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

Raphaël Giraud
EUREQua-Université Paris I,106-112 Bd de l'Hôpital 75005 PARIS FRANCE
THEMA-Université de Cergy-Pontoise
Raphael.Giraud@univ-paris1.fr

Abstract

We axiomatize a model of decision under objective ambiguity described by multiple probability distributions. 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 properties of the model are examined.

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

1 Introduction

Ambiguity and ambiguity aversion have become the center of attention in the last twenty years in decision theory. This interest has grown from the challenge levelled 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.

Now, ambiguity aversion may be interpreted as a form of pessimism: when favorable and unfavorable scenarios 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.

The approach we take to that effect is similar to the approaches of Gajdos, Tallon, and Vergnaud (2004) and Nehring (2002, 2001). We assume that decision makers have preferences on pairs $(f, \mathscr{P})$, where f is a Savagean act 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?

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:

$$V_{\mathscr{P}}(f) = \int_{\mathscr{P}} \int_S u(f(s)) \, \mathrm{d}P(s) \, \mathrm{d}\nu^{\mathscr{P}}(P),$$

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

Our model can be viewed as dual to the model in Klibanoff, Marinacci, and Mukerji (2003). The criterion axiomatized in this paper is the following:

$$V_{\mathscr{P}}^{KKM}(f) = \int_{\mathscr{P}} \varphi(\int_S u(f(s)) \, \mathrm{d}P(s)) \, \mathrm{d}\pi^{\mathscr{P}}(P),$$

where $\pi^{\mathscr{P}}$ is a probability measure. Thus, 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 Klibanoff et al. (2003)'s model in the sense that our φ is linear and our $\pi^{\mathscr{P}}$ is non-additive.

In our view, the model we propose has the advantage over Klibanoff et al.'s model of axiomatizing in a rather intuitive and simple way, 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. This is also done by Klibanoff et al., but they make use of the notion of second order acts and of preferences over these second order acts to generate the second order beliefs. We object to this notion because second order acts and preferences on them are not primitives of the problem in a real decision problem, they are derived constructions. Generally speaking, Klibanoff et al.'s model is a reconstruction of a more abstract decision problem that seems to us too far away from the natural syntax of a decision problem in real life.

Our approach is very much indebted in its inspiration by the reconsideration of the Ellsberg paradox to be found in Chateauneuf, Cohen, and Jaffray (forthcoming)[1]. These authors show how the Ellsberg paradox, 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 not only Savage's Sure Thing Principle is violated by this paradox, but also the Independence Axiom. Here we propose to generalize this approach in order to connect the Savage framework and the Anscombe-Aumann framework. This is also done in Nehring (2002), but in this paper the Anscombe-Aumann acts are defined on the same state space as Savage acts, whereas here they will be defined (in the proof of the theorem) on a different state space.

Most importantly, in line with Chateauneuf et al., we show 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. However we leave that for future research. In this paper we explore some simple properties of the criterion axiomatized.

The paper is organized as follows: section 2 introduces the set up and the axioms, in section 3 the main theorem is stated and the main properties of the functional are studied. Section 4 concludes, while the theorem is proved in section 5.

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 $(X, \mathscr{B})$. We denote by $\mathscr{F}$ the set of Savage acts, i.e. the set of finite-valued Σ-measurable functions f from S to X. Let $\mathscr{P}(S)$ be the set of all countably additive probability measures on Σ. An element $P \in \mathscr{P}(S)$ is said to be *convex-ranged* if for all $\alpha \in [0,1]$, for all $A \in \Sigma$, there exists $B \in \Sigma$ such that $P(B) = \alpha P(A)$. Let $\mathscr{P}_c(S) \subseteq \mathscr{P}(S)$ be the set of convex-ranged countably additive probability measures. We denote by $\mathfrak{P}_0$ the set of all non-empty finite subsets $\mathscr{P}$ of $\mathscr{P}_c(S)$ having the following separation property:

Property 1 (Separation property)**.** *For all $D \subseteq \mathscr{P}$, there exists $A_D \in \Sigma$ such that $P(A_D) = 1$ for all $P \in D$ and $P(A_D) = 0$ for all $P \notin D$.*

Following Gajdos et al. (2004), the objects of choice in our setting will be pairs $(f, \mathscr{P})$ in $\mathscr{F} \times \mathfrak{P}_0$. 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 $\mathscr{F} \times \mathfrak{P}_0$ and are represented by the relation $\succsim$.

[1] It seems however that the original idea was Schmeidler's (personal communication by J.-Y. Jaffray).

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 one makes a deal in a country, the probability that the deal will be enforced depends in particular on the legal system. It is more imprecise if the decision maker does not know the country well. 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 informations.

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 $s \notin A$. 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 $P^f \in \Delta(X)$.

Axiom 2 (Information-Contingent Continuity). *For all $\mathscr{P} \in \mathfrak{P}_0$, 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) $(fAh, \mathscr{P}) \succ (g, \mathscr{P}) \succ (fBh, \mathscr{P})$,

(ii) For all $P \in \mathscr{P}$, $P^{fAh} = \alpha P^f + (1-\alpha)P^h$ and $P^{fBh} = \beta P^f + (1-\beta)P^h$

In order to state the next axiom, we need the following lemma:

Lemma 1. *For all $\mathscr{P} \in \mathfrak{P}_0$, for all $\pi \in \Delta(X)$, there exists $k \in \mathscr{F}$ such that, $P^k = \pi$ for all $P \in \mathscr{P}$.*

Proof. As a preliminary remark, it must be noticed that, as a consequence of the Lyapunov convexity theorem the set $\mathscr{P}$ is convex-ranged in the sense of Nehring (2002), i.e. for all $\alpha \in [0, 1]$, for all $A \in \Sigma$, there exists $B \in \Sigma$ such that $P(B) = \alpha P(A)$ for all $P \in \mathscr{P}$.

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 convex rangedness 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 $\pi(x_1)\delta_{x_1} + (1 - \pi(x_1))\pi^{x_1} = \pi$. □

Let

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

Intuitively, acts in $\Lambda(\pi, \mathscr{P})$ differ only by the permutation of outcomes on events of equal probability, a manipulation that amounts to relabelling these events. From a normative point of view, such a relabelling should note affect the decision maker's preference, because this would amount to a framing effect. This is the intuition that motivates the next axiom:

Axiom 3 (No Framing Effect). *For all $\pi \in \Delta(X)$, for all $\mathscr{P} \in \mathfrak{P}_0$, for all $k, k' \in \Lambda(\pi, \mathscr{P})$, $(k, \mathscr{P}) \sim (k', \mathscr{P})$.*

In view of this axiom, we will abuse notation and write

$$\Lambda(\pi, \mathscr{P}) \succsim \Lambda(\pi', \mathscr{P})$$

for: there exists (and thus, for all) $k \in \Lambda(\pi, \mathscr{P})$, $k' \in \Lambda(\pi', \mathscr{P})$, $k \succsim k'$.

Definition 1. *Let $f, g \in \mathscr{F}$ and $\mathscr{P} \in \mathfrak{P}_0$. We shall say that f and g are $\mathscr{P}$-comonotonic if, for all $P, Q \in \mathscr{P}$,*

$$\Lambda(P^g, \mathscr{P}) \succsim \Lambda(Q^g, \mathscr{P})$$

whenever

$$\Lambda(P^f, \mathscr{P}) \succ \Lambda(Q^f, \mathscr{P}).$$

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:

$$P \succsim^f Q \Leftrightarrow \Lambda(P^f, \mathscr{P}) \succsim \Lambda(Q^f, \mathscr{P}).$$

This ordering answers the following question: "were I given the choice between an act that, conditional on my information, unambiguously induces the lottery P^f, and one that induces the lottery Q^f, which one would I choose?" If I choose P^f, this means that, given act f, I deem the scenario corresponding to P as more favorable than the scenario corresponding to Q. Therefore, this ordering amounts to rank scenarios according to the relative favorableness given act f. Now, two acts are $\mathscr{P}$-comonotonic as defined if, roughly speaking, they order the scenarios in the same way. If they order scenarios in the same way, they do not provide any hedging opportunity against each other, not in the sense of compensating bad states of nature of one act with good states of the other, as in usual hedging, but of compensating bad scenarios with good scenarios.

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. This normatively appealing behaviour is what the next axiom requires:

Axiom 4 (Information-Contingent Comonotonic Independence)**.** *For all $\mathscr{P} \in \mathfrak{P}_0$, for all $f, g, h \in \mathscr{F}$ pairwise $\mathscr{P}$-comonotonic, for all $\alpha \in [0,1]$, for all $A \in \Sigma$ such that, for all $P \in \mathscr{P}$, $P^{fAh} = \alpha P^f + (1-\alpha)P^h$ and $P^{gAh} = \alpha P^g + (1-\alpha)P^h$,*

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

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

Axiom 5 (Information-Contingent Dominance)**.** *For all $f, g \in \mathscr{F}$, for all $\mathscr{P} \in \mathfrak{P}_0$, if*

$$\Lambda(P^f, \mathscr{P}) \succsim \Lambda(P^g, \mathscr{P}),\ \forall P \in \mathscr{P},$$

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

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

Axiom 6 (Non-Degeneracy)**.** *For all $\mathscr{P} \in \mathfrak{P}_0$, there exist $f, g \in \mathscr{F}$ such that $(f, \mathscr{P}) \succ (g, \mathscr{P})$.*

The next axiom is essentially technical. It 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 7 (Certainty Equivalent)**.** *For all $\mathscr{P} \in \mathfrak{P}_0$, for all $f \in \mathscr{F}$, there exists $x_f \in X$ such that $(x_f, \mathscr{P}) \sim (f, \mathscr{P})$.*

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 8 (Preferences under Certainty)**.** *For all $\mathscr{P}, \mathscr{P}' \in \mathfrak{P}_0$, for all $x \in X$, $(x, \mathscr{P}) \sim (x, \mathscr{P}')$.*

3 The Representation Theorem

3.1 Statement

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

Definition 2. *Let Ω be a finite set. A* capacity *on Ω is a function $\nu : 2^\Omega \to \mathbb{R}$ such that:*

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

(ii) For all $A \subseteq B \subseteq \Omega$, $\nu(A) \leq \nu(B)$.

Let $\varphi : \Omega \to \mathbb{R}$. There exist families $(A_i)_{i \in I}$, $A_i \subseteq \Omega$ and $(x_i)_{i \in I}$ with $x_i \in \mathbb{R}$ such that $x_1 \leq x_2 \leq \ldots \leq x_n$ and φ takes the value x_i on the set A_i. Then, the Choquet integral *of φ with respect to ν is defined by:*

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

with $A_{n+1} = \varnothing$.

We can now state the main representation theorem:

Theorem 1. *If $\succsim$ satisfies axioms 1 through 8, then, there exist a non-constant function $u : X \to \mathbb{R}$ and, for all $\mathscr{P} \in \mathfrak{P}_0$, a capacity $\nu^{\mathscr{P}}$ on $\mathscr{P}$ such that:*

$$(f, \mathscr{P}) \succsim (g, \mathscr{P}') \Leftrightarrow V_{\mathscr{P}}(f) \geq V_{\mathscr{P}'}(g),$$

where

$$V_{\mathscr{P}}(f) = \int_{\mathscr{P}} \int_S u(f(s))\ dP(s)\ d\nu^{\mathscr{P}}(P) \quad (2)$$

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

The proof of the theorem appears in section 5.

Remark 1. *All axioms except axiom 7 (certainty equivalent) are also necessary, as it is easy to show.*

3.2 Interpretation and Properties of the Representation

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 (finite) 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.

3.2.1 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 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 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

[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 $\Pi = \{\pi \in \Delta(\mathscr{P}) \mid \pi \leq \nu^*\}$.

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.

3.2.2 Subjective Beliefs and Objective Information

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.

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 $(x, \mathscr{P}) \succ (y, \mathscr{P})$*

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

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 $A, B \in \Sigma$,

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

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

Proof. This proposition is a straightforward consequence of the properties of the Choquet integral: positive homogeneity, comonotonic additivity and monotonicity. Details are left to the reader. □

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. 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.

3.2.3 Ambiguity Aversion

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.

What is however the natural definition of ambiguity aversion in our setting? In order to answer this question, we introduce, for all $\mathscr{P} \in \mathfrak{P}_0$ the notion of $\mathscr{P}$-unambiguous acts:

Definition 3. *For all $\mathscr{P} \in \mathfrak{P}_0$, $f \in \mathscr{F}$ 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 $\mathscr{F}$. Let $\mathscr{U}_{\mathscr{P}}$ be the set of $\mathscr{P}$-unambiguous acts. We can now give the following definition of comparative uncertainty aversion given information $\mathscr{P}$.

Definition 4. *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 $k \in \mathscr{U}_{\mathscr{P}}$, for all $f \in \mathscr{F}$:*

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

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

Proposition 2. *Let $\succsim_1$ and $\succsim_2$ be the preference relations of two decision makers satisfying the axioms of the theorem. Then if decision maker 1 is more ambiguity averse than decision maker 2 given information $\mathscr{P}$, then $\rho_2^{\mathscr{P}} \geq \rho_1^{\mathscr{P}}$.*

Proof. Fix $\mathscr{P} \in \mathfrak{P}_0$. 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 $\mathscr{F}^{AA}$ introduced in the proof of the main theorem (section 5). Therefore, if decision maker 1 is more ambiguity averse than decision maker 2 given information $\mathscr{P}$, they have same utility function $U_{\mathscr{P}}$, therefore same utility function u.

Now, clearly, as any constant $z \in X$ is an element of $\mathscr{U}_{\mathscr{P}}$, if decision maker 1 is more ambiguity averse than decision maker 2 given information $\mathscr{P}$, then, for all $x, y, z \in X$ such that $x \succ y$, for all $A \in \Sigma$, $xAy \succsim_1 z \Rightarrow xAy \succsim_2 z$. In particular, let x, y be such that $u(x) = 1$ and $u(y) = 0$ and let z be a certainty equivalent of xAy given $\mathscr{P}$. Then, $\rho_1^{\mathscr{P}}(A) = V_{\mathscr{P},1}(xAy) = u(z) \leq V_{\mathscr{P},2}(xAy) = \rho_2^{\mathscr{P}}(A)$. □

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.

Now, we have analyzed the behavior of a decision-maker given a fixed information. How will changing information affect his behavior? One can assume that decision makers usually prefer having precise information. Therefore, we define, in the spirit of Gajdos et al. (2004), aversion to imprecision, though in a cruder way:

Definition 5. *A decision maker is averse to imprecision if, for all $\mathscr{P}, \mathscr{P}' \in \mathfrak{P}_0$,*

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

An obvious consequence of this definition is the following:

Proposition 3. *If a decision maker satisfying the axioms of the representation theorem is averse to imprecision, then $\rho^{\mathscr{P}} \geq \rho^{\mathscr{P}'}$ whenever $\mathscr{P} \subseteq \mathscr{P}'$.*

Proof. Take $f = xAy$. □

When information is more precise, one is surer of the likelihood of events. Therefore, one is more likely to bet on them, one will pay more for a given bet. This is the import of this proposition.

4 Conclusion

We have axiomatized in what seems to us a rather simple way a model of decision making under ambiguous objective information or imprecise risk where the decision maker maximizes the average expected utility of a given act with respect to some second order belief over beliefs.

Several paths are open for future research. The first and in our opinion most important one is to characterize the case where our model reduces to Choquet Expected Utility with respect to the willingness to bet $\rho^{\mathscr{P}}$. This is of course the case when $\rho^{\mathscr{P}}$ is additive for all $\mathscr{P}$, but is it possible to generalize this? This will open the possibility to characterize in the Savage setting Choquet Expected Utility and other related models that have simple axiomatizations in the Anscombe Aumann setting in a simpler way than what is generally to be found in the literature. Second, it would be interesting to further explore the consequences of imposing some classical properties on the capacity $\nu^{\mathscr{P}}$, linked in particular to the separation property. Thirdly, we have given only a one way characterization of ambiguity aversion and aversion to imprecision. Some technical problems arise that we will have to address. Finally, one would like to use the model proposed here to describe the way the decision maker behaves when he or she faces new information, i.e. how second order-beliefs are updated in a dynamical model. This is left for further research.

5 Proof of the Representation Theorem

The proof proceeds in several steps.

- **Step 1** Fix $\mathscr{P} \in \mathfrak{P}_0$. Given $f \in \mathscr{F}$, one can canonically associate a function:

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

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

$$\mathscr{F}^{AA}_{\mathscr{P}} := \{F \in \Delta(X)^{\mathscr{P}} \mid \exists f \in \mathscr{F}, F = F^f_{\mathscr{P}}\}$$

  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 subscript $\mathscr{P}$ and write only F^f and $\mathscr{F}^{AA}$.

- **Step 2** By axioms 3 and 5, if $F^f = F^g$, then $f \sim g$. Therefore, one can define a preference relation $\succsim^{\mathscr{P}}_{AA}$ on $\mathscr{F}^{AA}$ by setting:

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

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

  By convex-rangedness 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 $P \in \mathscr{P}$,

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

  Therefore, the set $\mathscr{F}^{AA}$ is convex. We wish to show that preference $\succsim_{AA}$ on $\mathscr{F}^{AA}$ 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 lemmas.

  Lemma 2 (Continuity)**.** *For all $F^f, F^g, F^h \in \mathscr{F}^{AA}$, if $F^f \succ_{AA} F^g \succ_{AA} F^h$, then there exists $\alpha, \beta \in]0, 1[$ such that:*

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

  and

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

  Proof. This follows automatically from axiom 2. □

  The following remark will be useful:

  Remark 2. *Identifying constant AA-acts in $\mathscr{F}^{AA}$ and elements of $\Delta(X)$, one can consider the restriction to $\Delta(X)$ of the relation $\succsim^{\mathscr{P}}_{AA}$. Notice that, by lemma 1, all constant AA-acts belong to $\mathscr{F}^{AA}$. Because, for $\pi \in \Delta(X)$, we have $F^{\Lambda(\pi,\mathscr{P})} = \pi$, for all $\pi, \pi' \in \Delta(X)$, we have, as a matter of fact:*

$$\pi \succsim_{AA} \pi' \Leftrightarrow \Lambda(\pi, \mathscr{P}) \succsim \Lambda(\pi', \mathscr{P}).$$

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

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

  This formula is equivalent to:

$$P^f \succ_{AA} Q^f \Rightarrow P^g \succsim_{AA} Q^g,$$

  i.e., by the previous remark, to:

$$\Lambda(P^g, \mathscr{P}) \succsim \Lambda(Q^g, \mathscr{P})$$

  whenever

$$\Lambda(P^f, \mathscr{P}) \succ \Lambda(Q^f, \mathscr{P})$$

  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 lemma:

  Lemma 3 (Comonotonic Independence)**.** *For all $F^f, F^g, F^h \in \mathscr{F}^{AA}$ pairwise comonotonic and for all $\alpha \in [0, 1]$:*

$$F^f \succsim_{AA} F^g$$

  if and only if

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

  Proof. This follows from the previous remarks, from axiom 4 and from the fact that, by convex-rangedness of $\mathscr{P}$, there exists $A \in \Sigma$ such that $F^{fAh} = \alpha F^f + (1 - \alpha)F^h$ and $F^{gAh} = \alpha F^g + (1 - \alpha)F^h$. □

The next lemma is a direct consequence of the previous remark and of axiom 5:

Lemma 4 (Dominance)**.** *For all* $F^f, F^g \in \mathscr{F}^{AA}$*, if* $F^f(P) \succsim_{AA} F^g(P)$ *for all* $P \in \mathscr{P}$*, then* $F^f \succsim_{AA} F^g$*.*

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

Lemma 5 (non-triviality)**.** *There exists* $F^f, F^g \in \mathscr{F}^{AA}$*, such that* $F^f \succ_{AA} F^g$*.*

- **Step 3** We will now proceed to construct the objects of the theorem. First, we know that, because $\succsim_{AA}$ is a weak order, because of lemmas 2,3 and 5 and because all constant AA-acts are pairwise comonotonic and belong to $\mathscr{F}^{AA}$, restricting $\succsim_{AA}$ to constant acts allows to show 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 $\pi, \pi' \in \Delta(X)$:

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

Choose some $x_0 \in X$ and normalize $U_{\mathscr{P}}$ so that $U_{\mathscr{P}}(x_0) = 0$. Now let

$$B_0(\mathscr{P}, \mathscr{F})$$

denote the set

$$\{U_{\mathscr{P}} \circ F^f \mid f \in \mathscr{F}\}.$$

This set is a convex subset, containing 0, of the vector space $B_0(\mathscr{P})$ of all finite valued functions from $\mathscr{P}$ to $\mathbb{R}$. Because of lemma 4, it is possible to define a relation $\succsim^*$ on $B_0(\mathscr{P}, \mathscr{F})$ setting

$$U_{\mathscr{P}} \circ F^f \succsim^* U_{\mathscr{P}} \circ F^g \Leftrightarrow F^f \succsim_{AA}^{\mathscr{P}} F^g.$$

Now, enumerate $\mathscr{P} = \{P_1, \dots, P_n\}$. Let $\mathfrak{S}_n$ be the set of permutations of $\mathscr{P}$. For all $\sigma \in \mathfrak{S}_n$, let $B_0^\sigma(\mathscr{P}, \mathscr{F})$ denote the set of all functions $\varphi \in B_0(\mathscr{P}, \mathscr{F})$ such that

$$\varphi(P_{\sigma(1)}) \succsim^* \varphi(P_{\sigma(2)}) \succsim^* \dots \succsim^* \varphi(P_{\sigma(n)}).$$

This set is convex and contains all the constants in $B_0(\mathscr{P}, \mathscr{F})$, and the sets $\{B_0^\sigma(\mathscr{P}, \mathscr{F})\}_{\sigma \in \mathfrak{S}_n}$ are the maximal comonotonicity sets in $B_0(\mathscr{P}, \mathscr{F})$, i.e. they are the maximal subsets having the property that, for any $\sigma \in \mathfrak{S}_n$, for any two acts F^f and F^g belonging to $B_0^\sigma(\mathscr{P}, \mathscr{F})$, F^f and F^g are comonotonic.

Because of lemma 3, the restriction of $\succsim^*$ to any of these subsets satisfies independence, and, therefore, by lemma 2 and by the mixture space theorem, there exists a monotonic (by lemma 4) linear functional

$$I^\sigma : B_0^\sigma(\mathscr{P}, \mathscr{F}) \to \mathbb{R}$$

representing $\succsim^*$ on $B_0^\sigma(\mathscr{P}, \mathscr{F})$. This functional can be uniquely extended to a monotonic linear functional J^σ defined on the vector space $B_0^\sigma(\mathscr{P})$ spanned by $B_0^\sigma(\mathscr{P}, \mathscr{F})$. By monotonicity and non-degeneracy, there exists therefore a unique probability μ^σ on $\mathscr{P}$ such that

$$J^\sigma(\varphi) = \int_{\mathscr{P}} \varphi \, \mathrm{d}\mu^\sigma$$

for all $\varphi \in B_0^\sigma(\mathscr{P})$. Because each φ in $B_0^\sigma(\mathscr{P}, \mathscr{F})$ is bounded as it is finite valued, and because of continuity (lemma 2), it has a certainty equivalent. Therefore, because the constants (elements of $U_{\mathscr{P}}(\Delta(X))$) all lie in $\cap_{\sigma \in \mathfrak{S}_n} B_0^\sigma(\mathscr{P}, \mathscr{F})$, we have, for all $\varphi \in B_0^\sigma(\mathscr{P}, \mathscr{F})$,$\varphi' \in B_0^{\sigma'}(\mathscr{P}, \mathscr{F})$:

$$\varphi \succsim^* \varphi'$$

if and only if

$$\int_{\mathscr{P}} \varphi \, \mathrm{d}\mu^\sigma \geq \int_{\mathscr{P}} \varphi' \, \mathrm{d}\mu^{\sigma'}.$$

Now, let $D \subseteq \mathscr{P}$ and let $\mathbb{1}_D$ be the indicator function of D. By the separation property, $\mathbb{1}_D \in B_0(\mathscr{P}, \mathscr{F})$. Indeed, if we take $x, y \in X$ such that $U(x) = 1$ and $U(y) = 0$, and if we set $f = xA_Dy$, where A_D is such that $P(A_D) = 1$ for all $P \in D$ and $P(A_D) = 0$ for all $P \notin D$, then $P^f = \delta_x$ for all $P \in D$ and $P^f = \delta_y$ for all $P \notin D$. Therefore $U \circ F^f = \mathbb{1}_D$.

If

$$\mathbb{1}_D \in B_0^\sigma(\mathscr{P}, \mathscr{F}) \cap B_0^{\sigma'}(\mathscr{P}, \mathscr{F}),$$

then

$$\mu^\sigma(D) = \mu^{\sigma'}(D).$$

We can therefore define the capacity $\nu^{\mathscr{P}}$ by:

$$\nu^{\mathscr{P}}(D) = \mu^\sigma(D)$$

for all $D \subseteq \mathscr{P}$ and for any $\sigma \in \mathfrak{S}_n$ such that $\mathbb{1}_D \in B_0^\sigma(\mathscr{P}, \mathscr{F})$. Notice that, as $B_0(\mathscr{P}, \mathscr{F}) = \cup_{\sigma \in \mathfrak{S}_n} B_0^\sigma(\mathscr{P}, \mathscr{F})$, by the separation property for each $D \subseteq \mathscr{P}$, there exists $\sigma \in \mathfrak{S}_n$ such that $\mathbb{1}_D \in B_0^\sigma(\mathscr{P}, \mathscr{F})$, so $\nu^{\mathscr{P}}$ is defined everywhere. It is easy to verify that, for any $\varphi \in B_0(\mathscr{P}, \mathscr{F})$, we have $I(\varphi) = \int_{\mathscr{P}} \varphi \, \mathrm{d}\nu^{\mathscr{P}}$.

Summarizing the results of this step of the proof, we have that, for any $f, g \in \mathscr{F}$,

$$f \succsim g \Leftrightarrow V_{\mathscr{P}}(f) \geq V_{\mathscr{P}}(g),$$

where

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

- **Step 4** Let $u_{\mathscr{P}} : X \to \mathbb{R}$ be defined by $u_{\mathscr{P}}(x) = U_{\mathscr{P}}(\delta_x)$. By axiom 8, for all $x, y \in X$, for all $\mathscr{P}, \mathscr{P}' \in \mathfrak{P}_0$, $(x, \mathscr{P}) \succsim (y, \mathscr{P})$ if and only if $(x, \mathscr{P}') \succsim (y, \mathscr{P}')$. But this is clearly equivalent to $(\delta_x, \mathscr{P}) \succsim^{\mathscr{P}}_{AA} (\delta_y, \mathscr{P})$ if and only if $(\delta_x, \mathscr{P}') \succsim^{\mathscr{P}'}_{AA} (\delta_y, \mathscr{P}')$. 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 7. We have:

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

  But, on the other hand,

$$\begin{aligned} u(x_f) &= V_{\mathscr{P}}(f) \\ &= \int_{\mathscr{P}} U_{\mathscr{P}}(P^f) \, \mathrm{d}\nu^{\mathscr{P}}(P) \\ &= \int_{\mathscr{P}} \int_X u \, \mathrm{d}P^f \, \mathrm{d}\nu^{\mathscr{P}}(P) \\ &= \int_{\mathscr{P}} \int_S u \circ f \, \mathrm{d}P \, \mathrm{d}\nu^{\mathscr{P}}(P) \end{aligned}$$

  because $U_{\mathscr{P}}$ is affine, and the same holds for g, and this completes the proof.

References

Alain Chateauneuf, Michèle Cohen, and Jean-Yves Jaffray. Décision dans l'incertain: les modèles classiques. In Denis Bouyssou, Didier Dubois, Marc Pirlot, and Henri Prade, editors, *Concepts et Méthodes pour l'aide à la décision*, page ? Hermes, forthcoming.

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

Thibault Gajdos, Jean-Marc Tallon, and Jean-Christophe Vergnaud. Decision making with imprecise probabilistic information. *Journal of Mathematical Economics*, 40(6):647–681, 9 2004.

Paolo Ghirardato and Massimo Marinacci. Ambiguity made precise: A comparative foundation. *Journal of Economic Theory*, 102(2): 251–289, February 2002.

Paolo Ghirardato and Massimo Marinacci. Risk, ambiguity, and the separation of utility and beliefs. *Mathematics of Operations Research*, 26:864–890, 2001.

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

Peter Klibanoff, Massimo Marinacci, and Sujoy Mukerji. A smooth model of decision making under ambiguity. Working Paper 11/03, ICER, 2003.

Klaus Nehring. Imprecise probabilistic beliefs. mimeo, November 2002.

Klaus Nehring. Ambiguity in the context of probabilistic beliefs. mimeo, November 2001.

Klaus Nehring. On the interpretation of Sarin and Wakker's 'A simple axiomatization of nonadditive expected utility'. *Econometrica*, 62(4):935–38, 1994.

Leonard J. Savage. *The Foundations of Statistics*. Wiley, New York, 1954.

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

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