

Case 3 : $\inf_{\mathscr{P}} P(A) < \rho_2(A) < \rho_1(A) = \sup_{\mathscr{P}} P(A)$. Similar.

Case 4 : $\inf_{\mathscr{P}} P(A) = \rho_2(A) < \rho_1(A) = \sup_{\mathscr{P}} P(A)$. Then there exists $P \in co(\mathscr{P})$ such that

$$\rho_2(A) < P(A) < \rho_1(A),$$

again contradicting the definition of imprecision aversion. □

References

AHN, D. S. (2003, revised 2005): "Ambiguity Without a State-Space," mimeo,UC Berkeley.

AMARANTE, M. (2006): "(Choquet-)Integrating over Priors: α(f)-MEU," Columbia University.

ANDERSON, E., L. HANSEN, AND T. SARGENT (1999): "Robustness, Detection and the Price of Risk," University of Chicago.

ANSCOMBE, F., AND R. AUMANN (1963): "A Definition of Subjective Probability," *Annals of Mathematical Statistics*, 34, 199–205.

ARROW, K., AND L. HURWICZ (1972): "An optimality criterion for decision-making under ignorance.," in *Uncertainty and Expectations in Economics*, ed. by C. F. Carter, and J. Ford. Basil Blackwell & Mott Ltd., Oxford, England.

CHATEAUNEUF, A., M. COHEN, AND J.-Y. JAFFRAY (2006): "Décision dans l'incertain: les modèles classiques," in *Concepts et Méthodes pour l'aide à la décision*, ed. by D. Bouyssou, D. Dubois, M. Pirlot, and H. Prade, vol. 2, p. ? Hermes.

CHATEAUNEUF, A., J. EICHBERGER, AND S. GRANT (2006): "Choice under uncertainty with the best and worst in mind: neo-additive capacities," mimeo.

COHEN, M. (1992): "Security Level, Potential Level, Expected Utility: a Three-Criteria Decision Model Under Risk," *Theory and Decision*, 33, 101–134.

ELLSBERG, D. (1961): "Risk, ambiguity, and the Savage axioms," *Quartely Journal of Economics*, 75, 643–669.

EPSTEIN, L. G. (1999): "A Definition of Uncertainty Aversion," *Review of Economic Studies*, 66, 579–608.

GAJDOS, T., T. HAYASHI, J.-M. TALLON, AND J.-C. VERGNAUD (2006): "Attitude toward Imprecise Information," Discussion paper, Centre d'Economie de la Sorbonne, Université Paris I- Department of Economics, University of Texas at Austin.

GAJDOS, T., J.-M. TALLON, AND J.-C. VERGNAUD (2004): "Decision making with imprecise probabilistic information," *Journal of Mathematical Economics*, 40(6), 647–681.

Ghirardato, P., F. Maccheroni, and M. Marinacci (2001): "Differentiating Ambiguity and Ambiguity Attitude," *Forthcoming in Journal of Economic Theory.*

Ghirardato, P., and M. Marinacci (2001): "Risk, ambiguity, and the separation of utility and beliefs," *Mathematics of Operations Research*, 26, 864–890.

——— (2002): "Ambiguity Made Precise: A Comparative Foundation," *Journal of Economic Theory*, 102(2), 251–289.

Gilboa, I. (1988): "A Combination of Expected Utility and Maxmin Decision Criteria," *Journal of Mathematical Psychology*, 32, 405–420.

Gilboa, I., and D. Schmeidler (1989): "Maxmin expected utility with a non-unique prior," *Journal of Mathematical Economics*, 18, 141–153.

Gilboa, I., and D. Schmeidler (1995): "Canonical representations of set functions," *Mathematics of Operations Research*, 20, 197–212.

Jaffray, J.-Y. (1988): "Choice under Risk and the Security Factor: An Axiomatic Model," *Theory and Decision*, 24, 169–200.

——— (1989): "Généralisation du critère de l'utilité à l'incertain régulier.," *Operations Research/ Recherche Opérationnelle*, 23(3), 237–267.

Klibanoff, P., M. Marinacci, and S. Mukerji (2005): "A Smooth Model of Decision Making under Ambiguity," *Econometrica*, 73(6), 1849–1892.

Kopylov, I. (2006): "A Parametric Model of Ambiguity Hedging," University of California at Irvine.

Lopes, L. L. (1987): "Between Hope and Fear: The psychology of risk," *Advances in Experimental Social Psychology*, 20, 255–295.

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

Nehring, K. (1994): "On the Interpretation of Sarin and Wakker's 'A Simple Axiomatization of Nonadditive Expected Utility'," *Econometrica*, 62(4), 935–38.

——— (2001): "Ambiguity in the Context of Probabilistic Beliefs," mimeo, University of California, Davis.

——— (2002): "Imprecise Probabilistic Beliefs," mimeo.

Olszewski, W. (2002, revised 2006): "Preferences Over Sets of Lotteries," University of Northwestern.

Rustichini, A. (1992): "Decision Theory with Higher Order Beliefs," in *Proceedings of TARK IV.*

Sarin, R., and P. Wakker (1992): "A Simple Axiomatization of Nonadditive Expected Utility," *Econometrica*, 60, 1255–1272.

Savage, L. J. (1954): *The Foundations of Statistics.* Wiley, New York.

Schmeidler, D. (1989): "Subjective probability and expected utility without additivity," *Econometrica*, 57, 571– 587.

Seo, K. (2006): "Ambiguity and Second Order Beliefs," mimeo,University of Rochester.

Stinchcombe, M. (2003): "Choice and games with ambiguity as sets of probabilities," mimeo,University of Texas, Austin.

Tapking, J. (2004): "Axioms for preferences revealing subjective uncertainty and uncertainty aversion," *Journal of Mathematical Economics*, 40, 771–797.

Wang, T. (2001, revised 2003): "A Class of Multi-Priors Preferences," University of British Columbia.

Yaari, M. (1987): "The dual theory of choice under risk," *Econometrica*, 55, 95–115.