CER-ETH – Center of Economic Research at ETH Zurich

A Dual Approach to Ambiguity Aversion

A. Bommier

Working Paper 14/207
December 2014

Economics Working Paper Series

Eidgenössische Technische Hochschule Zürich
Swiss Federal Institute of Technology Zurich

A Dual Approach to Ambiguity Aversion*

Antoine Bommier (ETH-ZURICH) †

December 19, 2014

In this paper, the assumption of monotonicity of Anscombe and Aumann (1963) is replaced by a weaker assumption of monotonicity with respect to first order stochastic dominance. I derive a representation result where ambiguous distributions of objective beliefs are first aggregated into "equivalent unambiguous beliefs" and then risk preferences are used to compute the utility of these equivalent unambiguous beliefs. Such an approach makes it possible to disentangle ambiguity aversion, related to the treatment of information, and risk aversion, related to the evaluation of the equivalent unambiguous beliefs. An application shows the tractability of the framework and its intuitive appeal.

Keywords: ambiguity aversion, first-order stochastic dominance, separability, comonotonic sure-thing principle, rank-dependent utility, saving behavior.

JEL codes: D81.

1 Introduction

A convenient framework to model choice under uncertainty is the horse-roulette lottery introduced by Anscombe and Aumann (1963) (thereafter "AA"). There is a weakly ordered set of consequences $(X, \geq)$, a (thereafter finite) set of states of the world: $S = \{1, ...\Omega\}$. Acts are defined as applications from S into $\mathcal{L}(X)$, where $\mathcal{L}(X)$ is the set of "roulette lotteries" with consequences in X. The set of acts is therefore isomorphic to $(\mathcal{L}(X))^{\Omega}$. The issue is to adequately model preferences over such a set. One of the structuring assumptions introduced by AA is that of "monotonicity". More precisely, from a preference relation $\succeq$ over $\mathcal{L}(X)^{\Omega}$,

*This work was motivated by numerous discussions with Marie-Charlotte Guetlein. This was her work (Guetlein, 2014) that convinced me that it would be worth exploring other routes, and in particular those relaxing the assumption of monotonicity of Anscombe and Aumann (1963). I would like to thank her for the role she had in this project. I am also grateful to Thibault Gajdos, Peter Klibanoff and Jean-Marc Tallon for their comments.

†Antoine Bommier, Chair of Integrative Risk Management and Economics at ETH Zurich. E-mail: abommier@ethz.ch. I gratefully acknowledge financial support from Swiss-Re.

one builds a preference relation $\succeq_{\mathcal{L}}$ over $\mathcal{L}(X)$ by looking at constant acts.[1] Monotonicity in the sense of AA involves assuming that the preference relation $\succeq$ (defined over $\mathcal{L}(X)^{\Omega}$) is monotonic with respect to the preference relation $\succeq_{\mathcal{L}}$ (defined over $\mathcal{L}(X)$). Formally that means if we have two acts $L = (l_1, l_2, .., l_{\Omega})$ and $M = (m_1, m_2, ..., m_{\Omega})$ such that $l_i \succeq_{\mathcal{L}} m_i$ for all $i \in \{1, ...\Omega\}$, it must be the case that $L \succeq M$. Such an assumption has been maintained in about all papers dealing with ambiguity aversion in such a horse-roulette lottery setting.

One may however look at AA's monotonicity assumption from a different perspective. Indeed that assumption implies that:

$$((l_1, l_2, .., l_{\Omega}) \succeq (m_1, l_2, .., l_{\Omega})) \Leftrightarrow (l_1 \succeq_{\mathcal{L}} m_1) \Leftrightarrow ((l_1, m_2, ..., m_{\Omega}) \succeq (m_1, m_2, ..., m_{\Omega}))$$

In other words, preferences over what may happen in the state of the world "1" (and more generally in any given state of the world $i \in \{1, ...\Omega\}$) are independent of what may happen in the other states of the world. Thus AA's monotonicity criterion is also imposing a state separability assumption.[2] It is worth contrasting this assumption with one the statements made by Mark Machina:

> If there is a general lesson to be learned from Ellsberg's examples and the examples here, it is that the phenomenon of ambiguity aversion is intrinsically one of nonseparable preferences across mutually exclusive events, and that models that exhibit full—or even partial—event-separability cannot capture all aspects of this phenomenon.
>
> (Machina, 2009, p 390)

In a horse-roulette setting, states of the world are nothing else than mutually exclusive events. Why then impose state-separability in ambiguity models if ambiguity aversion is precisely about non separability of preferences across mutually exclusive events?

In order to illustrate the implications of the separability property inherent to AA's monotonicity assumption, let us consider the simple fictive experiment. Think that there are two urns A and B with in each urn an undetermined even number of balls. Importantly, the number of balls in urn A needs not be the same as the one in urn B. The only information that is provided is that in urn A, the balls are black or white in equal proportions, and in urn B the balls are red or green in equal proportions. Assume that a third urn C may be built by putting all the balls of urns A and B into a single urn. Since the number of balls in A and B

[1] The preference relation $\succeq_{\mathcal{L}}$ is defined by $l \succeq_{\mathcal{L}} m \Longleftrightarrow (l, ..., l) \succeq (m, ..., m)$.

[2] Several papers, like Schmeidler (1989) or Maccheroni, Marinacci, and Rustichini (2006), introduce the monotonicity assumption of AA as a "state independence condition", which is an even stronger property than state separability.

are unknown, the composition of urn C is uncertain. For example, the proportion of red balls in urn C may be any rational number in $(0, \frac{1}{2})$. Now, assume that a decision maker (DM, thereafter) has to choose an urn (A, B or C), knowing that (i) a single ball will be drawn out that urn, and (ii) the color of the drawn ball will determine his pay-off. In order to avoid too obvious comparisons, assume that the pay-offs are such that:

$$\text{payoff(white)}<\text{payoff(green)}<\text{payoff(red)}<\text{payoff(black)}$$

Acts associated to urns A and B are simple unambiguous 50/50 risky bets, with A being a spread of B. The composition of urn C is uncertain, and one may think that ambiguity aversion would decrease the value associated to urn C. However, if AA-monotonicity is assumed (and preferences over constant acts are of the expected utility kind), one should have $A \succeq C \succeq B$ or $B \succeq C \succeq A$. In particular, if the decision maker is indifferent between the unambiguous urns A and B, he should also be indifferent between the urn A and the uncertain urn C, whatever his degree of ambiguity aversion. Whether this is an appealing feature of ambiguity models is questionable. It might indeed make sense to think that ambiguity aversion could make the unambiguous urns A and B more valuable than the uncertain urn C.

The objective of this paper is to explore an alternative route where ambiguity aversion is modeled without imposing state separability, so as to have a framework where the indifference between the urns A and B in the fictive experiment described above does not necessarily imply the indifference between the urns A and C when $A \sim B$. An assumption of monotonicity will still be introduced, but a weaker version than that of AA. Namely, monotonicity will be expressed with respect to first-order stochastic dominance, which is only a partial weak order on $\mathcal{L}(X)$, instead of being expressed with respect to a complete weak order on $\mathcal{L}(X)$.

Substituting AA's monotonicity assumption by monotonicity with respect to first order stochastic dominance leaves us with the possibility to impose other structuring assumptions without necessarily boiling down to the subjective expected utility framework axiomatized by AA. In the current paper, I consider the case of the "comonotonic sure-thing principle" of Chew and Wakker (1996), where comonotonicity is meant with respect to the order over X, the set of deterministic consequences.[3] Our main result is that rational preferences can be represented by utility functions where: (i) in a first stage, ambiguous beliefs are aggregated to form an "equivalent unambiguous belief"; (ii) in the second stage, this "equivalent unambiguous belief" is plugged into risk preferences to get a final evaluation. The procedure is very general, but as we assume the comonotonic sure-thing principle, we obtain that risk preferences must

[3] This is different from the approach of Schmeidler (1989) where comonotonic independence was assumed considering the order $\succeq_{\mathcal{L}}$ over roulette lotteries.

be rank-dependent in the sense of Green of Jullien (1988).[4]

The contrast between our approach and the one that assumes AA-monotonicity is pretty straightforward. In "AA-monotone" models, like those of Schmeidler (1989), Gilboa and Schmeidler (1989), Klibanoff, Marinacci and Mukerji (2005), or more generally any "MBA model" of Cerreia-Vioglio, Ghirardato, Maccheroni, Marinacci, and Siniscalchi (2011), "risk preferences" are used to relate subjective distributions of roulette lotteries to subjective distributions of utility levels. Then, some particular aggregation procedure is used to evaluate this subjective distribution of utility levels. The diversity of AA-monotone ambiguity models results from the way this second stage aggregation is achieved, but the two stage modus operandi is maintained in most contributions, as a direct consequence of the assumed state separability. What we propose is to follow a symmetric path, where a subjective distribution of roulette lotteries is first aggregated to provide a single roulette lottery and then risk preferences used to estimate the utility of that roulette lottery. Figure 1 illustrates the difference between the two ways of modeling uncertainty aversion.

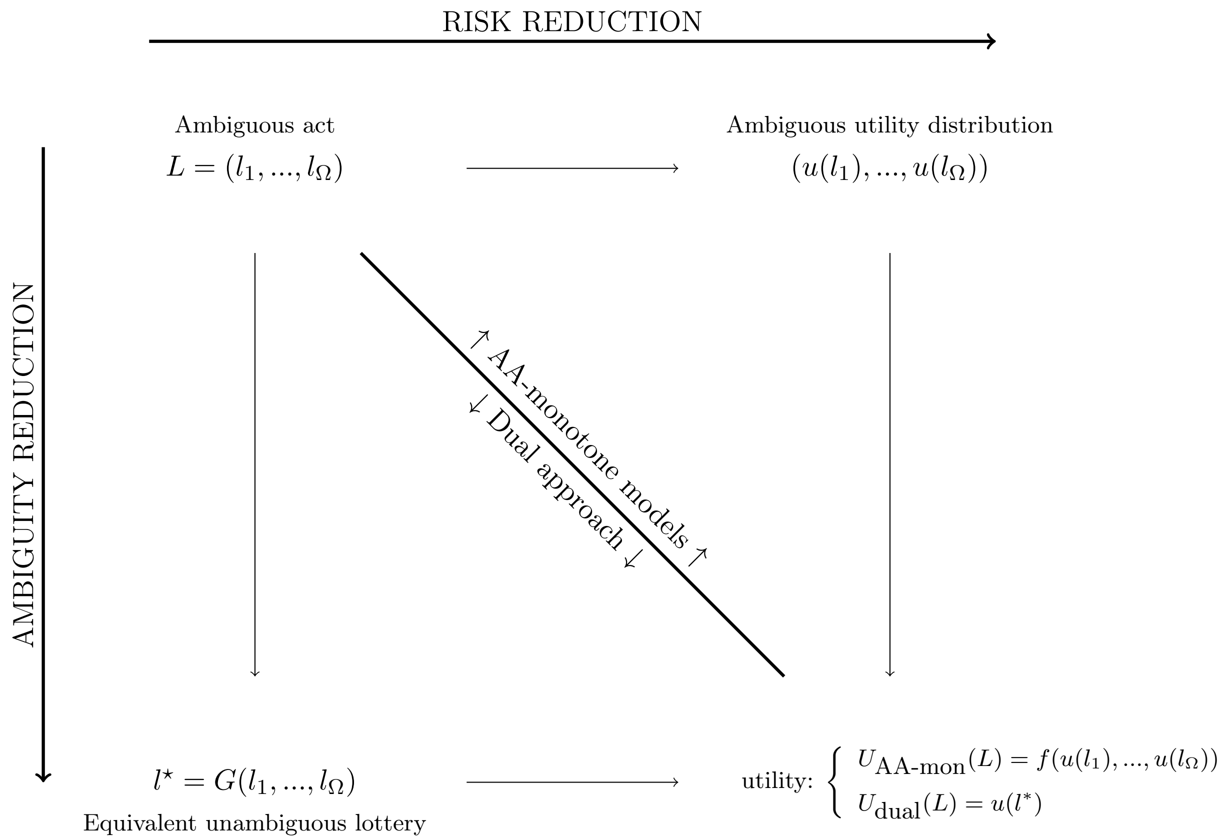

Figure 1: Dual approaches to ambiguity aversion

[4]That includes the expected utility and rank-dependent expected utility frameworks as special cases.

A key feature of the dual approach is that it is indistinguishable from the traditional approach when the set of consequences contains only two elements. Indeed, when the set of consequences has only two elements, first-order stochastic dominance generates a complete order over "roulette lotteries" and both the traditional and the dual routes are totally identical. The dual approach will therefore provide models that do just as well as the AA-monotone models to explain the results of standard two-outcome Ellsberg's urn experiments.[5]

Compared to the AA-monotone models, the dual approach admittedly provides just another way to model Ellsberg's paradox. It is worth however to emphasize two aspects of such an alternative approach. The first one is its ability to separate risk aversion and ambiguity aversion. Importantly, the way subjective beliefs are transformed into an equivalent objective belief is independent of risk preferences. Ambiguity aversion only plays a role in that first stage, with stronger ambiguity aversion leading to using lower unambiguous equivalent beliefs (lower in the sense of first-order stochastic dominance). It is then obvious that agents with different risk aversion can be compared in terms of ambiguity aversion - something that could not be obtained in models like that of Klibanoff, Marinacci and Mukerji (2005).

The second and related aspect worth being mentioned is that the dual approach is very simple to use in applications. In fact, an increase in ambiguity aversion is equivalent to using a more pessimistic information in the sense of first-order stochastic dominance. The impact of an increase in ambiguity aversion is then totally similar to that of a change in risk, where one goes from a given distribution to another dominated one. The literature on behavior under risk that has explored the impact of such changes in risk can then be directly used to conclude about the impact of ambiguity aversion. As an example we will study the saving behavior of an ambiguity averse agent who receives ambiguous information about his future income.

While I was circulating a preliminary version of this paper, Peter Klibanoff kindly mentioned me the model of Expected Utility with Uncertain Probabilities (thereafter "EUUP") of Izhakian (2014a, 2014b)[6], which I was not aware of. There are common ideas in Izhakian (2014a) and the current paper but also fundamental differences. In common, is the idea that ambiguity aversion does not necessarily have to come into play when aggregating utility levels of objective lotteries, but could directly come into play at the beliefs aggregation stage. In other words, both papers employ the dual route that passes through the south-west corner in Figure 1. However, a fundamental difference is that in the EUUP model, ambiguity aversion involves lowering the certainty equivalent probability attributed to uncertain event probabilities, whatever the consequences of the events. It is thus logically related, in Izhakian

[5]To find a difference between the dual and the AA-monotone models one would have to consider experiments with at least three outcomes. Machina (2014) emphasizes how limited is our knowledge in that case. I could not find an existing experiment that would help to discriminate between the dual and an the traditional approach. I will explain in Section 6.1 the kind of experiments that would make it possible to do so.

[6]These are companion papers built upon an earlier joint version (Izhakian, 2012).

(2014b), to the support theory of Tversky and Koehler (1994) and Rotttenstreich and Tversky (1997). Instead, the current paper sticks to the notion of uncertainty aversion of Schmeidler (1989). Ambiguity aversion then precisely consists in opting for an aggregation of probabilities of events that depends on their consequences. The divergence eventually translates into the specifications that are suggested. While the model suggested by Izhakian writes as a particular case of our second representation result, the standard EUUP specification exhibits ambiguity aversion in the sense of Izhakian, but not in the sense of Schmeidler. A more thorough discussion of differences and commonalities of both approaches is provided in Section 7.

2 Setting

We consider a connected compact set of consequences X. We denote by $\mathcal{L}(X)$ the set of simple lotteries on X. For any simple lottery $l \in \mathcal{L}(X)$ we will denote by $supp(l)$ its support, that is the finite set of elements of X to which l assigns a positive probability. For any $x \in X$ we denote by $\delta_x \in \mathcal{L}(X)$ the lottery that assigns probability 1 to the consequence x.

There is a finite set $S = \{1, ..., \Omega\}$ of states of the world. An act is a function from S into $\mathcal{L}(X)$. We thus denote $\mathcal{F} = \mathcal{L}(X)^\Omega$ the set of acts. This set is endowed with the weak topology. For any act $L = (l_1, ...l_\Omega) \in \mathcal{F}$, we denote by $supp(L)$ the support of L, formally defined by $supp(L) = \cup_{1 \leq i \leq \Omega} supp(l_i)$. For any $x \in X$ we denote by Δ_x the act that provides the degenerate lottery δ_x in all states of the world. That is: $\Delta_x = (\delta_x, ..., \delta_x)$.

For any subset $A \subset X$, and any act $L = (l_1, ..., l_\Omega) \in \mathcal{F}$, we denote by $Prob(L \in A) \in [0,1]^\Omega$ the vector $(Prob(l_1 \in A), ..., Prob(l_\Omega \in A))$. A similar notation will be used when "$\in A$" is replaced by some other logical operation. For example, for $x \in X$, we denote by $Prob(L = x)$ the vector $(Prob(l_1 = x), ..., Prob(l_\Omega = x)) \in [0,1]^\Omega$.

The set of acts $\mathcal{F}$ is endowed with a natural mixture operation:

$$\alpha(l_1, ..., l_\Omega) + (1-\alpha)(m_1, ..., m_\Omega) = (\alpha l_1 + (1-\alpha)m_1, ..., \alpha l_\Omega + (1-\alpha)m_\Omega)$$

An act is said to be constant if and only if it is a diagonal element (i.e., an element of the form $(l, ..., l)$ for some $l \in \mathcal{L}(X)$). We denote by $\mathcal{F}^c$ the set of constant acts. An act is said to be deterministic if it is of the form Δ_x for some $x \in X$. Deterministic acts are therefore degenerate constant acts.

For any scalar $q \in [0,1]$ we will denote by $\vec{q} \in [0,1]^\Omega$ the constant vector $(q, ...q)$. In particular, $\vec{0}$ and $\vec{1}$ will denote the constant vectors $(0, ..., 0)$ and $(1, ..., 1)$. As is standard, for any $p = (p_1, ..., p_\Omega)$ and $p' = (p'_1, ..., p'_\Omega)$ in $[0,1]^\Omega$, the statement $p \geq p'$ is to be understood as $p_i \geq p'_i$ for all $i \in \{1, ..., \Omega\}$.

The first representation result that will be provided (Proposition 1) does not require preferences over constant acts to be expected utility. This gain of generality comes with the need to consider utility representations that take values in $\bar{\mathbb{R}} = \mathbb{R} \cup \{-\infty, +\infty\}$, just like in the rank-dependent models of Wakker (1993) and Chateauneuf (1999) for preferences over lotteries. To state our result rigorously, we need to mention some specific conditions of continuity and monotonicity, that we group under the appellation "Ω-admissibility".

Definition 1 *Assume that X is provided with a weak order $\geq$. The function $v : X \times [0,1]^{\Omega} \to \bar{\mathbb{R}}$ is said to be Ω-admissible if it fulfills the following properties:*

1. $v(x, \vec{0}) = 0$ *for all* $x \in X$.
2. $v(x,p) < +\infty$ *expect possibly when* $x \in max(X)$ *if* $p = \vec{1}$.[7]
3. $v(x,p) > -\infty$ *expect possibly when* $x \in min(X)$ *if* $p \neq \vec{0}$.
4. *if* $v(x,p) = -\infty$ *for some* $x \in min(X)$ *and some* $p \neq \vec{0}$, *then it is the case for all* (y,q) *with* $y \in min(X)$ *and* $q \neq \vec{0}$.
5. *For any* $p \neq \vec{0}$ *the function* $x \to v(x,p)$ *is continuous and strictly increasing.*
6. *For any* p *different from* $\vec{0}$ *and* $\vec{1}$, *the function* $x \to v(x,\vec{1}) - v(x,p)$ *is continuous and strictly increasing.*
7. *For any* $x, y \in X$, *with* $x > y$, *the function* $p \to v(x,p) + (v(y,\vec{1}) - v(y,p))$ *is continuous and strictly increasing.*[8]

When $\Omega = 1$ (which provides the notion of "1-admissibility"), these conditions coincide with those listed in Definition 1 of Chateauneuf (1999). The notion of Ω-admissibility can be viewed as meaning "continuous and strictly increasing", in a sense that has been fine-tuned to deal with the normalization $v(x,\vec{0}) = 0$ and the possibility of infinite valuations.

3 Representation result

3.1 Assumptions

We consider a binary relation $\succeq$ over $\mathcal{F}$ which is assumed to fulfill the following axioms:

Axiom 1 (rationality) *The relation $\succeq$ is complete, transitive.*

[7] We will denote $min(X) = \{x | y \geq x \text{ for all } y \in X\}$ and $max(X) = \{x | x \geq y \text{ for all } y \in X\}$.

[8] If y is a minimal element the term $(v(y,\vec{1}) - v(x,p))$ may look like $-\infty - (-\infty)$. In that case the convention, $-\infty - (-\infty)$=0 should be used, meaning then that the function $p \to v(x,p)$ has then to be continuous and strictly increasing.

Axiom 2 (non-triviality) *There exists $L, M \in \mathcal{F}$ such that $M \not\preceq L$.*

Axiom 3 (continuity) *For any $L \in \mathcal{F}$ the sets $\{M \in \mathcal{F} | M \succeq L\}$ and $\{M \in \mathcal{F} | L \succeq M\}$ are closed subsets of $\mathcal{F}$.*

From the preference relation $\succeq$, the strict preference relation $\succ$ and the indifference relation $\sim$ are build as usual.

The above axioms make it possible to define a weak order $\geq$ on X. Formally, we define $\geq$ by setting that for all $x, y \in X$:

$$x \geq y \Leftrightarrow \Delta_x \succeq \Delta_y$$

Due to Axiom 3 this weak order is continuous on X.[9] Moreover since X is compact it always admits extreme elements and the sets $min(X)$ and $max(X)$ are non empty.

Based on the weak order relation we can express our assumption of monotonicity:

Axiom 4 (monotonicity with respect to first-order stochastic dominance) *For any $L, M \in \mathcal{F}$, if $Prob(L \geq x) \geq Prob(M \geq x)$ for all $x \in X$, then $L \succeq M$. Moreover if, in addition, there exists some $x \in X$ for which the inequality is strict, then $L \succ M$.*

Axiom 4 is a monotonicity property that is weaker than that of AA. Indeed, Axiom 4 imposes some consistency between the preference relation $\succeq$ and its restriction to the set of deterministic acts, while AA-monotonicity imposes some consistency between the preference relation $\succeq$ and its restriction to the set of constant acts.

Both assumptions can be viewed as imposing different degrees of separability. When Axiom 4 is introduced, one may associate a utility level in $[0, 1]$ to any element $x \in X$, and then evaluates acts on X as if they were the real-valued acts obtained by replacing consequences by their utility levels. Thus, instead of evaluating elements of $\mathcal{L}(X)^\Omega$ one has to evaluate elements of $\mathcal{L}([0, 1])^\Omega$. With AA monotonicity an additional separability assumption is introduced. To each lottery in $\mathcal{L}([0, 1])$ one can associate a utility level in $[0, 1]$. Then acts are evaluated as vectors in $[0, 1]^\Omega$.[10]

We can now introduce the following axiom:

Axiom 5 (comonotonic sure-thing principle) *Consider n vectors of probabilities, $p_i = (p_{i1}, ...p_{i\Omega}) \in [0, 1]^\Omega$, with $i \in \{1, ...n, \}$, that fulfills $\sum_{i=1}^n p_i = \vec{1}$. And consider two acts $\sum_{i=1}^n p_i\delta_{x_i} = (\sum_{i=1}^n p_{i1}\delta_{x_i}, ..., \sum_{i=1}^n p_{i\Omega}\delta_{x_i})$ and $\sum_{i=1}^n p_i\delta_{y_i} = (\sum_{i=1}^n p_{i1}\delta_{y_i}, ..., \sum_{i=1}^n p_{i\Omega}\delta_{y_i})$,*

[9] See for example Lemma 1 in Chateauneuf (1999).

[10] The implication of AA monotonicity is for example explicit in Cerreia-Vioglio, Ghirardato, Maccheroni, Marinacci, and Siniscalchi (2011), where this second-level separability is visible in the representation result shown in their Theorem 1.

where x_1,x_2,...,x_n *and* $y_1, y_2, ..., y_n$ *are two non decreasing sequences of elements of* X*. Assume moreover that for some* $j \in \{1, ..., n\}$ *we have* $x_j = y_j$*. Then for all* z *such that:*

$$\begin{cases} x_{j-1} \leq z \leq x_{j+1} & \textit{and} \quad y_{j-1} \leq z \leq y_{j+1} & \textit{if} \quad j \neq 1 \\ z \leq x_{j+1} & \textit{and} \quad z \leq y_{j+1} & \textit{if} \quad j = 1 \\ x_{j-1} \leq z & \textit{and} \quad y_{j-1} \leq z & \textit{if} \quad j = n \end{cases} \tag{1}$$

we have:

$$\left(\sum_{i=1}^{n} p_i \delta_{x_i} \succeq \sum_{i=1}^{n} p_i \delta_{y_i} \right) \Leftrightarrow \left(p_j \delta_z + \sum_{i \neq j} p_i \delta_{x_i} \succeq p_j \delta_z + \sum_{i \neq j} p_i \delta_{y_i} \right).$$

This last axiom is directly imported from Chateauneuf (1999) who used the comonotonic sure-thing principle to axiomatize the rank-dependent utility model of Green of Jullien (1988). This axiom can be viewed as a restricted version of Savage sure-thing principle that would be adapted to the horse-roulette setting. It is only a restricted version of it, as we introduce the requirement (1) stipulating that common consequences of two acts can be changed without impacting the comparison between these acts, only when these changes are rank preserving. Interestingly, an assumption that looks similar to Axiom 5 can be found in Schmeidler (1989). However, in Schmeidler (1989), comonotonicity was considered with respect to the order $\succeq_{\mathcal{L}}$ over the set $\mathcal{L}(X)$, while the notion of comonotonicity that is behind Axiom 5 refers to the order over X. It is therefore not contingent on some risk preferences. In fact, the comonotonic independence axiom of Schmeidler (1989) could be viewed as an expression of the comonotonic sure-thing principle when roulette lotteries are considered as "sure-things". Our approach is different as we take the "sure-things" as being the deterministic consequences, that is the elements of X.

What is the right degree of separability to assume may of course be a matter of discussion. If we combine "natural separability assumptions" like AA-monotonicity and the sure-thing principle of Savage –which could be stated as in Axiom 5 without introducing the requirement (1)– we end up with the subjective expected utility model which cannot account for Ellsberg's paradox. Some assumptions have to be relaxed. About all papers decided to maintain AA-monotonicity and to forget about the sure-thing principle. In the current paper, a different approach is taken. AA-monotonicity is replaced by monotonicity with respect to first-order stochastic dominance, and the sure-thing principle, replaced by the comonotonic version of it. Thus instead of maintaining one assumption and removing the other, we use weaker versions for both.

A key difference between Axiom 4 and the monotonicity property of AA, is that the formulation of Axiom 4 does not rely on some preferences over $\mathcal{L}(X)$. In other words this notion of monotonicity is independent of some imputed risk preferences. This is crucial as this is what makes it possible to easily disentangle risk and ambiguity aversion. In fact, one of the conceptual differences between the AA-monotone models and the dual approach is that, in AA-monotone models, the assumptions regarding risk and ambiguity preferences are nested in each other. One can indeed view AA-monotone models as cases where: (i) a separability assumption is made, implying the existence of (state independent) risk preferences; (ii) additional assumptions whose formulations make use of these risk preferences are introduced. Among these additional assumptions, we would for example find "monotonicity with respect to risk preferences" (i.e., AA-monotonicity), Schmeidler (1989)'s assumption of comonotonic independence, or assumptions related to preferences over second order acts in Klibanoff, Marinacci and Mukerji (2005). This nested approach does not call for a separation of risk and ambiguity preferences. The dual approach that is introduced in the current paper avoids such a nested procedure. None of the axioms is related to preferences over constant acts.

3.2 A rank-dependent ambiguity model

We can state our main representation result. For that result we do not impose that preferences over constant acts fit in the expected utility framework, implying that we have to consider possibly non-finite representations and the notion of Ω-admissibility introduced in Definition 1. A simpler but more restrictive representation result will be stated in Proposition 2, assuming that preferences over constant acts fulfill the independence axiom.

Proposition 1 *Axioms 1 to 5 are fulfilled if and only the preference relation $\succeq$ can be represented by:*

$$U(L) = \sum_{x \in S(L)} (v(x, Prob(L \geq x)) - v(x, Prob(L > x)).) \tag{2}$$

where v is an Ω-admissible function and $S(L)$ is any finite subset of X, such that for any $y \in supp(L)$, there exists a unique $x \in S(L)$ such that $y \sim x$.

Moreover if $v(x, p) > -\infty$ for all $x \in X$ and $p \in [0, 1]^\Omega$, the preference relation $\succeq$ can be also represented as in (2) with the function v replaced by any function w such that $w(x, p) = \alpha v(x, p) + \beta(p)$, where α is a positive scalar and β a real valued function.

Proposition 1 is a multidimensional extension of Theorem 1 of Chateauneuf (1999). While Chateauneuf (1999) was discussing preferences defined over $\mathcal{L}(X)$ we discuss preferences over $\mathcal{L}(X)^\Omega$. Proposition 1 simply reformulates the main result of Chateauneuf (1999) in that setting.

Note that if there is an $x \in S(L)$ for which no $y \in supp(L)$ fulfills $x \sim y$, then $Prob(L \geq x) = v(x, Prob(L > x))$. The corresponding term in the sum (2) is then equal to 0. Moreover, from property 5 of $\Omega-$admissible functions, we know that if we have $x, y \in$X such that $x \sim y$ then $v(x,p) = v(y,p)$ fore any $p \in [0,1]^{\Omega}$. These two points guarantee that the sum (2) is independent of the choice of a particular set $S(L)$.

Since $S(L)$ is finite, there is always a finite numbers of terms in the sum (2). However, due to properties 2 to 4 of Ω-admissible functions, some of these elements can be infinite. It may happen (because of the property 4 in Definition 1) that for minimal outcomes x the term $v(x, Prob(L \geq x)) - v(x, Prob(L > x))$ writes as $-\infty - (-\infty)$. In that cases, the convention $-\infty - (-\infty) = 0$ should be adopted. In fact, the only acts that may provide an utility equal to $+\infty$ are the acts that provide maximal outcomes in all circumstances. And the only acts that may provide utility $-\infty$ are those that provide minimal elements in all circumstances. Thus utility is finite, except possibly for minimal and maximal acts.

Proof. Let us show that representation by (2) implies that Axioms 1 to 5 are fulfilled. Axiom 1 is satisfied whenever there exists a utility representation. Axioms 2 to 4 results from the monotonicity and continuity properties of v. As for Axiom 5, applying representation (2) to elements of the form $\sum_{i=1}^{n} p_i \delta_{x_i}$, with $x_n \geq x_{n-1} \geq .. \geq x_1$, we obtain:

$$U(\sum_{i=1}^{n} p_i \delta_{x_i}) = \sum_{i=1}^{n} \left(v(x_i, \sum_{k=i}^{n} p_k) - v(x_i, \sum_{k=i+1}^{n} p_k) \right)$$

A similar expression holds for $U(\sum_{i=1}^{n} p_i \delta_{y_i})$. Thus if $x_j = y_j = z$ we have $U(\sum_{i=1}^{n} p_i \delta_{x_i}) - U(\sum_{i=1}^{n} p_i \delta_{y_i})$ which is independent of z and Axiom 5 holds.

The difficult part of the proof is to show that Axioms 1 to 5 do imply representation (2). It is found in appendix. For that part, we heavily rely on the proof found in Chateauneuf (1999) which itself uses a key result of Wakker (1993). ∎

3.3 On beliefs aggregation

In order to emphasize that what we propose is a dual approach where ambiguous distributions of beliefs are first aggregated into an equivalent unambiguous belief, which is then evaluated with some risk preferences, we may want to rewrite our representation result as follows:

Corollary 1 *If Axioms 1 to 5 are fulfilled, there exists a 1-admissible function $u : X \times [0,1] \to \mathbb{R}$ and, for all $x \in X$, a strictly increasing function $G_x : [0,1]^{\Omega} \to [0,1]$, fulfilling $G_x(\vec{q}) = q$ for all $q \in [0,1]$, such that the preference relation $\succeq$ can be represented by:*

$$U(L) = \sum_{x \in S(L)} \left(u(x, G_x(Prob(L \geq x))) - u(x, G_x(Prob(L > x))) \right) \tag{3}$$

where $S(L) \subset X$ is as in Proposition 1.

Proof. Let us use representation (2). First observe that v can always be normalized so that, for any non minimal element $x \in X$, the function $p \rightarrow v(x,p)$ is increasing in p. Indeed, if $v(y,p) = -\infty$ for some y, then the property 7 of Ω-admissible functions applied to minimal elements y gives the result. And if $v(y,p)$ is never equal to $-\infty$, we can replace $v(x,p)$ by $v(x,p) - v(y_0,p)$ where y_0 is a given minimal element, to get another utility representation of the preference relation. When normalized in that way, $v(x,p)$ is equal to zero when $x = y_0$. Applying property 7 with $y = y_0$ we then obtain that for any non minimal element x of X the function $p \rightarrow v(x,p)$ is increasing in p.

Assume that v has been normalized as suggested above. Define $u : X \times [0,1] \rightarrow \mathbb{R}$ by $u(x,q) = v(x,\vec{q})$ for all $x \in X$ and $q \in [0,1]$. And for any non minimal element x, define $G_x : [0,1]^\Omega \rightarrow [0,1]$ as the solution of:

$$u(x, G_x(p)) = v(x,p)$$

The fact that there is always a unique solution to this equation is due to the monotonicity of u (implied by that of v) and the continuity axiom. Moreover $G_x(p_1,...p_\Omega) \in [0,1]$ because $v(y,\vec{0}) \leq v(y,p_1,....p_\Omega) \leq v(y,\vec{1})$. It is moreover obvious, from the definition of u, that $G_x(p)$ is strictly increasing in p and that for any $q \in [0,1]$ we have $G_x(\vec{q}) = q$.

For any minimal element y, let us define $G_y(p)$ by $G_y(p) = \frac{1}{\Omega}\sum_{i=1}^{\Omega} p_i$, which is a strictly increasing function that fulfills $G_y(\vec{q}) = q$ for all $q \in [0,1]$. The definition of G_y for minimal elements y looks arbitrary but, given the normalization of v, this has now impact. Indeed for any $p, p' \in [0,1]^\Omega$ with $p \neq \vec{0}$ and $p' \neq \vec{0}$, and any minimal element y we have $v(y,p) = v(y,p')$ (the value being either always 0 or always $-\infty$).

With the above definition we have:

$$u(x, G_x(p)) - u(y, G_x(p')) = v(x,p) - v(x,p')$$

for all non minimal element x, and all $p, p' \in [0,1]^\Omega$. The equality remains moreover true for minimal elements when $p \neq \vec{0}$ and $p' \neq \vec{0}$. Thus representation (2) rewrites as in (3) whenever there are non minimal elements in $supp(L)$. We can easily check that (2) and (3) coincide when all elements of $supp(L)$ are minimal elements. Indeed in such a case both representations provide either $U(L) = -\infty$ or $U(L) = 0$, depending on whether the function v in (2) is bounded from below. ■

With representation (3) it is clear that when we consider an act $L = (l_1, ..., l_\Omega)$, one first aggregates the decumulative distributions $Prob(l_i \geq x)$ into a single decumulative distribution function $G_x(Prob(l_1 \geq x), ..., Prob(l_\Omega \geq x))$ and then plugs this unambiguous decumulative distribution into risk preferences of the rank-dependent utility kind.

Our approach may be connected to the vast literature on beliefs aggregation, in which the issue is to find a "well-behaved" application $\phi : \mathcal{L}(X)^{\Omega} \to \mathcal{L}(X)$ that would aggregate Ω beliefs into a single one. What should mean "well-behaved" in that stetting is of course a focal point in the debate. McConway (1981) and Wagner (1982) suggested that there should be a function $G : [0,1]^{\Omega} \to [0,1]$ such that for any subset $A \subset X$:

$$Prob(\phi(l_1, ..., l_{\Omega}) \in A) = G(Prob(l_1 \in A), ..., Prob(l_{\Omega} \in A)) \quad (4)$$

They showed however that this is only consistent with a linear aggregation of beliefs. That would lead us back to the subjective expected utility model of AA, or some extensions of it where risk preferences would not necessarily fit in the expected utility framework. A less demanding restriction, discussed in Bordley and Wollf (1981) and McConway (1981), is obtained when allowing the function G in (4) to depend on the set A. In other words, one may want to assume that for any subset $A \subset X$, there exists a function $G_A : [0,1]^{\Omega} \to [0,1]$ such that:

$$Prob(\phi(l_1, ...l_{\Omega}) \in A) = G_A(Prob(l_1 \in A), ..., Prob(l_{\Omega} \in A)) \quad (5)$$

But, this was also shown to be only consistent with linear aggregation, though with the possibility to have negative weights (McConway, 1981). Our approach leads us to consider a type of aggregation of beliefs where equation (5) holds for any uppers-sets (i.e. sets of the form $A_x = \{y \in X | y \geq x\}$), but not necessarily for all sets.

It may seem the natural assumption to impose (5) to hold on upper-sets, especially if one wants to constrain the procedure of beliefs aggregation to be monotone with respect to first-order stochastic dominance. I am not aware, however, of contributions in the literature on beliefs aggregation that specifically consider the case where (5) is only assumed to hold for upper- (or lower-) sets.[11] In fact the literature on beliefs aggregation generally considers settings where there is no natural order, or preference relation, over the set of consequences X. The notion of upper- (or lower-) sets is then simply undefined. In those cases, stipulating that (5) should be restricted to upper-sets would simply be meaningless.

3.4 Additional structuring assumptions

The dual procedure that is suggested is very general and quite flexible. In particular ambiguity aversion may depend on outcome levels since the function G_x does not need to be the same for all x. This answers a concern raised in Machina (2014), who argued that one of the drawbacks of the usual ambiguity model was their inability to exhibit a degree of ambiguity aversion varying with income levels. Moreover risk preferences need not to be of the expected utility

[11] If equation (5) holds for uppers-sets it also holds for lower-sets, and vice-versa.

form, nor even of the rank-dependent expected utility form. They just need to fit in the rank-dependent setting of Green of Jullien (1988), as a consequence of Axiom 5.

One may however want to add more structure to the model. A possibility is to assume that the independence axiom holds for constant acts:

Axiom 6 (independence axiom for constant acts) *For any $L_1, L_2, L_3 \in \mathcal{F}^c$ and any $\lambda \in$(0,1)*

$$(L_1 \succeq L_2) \Leftrightarrow (\lambda L_1 + (1-\lambda)L_3 \succeq \lambda L_2 + (1-\lambda)L_3)$$

Then we can state the following result:

Proposition 2 *If Axioms 1 to 6 are fulfilled, there exist a continuous increasing function $u : X \to \mathbb{R}$, and for all $x \in X$, a continuous increasing function $G_x : [0,1]^\Omega \to [0,1]$, fulfilling $G_x(\vec{q}) = q$ for all $q \in [0,1]$ such that the preference relation $\succeq$ is represented by the utility function:*

$$U(L) = \sum_{x \in S(L)} u(x)(G_x(Prob(L \geq x)) - G_x(Prob(L > x))) \tag{6}$$

where $S(L) \subset X$ is as in Proposition 1. The function u is moreover unique, up to a positive affine transformation.

Proof. As was mentioned in the introduction, from a preference relation $\succeq$ over the set of acts $\mathcal{F}$, one may define a preference relation $\succeq_\mathcal{L}$ over $\mathcal{L}(X)$ by $m \succeq_\mathcal{L} l \Leftrightarrow (m, ...m) \succeq (l, ..., l)$. When Axioms 1 to 6 are fulfilled, this preference relation fulfills the axioms of expected utility. Thus, from Theorem 8.2 of Fishburn (1970), we know that there must exist a function $w : X \to \mathbb{R}$ (unique up to a positive affine transformation) such that the preferences relation $\succeq_\mathcal{L}$ is represented by:

$$V(l) = \sum_{x \in S(L)} ((\text{Prob}(l \geq x))) - (\text{Prob}(l > x))w(x)$$

By normalization we can assume that $w(y) = 0$ for any $y \in min(X)$. From (3), we also know that $\succeq_\mathcal{L}$ is also represented by

$$U(l) = \sum_{x \in S(L)} (u(x, Prob(l \geq x)) - u(x, Prob(l > x)))$$

Let us moreover assume that v (and therefore u) has been normalized as in the proof of Corollary 1. The functions $U(.)$ and $V(.)$ are two representations of the same preference relation. This implies that there exists an increasing function Φ such that $U = \Phi(V)$. Applying this to lotteries $l = p\delta_x + (1-p)\delta_y$ for some $y \in min(X)$ we get $u(x,p) = \Phi(pw(x))$ for all $p \in (0,1)$

and $x \notin min(X)$. Now applying it again to elements of the form $l = p\delta_x + (1-p)\delta_z$, where x and z are two non minimal elements with $x > z$, we obtain:

$$\Phi(pw(x)) + \Phi(w(z)) = \Phi(pw(x) + (1-p)w(z)) + \Phi(pw(z))$$

which implies that Φ is an affine function.[12] Representation (6) is then obtained. ■

We recognize in (6) a special case of the dual approach where the decumulative distribution functions of an act are first aggregated into an equivalent unambiguous decumulative distribution function, which is then evaluated using expected utility. In this formulation it is possible to have a degree of ambiguity aversion which is level dependent as the function G_x may depend on x. This may be viewed as affording an interesting flexibility. However one could also argue that the way a DM ranks urns in a two-outcome Ellsberg's urn experiment, should not be impacted when changing outcomes levels in a rank preserving way. For example, think of a DM who has to choose between two urns out of which a single ball will be drawn. The DM is told that he will get a positive payment if the drawn ball is red, and no payment otherwise. The choice of the DM will of course depend on the information that is given regarding the composition of each urn. But it might make sense to think that his choice should be independent of how large the payment may be. Indeed, as long as the payment is positive, the DM should choose the urn that he thinks gives a higher chance of drawing a red ball. One might then want to assume that the magnitude of the payment does not impact the DM's views about the likelihood to draw a red ball. This kind of independence assumption is formalized below.

Axiom 7 (level independent ambiguity aversion) *For any vectors $p, q \in [0,1]^\Omega$ and any $x, y, z, t \in X$ with $x > y$ and $z > t$, we have*

$$p\delta_x + \left(\vec{1} - p\right)\delta_y \succeq q\delta_x + \left(\vec{1} - q\right)\delta_y \Leftrightarrow p\delta_z + \left(\vec{1} - p\right)\delta_t \succeq q\delta_z + \left(\vec{1} - q\right)\delta_t \tag{7}$$

Axiom 7 says that if we consider two-outcome acts, then changing the outcome levels in a rank preserving way is not going to change the ranking of these acts.

We have then the following result:

Proposition 3 *If Axioms 1 to 7 are fulfilled, there exist a continuous increasing function $u : X \to \mathbb{R}$ and a continuous increasing function $G : [0,1]^\Omega \to [0,1]$, fulfilling $G(\vec{q}) = q$ for all $q \in [0,1]$, such that the preference relation $\succeq$ is represented by the utility function:*

$$U(L) = \sum_{x \in S(L)} u(x)\left(G(Prob(L \geq x)) - G(Prob(L > x))\right) \tag{8}$$

[12] See Theorem 1 of Chapter 2, page 132, in Aczel (2006).

where $S(L) \subset X$ is as in Proposition 1.

Proof. Assume that the preference relation is represented as in (6) with a function u which has been normalized so that $u(y) = 0$ for minimal elements y. Consider $y \in min(X)$ as given. For any $x \notin min(X)$ we have:

$$U\left(p\delta_x + (1-p)\,\delta_y\right) \quad = \quad u(x)G_x(p) \tag{9}$$

Moreover for any $q \in [0,1]$ we have:

$$U\left(\vec{q}\delta_x + (\vec{1} - \vec{q})\delta_y\right) = qu(x) \tag{10}$$

For any $x \notin min(X)$ and any $p \in [0,1]^\Omega$, we know from Axiom 3 that there exists $q \in [0,1]$ such that $p\delta_x + (1-p)\,\delta_y \sim \vec{q}\delta_x + (1-\vec{q})\,\delta_y$. Moreover Axiom 7 imposes that this q is independent of x. The scalar q thus only depends on p and we define the function $p \to G(p)$ by setting $q = G(p)$. From equations (9) and (10) we obtain that for all $x > y$ one has $u(x)G_x(p) = G(p)u(x)$ and thus $G_x(p) = G(p)$. For minimal elements $y \in min(X)$ we have $u(y) = 0$ and therefore $u(y)G_y(p) = u(y)G(p)$ for any $p \in [0,1]^\Omega$. Representation (8) then directly follows from representation (6).

The other way around, one can easily check that preferences represented by (8) fulfill Axioms 1 to 7. ■

4 Examples

As mentioned in the introduction, the diversity of AA-monotone ambiguity models is related to the variety of aggregation procedures that were suggested to associate single utility levels to subjective distributions of utility levels. A similar diversity will be found in the dual approach, but related to the way several decumulative distribution functions are aggregated into a single decumulative distribution function. In both cases one has to associate a single number to a set of possibly different numbers. The methods that were proposed for AA-monotone models may be used in the dual approach. Consider for example the "max-min" approach suggested by Gilboa and Schmeidler (1989). In that model agents behave as if they are provided with a class Π of probability distributions over S, a utility index $u : X \to \mathbb{R}$ and compute the utility of act $L = (l_1, ...l_\Omega)$ with the following formula:

$$U^{\text{MEU}}(L) = \min_{(\pi_i)\in\Pi}\left(\sum_{i=1}^{\Omega}\pi_i \sum_{x\in S(L)}\left(Prob(l_i \geq x) - Prob(l_i > x)\right)u(x)\right)$$

Models		Specification
SEU	AA-mon.	$U^{\text{SEU}}(L) = \sum_{i=1}^{\Omega} \pi_i \left(\sum_{x \in S(L)} (Prob(l_i \geq x) - Prob(l_i > x))\, u(x) \right)$
	Dual	$U^{\text{SEU}}(L) = \sum_{x \in S(L)} u(x) \left(\sum_{i=1}^{\Omega} \pi_i \,(Prob(l_i \geq x) - Prob(l_i > x)) \right)$
max-min	AA-mon.	$U^{\text{MEU}}(L) = \min_{(\pi_i) \in \Pi} \left(\sum_{i=1}^{\Omega} \pi_i \sum_{x \in S(L)} (Prob(l_i \geq x) - Prob(l_i > x))\, u(x) \right)$
	Dual	$U^{\text{MEU}}_{\text{dual}}(L) = \sum_{x \in S(L)} u(x) \left(\min_{(\pi_i) \in \Pi} \sum_{i=1}^{\Omega} \pi_i Prob(l_i \geq x) - \min_{(\pi_i) \in \Pi} \sum_{i=1}^{\Omega} \pi_i Prob(l_i > x) \right)$
CEU	AA-mon.	$U^{\text{CEU}} = \sum_{i=1}^{\Omega}{}_{Choquet} \left(\sum_{x \in S(L)} u(x)\, (Prob(l_i \geq x) - Prob(l_i > x)) \right)$
	Dual	$U^{\text{CEU}}_{\text{dual}} = \sum_{x \in S(L)} u(x) \left(\sum_{i=1}^{\Omega}{}_{Choquet} Prob(l_i \geq x) - \sum_{i=1}^{\Omega}{}_{Choquet} Prob(l_i > x) \right)$
KMM	AA-mon.	$U^{\text{KMM}} = \phi^{-1} \left(\sum_{i=1}^{\Omega} \pi_i \phi \left(\sum_{x \in S(L)} (Prob(l_i \geq x) - Prob(l_i > x))\, u(x) \right) \right)$
	Dual	$U^{\text{KMM}}_{\text{dual}} = \sum_{x \in S(L)} u(x) \left(\phi^{-1}(\sum_{i=1}^{\Omega} \pi_i \phi(Prob(l_i \geq x))) - \phi^{-1}(\sum_{i=1}^{\Omega} \pi_i \phi(Prob(l_i > x))) \right)$

Table 1: Some AA-monotone models and their dual

Note: "SEU" refers to the model that was axiomatized in Anscombe and Aumann (1963), "max-min" to that Gilboa and Schmeidler (1989), "CEU" to that of Schmeidler (1989) and "KMM" to that of Klibanoff, Marinacci and Mukerji (2005).

A natural dual formulation is obtained by considering the case where the utility of an act is computed by:

$$U^{\text{MEU}}_{\text{dual}} = \sum_{x \in S(L)} u(x) \left(\min_{(\pi_i) \in \Pi} \sum_{i=1}^{\Omega} \pi_i Prob(l_i \geq x) - \min_{(\pi_i) \in \Pi} \sum_{i=1}^{\Omega} \pi_i Prob(l_i > x) \right)$$

Actually, for many of the well-known AA-monotone ambiguity models it is relatively straightforward to formulate their dual version. Table 1 illustrates the correspondence between some AA-monotone models and their duals. One should note that in the case of the subjective expected utility, the AA-monotone and the dual specifications are equivalent. For the others, the equivalence holds only when the set of consequences contains at most two elements.

5 Ambiguity aversion and comparative ambiguity aversion

In this section, we explain how some properties of the functions that appear in our utility representations can be interpreted in terms of ambiguity aversion.

5.1 Ambiguity aversion

We will use the definition of ambiguity aversion of Schmeidler (1989).

Definition 2 (ambiguity aversion) *The preference relation $\succeq$ is said to exhibit ambiguity aversion if and only if for all acts $L, M \in \mathcal{F}$, and all $\lambda \in (0,1)$, we have:*

$$(\lambda L + (1-\lambda)M) \succeq L \text{ or } (\lambda L + (1-\lambda)M) \succeq M \tag{11}$$

Moreover preferences exhibit strict ambiguity aversion if the above holds, and for some acts $L, M \in \mathcal{F}$ we have $(\lambda L + (1-\lambda)M) \succ L$.

We can then state the following result:

Proposition 4 *If preferences are represented by the utility function (2), the quasi-concavity (resp. concavity) for all $x > y$ of the function $p \to v(x,p) + (v(y,\vec{1}) - v(y,p))$ on $[0,1]^\Omega \backslash \{\vec{0}\}$ is a necessary (resp. sufficient) condition for preferences to exhibit ambiguity aversion.*[13]

If preferences are represented by the utility function (8) the quasi-concavity (resp. concavity) of the function $p \to G(p)$ is a necessary (resp. sufficient) condition for preferences to exhibit ambiguity aversion.

Proof. First, let us show that if (2) is used, then ambiguity aversion implies that for all $x > y$ the function $p \to v(x,p) + (v(y,\vec{1}) - v(y,p))$ must be quasiconcave. Consider $x > y$. For any $p \in [0,1]^\Omega$ we have:

$$U(p\delta_x + (\vec{1}-p)\delta_y) = v(x,p) + v(y,\vec{1}) - v(y,p)$$

Ambiguity aversion imposes that for any p and q in $[0,1]^\Omega$ we have:

$$U(\lambda(p\delta_x + (\vec{1}-p)\delta_y) + (1-\lambda)(q\delta_x + (\vec{1}-q)\delta_y)) \geq \min(U(p\delta_x + (\vec{1}-p)\delta_y), U(q\delta_x + (\vec{1}-q)\delta_y))$$

[13]The term $v(y,\vec{1})$ which is independent of p, is maintained here as it makes it possible to deal with the cases where v is not bounded from below. Again the convention $-\infty - (-\infty) = 0$ should be used in those cases.

Thus we must have:

$$v(x, \lambda p + (1-\lambda)q) + (v(y, \vec{1}) - v(y, \lambda p + (1-\lambda)q))) \geq$$
$$\min(v(x,p) + (v(y,\vec{1}) - v(y,p)), v(x,q) + (v(y,\vec{1}) - v(y,q)))$$

meaning that the function $p \to v(x,p) + (v(y,\vec{1}) - v(y,p))$ is quasiconcave.

The other way around, let us assume that for any $x > y$ the function $p \to v(x,p)+(v(y,\vec{1})-v(y,p))$ is concave and show that this implies ambiguity aversion. For any two acts L and M in $\mathcal{F}$ it is always possible to find a non-decreasing finite sequence $\{x_1, ..., x_n\}$ of elements of X that contains $supp(L)$, $supp(M)$ and at least one non minimal element. For $K = L, M$ denote $p_i^K = Prob(K = x_i)$ so that $K = \sum p_i^K \delta_{x_i}$. Denote by $i_0 = min\{i | x_i \notin min(X)\}$. Assume that the function v is renormalized as in the proof of Corollary 1, which has no impact on the concavity of $p \to v(x,p) + (v(y,\vec{1}) - v(y,p))$. With such a normalization $v(x,p)$ is constant and independent of p (except for $p = \vec{0}$) on $min(X)$, being equal to 0 or $-\infty$ depending on whether v is bounded from below. We can reorder the sum in (2) to obtain, for $K = L, M$:

$$U(K) = v(x_{i_0}, \sum_{j=i_0}^{n} p_j^K) + \sum_{i=i_0+1}^{n} \left(v(x_i, \sum_{j=i}^{n} p_j^K) - v(x_{i-1}, \sum_{j=i}^{n} p_j^K) \right) \tag{12}$$

We know that $p \to v(x,p) + (v(y,\vec{1}) - v(y,p))$ is concave for all $x > y$. Taking $y \in min(X)$ we obtain that $p \to v(x_{i_0}, p)$ is concave. Moreover if $y \notin min(X)$ we have $p \to v(x,p) - v(y,p)$ which is concave. Thus the utility function, written as in (12), is a sum of functions that are concave in probabilities, implying that (11) holds.

The representation (8) corresponds to the particular case where $v(x,p) = u(x)G(p)$. Then, for any $x > y$, the quasiconcavity or concavity of $v(x,p) + (v(y,\vec{1}) - v(y,p))$ is equivalent to those of G. ■

5.2 Comparative ambiguity aversion

A difficulty for comparing ambiguity is to think of decision problems where the ambiguity dimension can be separated from the risk dimension. A possibility, introduced in Gajdos, Hayashi, Tallon and Vergnaud (2008), involves considering acts that assume at most two possible outcomes. Formally for any $x, y \in X$ denote by $\mathcal{F}_{x,y}$ the set of acts whose supports are included in $\{x, y\}$ and by $\mathcal{F}_{x,y}^c = \mathcal{F}_{x,y} \cap \mathcal{F}^c$. Then a definition of comparative ambiguity aversion can be provided as follows:

Definition 3 (comparative ambiguity aversion) *Consider $\succeq^A$ and $\succeq^B$ two preference relations on $\mathcal{F}$. We will say that the preference relation $\succeq^A$ exhibits greater ambiguity aversion* *than* $\succeq^B$ *if and only if for all* $x, y \in X$ *and any* $L \in \mathcal{F}^c_{x,y}$ *we have:*

$$\{M \in \mathcal{F}_{x,y} | M \succeq^A L\} \subset \{M \in \mathcal{F}_{x,y} | M \succeq^B L\}. \tag{13}$$

Then we have the following result:

Proposition 5 *Consider two preference relations* $\succeq^A$ *and* $\succeq^B$*, fulfilling Axioms 1 to 7, and represented as in (8) with functions* u^A *and* G^A *for the representation of* $\succeq^A$ *and* u^B *and* G^B *for the representation of* $\succeq^B$*. Then, the preference relation* $\succeq^A$ *exhibits more ambiguity aversion than the preference relation* $\succeq^B$ *if and only if* $G^A(p) \leq G^B(p)$ *for all* $p \in [0,1]^\Omega$ *and* $u^B = \phi(u^A)$ *for some increasing function* ϕ *.*

Proof. Consider two elements $x, y \in X$ with $\Delta_x \succeq^A \Delta_y$. For any $p \in [0,1]^\Omega$, applying representation (3) we obtain:

$$U^A(p\delta_x + (1-p)\delta_y) = u^A(x)G^A(p) + (1 - G^A(p))u^A(y)$$

Consider $q \in [0,1]$. We have $(\vec{q}\delta_x + (1-\vec{q})\delta_y) \in \mathcal{F}^c_{x,y}$ and

$$\left(p\delta_x + (\vec{1}-p)\delta_y \succeq \vec{q}\delta_x + (\vec{1}-\vec{q})\delta_y\right) \Leftrightarrow \left(G^A(p) - q\right)\left(u^A(x) - u^A(y)\right) \geq 0$$

Now, assume first that we have $\Delta_y \succ^B \Delta_x$ and let us show that $\succeq^B$ cannot exhibit more ambiguity aversion than $\succeq^A$. Consider $q_1, q_2 \in (0,1)$ with $q_1 > q_2$. We have $U^A(\vec{q_1}\delta_x + (1-\vec{q_1})\delta_y) \geq U^A(\vec{q_2}\delta_x + (1-\vec{q_2})\delta_y)$ and $U^B(\vec{q_1}\delta_x + (1-\vec{q_1})\delta_y) < U^B(\vec{q_2}\delta_x + (1-\vec{q_2})\delta_y)$, contradicting (13). Thus for $\succeq^A$ and $\succeq^B$ to be comparable in terms of ambiguity aversion they must generate the same rankings of deterministic acts, which implies that $u^B = \phi(u^A)$ for some increasing function ϕ. Assume now that it is the case. Using $\Delta_x \succeq^B \Delta_y$, we compute:

$$U^B(p\delta_x + (\vec{1}-p)\delta_y) - U^B(\vec{q}\delta_x + (\vec{1}-\vec{q})\delta_y) = \left(u^B(y) - u^B(x)\right)\left(G^B(p) - q\right)$$

It directly follows that $\succeq^A$ exhibits more ambiguity aversion than $\succeq^B$ if and only if $G^A(p) \leq G^B(p)$ for all $p \in [0,1]^\Omega$. ■

Proposition 5 first shows that two agents may be comparable in terms of ambiguity aversion only if they have the same ranking over deterministic acts. In that respect, we parallel the literature on comparative risk aversion where only individuals with identical preferences over deterministic consequences can be compared in terms of risk aversion. Whenever it is the case, comparative ambiguity aversion can be assessed by comparing the functions G^A and G^B. Importantly there is no need to assume that the function u^A and u^B are the same. Thus individuals may be comparable in terms of ambiguity aversion even if they do not have the same preferences over constant acts.[14]

6 Discussion

The dual approach that is suggested is obtained by relaxing the assumption of AA monotonicity (replaced by Axiom 4) and introducing the comonotonic sure-thing principle (Axiom 5). Below, I discuss in turn these assumptions.

6.1 On AA's monotonicity assumption

In order to discuss the intuitive meaning of relaxing AA's monotonicity assumption, let us consider that states of the world can be viewed as different experts. An act $L = (l_1, ...l_i, ...)$ is a sequence of lotteries, where the lottery l_i represents the beliefs of expert i about the (random) consequences associated to act L. Ambiguity results from disagreement between experts. The DM has to compare acts knowing that a single state of the world will occur: in other words the DM knows that there is an expert whose predictions are exact. Now, let us think of a DM who is a patient having to choose between two medical treatments L and M and who consults two experts. For treatment L, both experts fully agree. However for treatment M they are divided. Expert 1 thinks that M will have exactly the same (random) impact as L. But Expert 2 thinks that M may actually lead to better health conditions than L, but may also have significant negative side effects. Assume that it occurs that from the DM's point of view, the potential positive effects that are described by Expert 2 are exactly balancing the negative ones described by this same expert. Thus the messages of Expert 1 and Expert 2 lead both to conclude that treatment M is as good as treatment L. The reasons provided to support this statement are nonetheless very different. AA-monotone ambiguity models would conclude that the decision maker should be indifferent between L and M. In contradistinction, the dual approach would conclude that L is to be preferred to M. In other words, even if the DM knows that one expert is correct and all experts say that L and M are equally good, the DM may feel safer choosing the treatment for which there is no disagreement between experts regarding the whole distribution of consequences. If we consider the case of an extremely ambiguity averse DM, his unambiguous equivalent beliefs would be formed by gathering the pessimistic aspects of Expert 1's predictions (that treatment M cannot have positive effects) and the pessimistic aspects of Expert 2 's predictions (that treatment M can

[14] In the case of AA-monotone models, a similar separation of ambiguity aversion and risk aversion is obtained with both the max-min model (see Gajdos, Hayashi, Tallon and Vergnaud 2008) and the Choquet expected utility model of Schmeidler (1989). But there are other specifications, as with that of Klibanoff, Marinacci and Mukerji (2005), where such a separation is impossible.

have negative side effects). The DM would then naturally prefer L to M.

This pessimistic aggregation of information may of course be viewed as indicating that the DM has difficulties to account for mutually exclusive events. Indeed, the DM knows that one expert is right (only one state of the world realizes). It looks then natural to think that considerations of the possible negative side effects (only predicted by Expert 2) have to go along with considerations of the possible positive effects mentioned by the same expert. I readily admit that the aggregation of beliefs proposed by the dual approach does not look logical. However, very similar illogical features occur when considering AA-monotone models.

Consider for example a similar problem with a DM who is consulting two experts to compare two medical treatments. For treatment L both agree that in all circumstance it will provide a utility of 1 to the DM. As for treatment M both experts agree that the results of the treatment depend on whether the DM has a particular gene, which cannot be observed, but is known to be present with probability 0.5. They provide, however, completely opposite views regarding the impact of the gene. Expert 1 says that M will provide the DM with utility 0 if he has the gene and utility 3 if he does not have it. Expert 2 has a symmetric evaluation: according to Expert 2, treatment M provides a utility of 3 if the DM has the gene and 0 otherwise.

Let's now consider a max-min DM with maximal ambiguity aversion. If such a DM were sure to have the gene, he would strictly prefer treatment L to M since $\min(1, 1) > \min(0, 3)$. If the DM were sure of not having the gene he would also prefer L to M, since $\min(1, 1) > \min(3, 0)$. But when the DM is told that he may or may not have the gene (two exclusive possibilities) his preference is reversed. Treatment M looks indeed better than treatment L since $\min(1, 1) < \min(\frac{0+3}{2}, \frac{3+0}{2})$. This preferences reversal does not look more, or less, rational than the one previously discussed.

In both the above examples we could think of two mutually exclusive events where the decision models that we were considering were exhibiting an apparent "monotonicity breakdown". But, as Machina (2009) insisted, ambiguity aversion is about non-separability of preferences across mutually exclusive preferences. Such kind of monotonicity breakdown is thus inherent to any ambiguity model.

In fact this is already apparent in the definition of ambiguity aversion formulated in Definition 2. Indeed for preferences to exhibit strict ambiguity aversion there must be acts L and M in $\mathcal{F}$ such that:

$$L \sim M \quad \text{and } \lambda L + (1 - \lambda)M \succ L \tag{14}$$

Assume $\lambda = \frac{1}{2}$. Writing $L = \frac{1}{2}L + \frac{1}{2}L$ and $M = \frac{1}{2}M + \frac{1}{2}M$, (14) rewrite as:

$$\begin{array}{rcl} \frac{1}{2}L + \frac{1}{2}M & \succ & \frac{1}{2}L + \frac{1}{2}L \\ \frac{1}{2}M + \frac{1}{2}M & \prec & \frac{1}{2}M + \frac{1}{2}L \end{array}$$

If one views these mixtures as representing the case where there are two mutually exclusive events occurring with probability $\frac{1}{2}$ (e.g., having the relevant gene or not), the above comparison emphasizes that strict ambiguity aversion in the sense of Definition 2 cannot be consistent with separability of preferences across these mutually exclusive events. Ambiguity models that introduce AA-monotonicity in their assumptions are in fact built upon a rather asymmetric basis, where separability of preferences across mutually exclusive events is assumed for events that occur with unknown probabilities (states of the world) but relaxed for events that occur with known probabilities.

An experimental design to test AA's monotonicity assumption Since the main contrast between the dual approach and the traditional one is the adherence to AA's monotonicity assumption, we detail a simple urn-balls experiment that could be designed to test this assumption. The principle is the one that was mentioned in the introduction. Think that we have a set of colors $\{a, b..., z\}$ (with at least three distinct colors), and we present three urns A, B and C to the DM. Composition of urn A is known and given by $(\pi_a^A, ..., \pi_z^A)$ where π_j^A denotes the proportion of balls of color j in urn A. Composition of urn B is also known and given by $(\pi_a^B, ..., \pi_z^B)$. The numbers of ball in urns A and B are both unknown (and not constrained to be equal). The urn C is obtained by pooling urns A and B into a single urn. Assume that the pay-offs given to the DM depends on the color of the single ball that will be drawn. These color specific pay-offs may of course be varied along the experiment. For each set of pay-offs, ask the DM which urn he prefers to use. The urns A, B and C generate acts that we naturally call A, B and C. The acts A and B are constant acts. But the act C is, in general, not constant. In fact C looks like $\lambda A + (1 - \lambda)B$ for some ambiguous λ. If preferences over constant acts are of the expected utility kind, AA-monotone models would predict that C cannot look strictly worse than both A and B. On the contrary, according to the dual approach it may happen that C looks strictly worse than both A or B. This would actually be systematically the case when the DM is strictly ambiguity averse and that A and B, though different, look equivalently good to him. The DM, while being indifferent between the constant acts A and B would be reluctant to choose C, which is an uncertain mixture of A and B.

6.2 On the comonotonic sure-thing principle

The comonotonic sure-thing principle is a structuring assumption introduced by Chew and Wakker (1996) and used by Chateauneuf (1999) to axiomatize rank-dependent utility. Our definition is a direct extension of that of Chateauneuf (1999) to the horse-roulette setting.

One should first remark that if there were only two possible consequences, the comonotonic sure-thing principle would be a trivial assumption. It is only in the case that there are three or more outcomes that the comonotonic sure-thing principle requires preferences to have some structure. Below, I explain the kind of experiment that could be designed to test the comonotonic sure-thing principle.

Consider an urn with four balls, out of which one is black and the others are red or blue (in unknown proportions). In the experimental game the DM will be asked to draw two balls out of the urn, and a third one if one of the two balls already drawn is black. At the end the DM will therefore always have two non-black balls, that can be of same or different colors, and zero or one black ball. We will group the results of the draw in 4 categories: "same color, no black", "different colors, no black", "same color + one black", "different colors + one black". And propose the following four type of pays-offs, generating therefore four non-constant acts:

	Pay-offs when drawing:			
Two non-black balls of:	same color	diff colors	same color	diff colors
and:	no black ball	no black ball	one black ball	one black ball
Act 1	x	y	z	t
Act 2	y	x	z	t
Act 3	x	y	t	z
Act 4	y	x	t	z

The design of this fictive experiment is done so that we only compare non-constant acts, which by construction are themselves 50/50 mixtures of non-constant acts. Indeed, we can view this experiment as asking the DM to bet on whether the drawn non-black balls will have the same color and making the pay-offs depend on whether a black ball will be drawn or not (which occurs with probability 0.5).

The comonotonic sure-thing principle indicates that if the intervals $[x, y]$ and $[z, t]$ do not overlap, then one should have:

$$(\text{Act } 1 \succeq \text{Act } 2) \Leftrightarrow (\text{Act } 3 \succeq \text{Act } 4)$$

The intuition is that Acts 3 and 4 are obtained from Acts 1 and 2 by changing their common realization on the event "a black ball is drawn". The sure-thing principle of Savage (1954) would require the equivalence to hold in all cases. We only assume that it is the case only if the events "a black ball" and "no black ball" generates acts with supports in non-overlapping intervals. One can check that models like those of Schmeidler (1989), Gilboa and Schmeidler (1989) or Klibanoff, Marinacci and Mukerji (2005) would not support this equivalence.

7 Comparison with the EUUP model

The EUUP model of Izhakian (2014a) and the current paper rely on quite different settings. Comparing the assumptions, and the reasoning of both papers would require to bridge the horse-roulette lottery setting with the "primary outcome space and a secondary probability space" setting of Izhakian (2014a), which may be quite cumbersome. The suggested specifications can however be easily compared, at least if using either a finite support version of the EUUP model, or a non-finite support version of the specification (6) derived in Proposition 2 of the current paper. Let us use this latter possibility. To deal with non finite support one has to use an integral version of (6), that is:

$$U(L) = -\int_X u(x)d(G_x(Prob(L \geq x))) \tag{15}$$

To get an expression that is closer to that of the EUUP model, one may (i) split the integral (15) in two integrals, (ii) proceed to an integration by parts in each integral and (iii) make a change of variable $y = u(x)$. Formally, for any given $x_0 \in X$:

$$\begin{aligned}
U(L) &= \int_{x \leq x_0} u(x)d(1 - G_x(Prob(L \geq x))) - \int_{x \geq x_0} u(x)d(G_x(Prob(L \geq x))) \\
&= u(x_0) - \int_{x \leq x_0} u'(x)(1 - G_x(Prob(L \geq x)))dx + \int_{x \geq x_0} u'(x)G_x(Prob(L \geq x))dx \\
&= u(x_0) - \int_{y \leq u(x_0)} (1 - G_{u^{-1}(y)}(Prob(u(L) \geq x)))dx + \int_{y \geq u(x_0)} G_{u^{-1}(y)}(Prob(u(L) \geq y))dy
\end{aligned}$$

If the function u is normalized so that $u(x_0) = 0$, we eventually obtain:

$$U(L) = \int_{y \geq 0} G_{u^{-1}(y)}(Prob(u(L) \geq y))dy - \int_{y \leq 0} (1 - G_{u^{-1}(y)}(Prob(u(L) \geq y)))dy \tag{16}$$

The EUUP model, on the other hand, corresponds to the case where:

$$U^{EUUP}(L) = \int_{y \geq 0} \Psi(Prob(u(L) \geq y))dy - \int_{y \leq 0} \Psi(Prob(u(L) \leq y))dy \tag{17}$$

The function $\Psi : [0,1]^{\Omega} \to [0,1]$ is given by:

$$p = (p_1, ..., p_{\Omega}) \to \Psi(p) = \Gamma^{-1}(\sum_{i=1}^{\Omega} \pi_i \Gamma(p_i))$$

where the function $\Gamma : [0,1] \to [0,1]$ is an increasing function and $(\pi_1, ..., \pi_{\Omega})$ a subjective probability distribution over $S = \{1, ..., \Omega\}$.

It is then clear that that the EUUP model shown in (17) is a particular case of (16) where the function G_x is given by:

$$\begin{cases} G_x(p) = \Psi(p) & \text{for all } x \geq x_0 \\ G_x(p) = 1 - \Psi(\vec{1} - p) & \text{for all } x \leq x_0 \end{cases}$$

A particularity of this specification, is that assuming that the function $p \to G_x(p)$ is quasiconcave when $x \geq x_0$ would imply that it is quasiconvex for $x \leq x_0$, and vice versa. It then directly follows from Proposition 4 that the standard EUUP specification does not exhibit ambiguity aversion in the sense of Definition 2, which is that of Schmeidler (1989). In other words, the model used in Izhakian (2014b), and other related applications, is a dual model, but not an ambiguity aversion model in the usual sense.

There are several ways to adapt the EUUP model so that it exhibits ambiguity aversion in the sense of Schmeidler. A first possibility involves assuming that the "anchoring point" x_0 is a minimal element of X, so that all consequences are "favorable". In that case, the second term in (17) disappears, and the EUUP model is equivalent to the dual KMM model shown in Table 1. Then, if Γ is such that the function Ψ is concave, preferences exhibit ambiguity aversion.[15] A second possibility involves using the asymmetric representation provided in Theorem 2 of Izhakian (2014a) and assuming a well chosen concave function Γ_{FV} and a well chosen convex function Γ_{UF}.[16] Interestingly, ambiguity aversion in the sense of Schmeidler is then obtained when the DM would be "ambiguity averse for gains and ambiguity loving for losses" in the terminology of Izhakian. The notion of ambiguity aversion discussed in Izhakian (2014a) is clearly different from that of Schmeidler.

The message of the current paper is that ambiguity aversion in the sense of Schmeidler may be obtained by aggregating probabilities of events in a specific way. But a key principle is that this aggregation has to depend on the consequences associated to the events. This is what distinguishes ambiguity aversion from the "support theory" of Tversky and Koehler

[15]The function Ψ is concave if the function $q \to \frac{-\Gamma'(q)}{\Gamma''(q)}$ is concave (or linear). See, e.g., Proposition 82 in Gollier (2004).

[16]The notation Γ_{FV} and Γ_{UF} correspond to those of Izhakian. The dual KMM model of Table 1 corresponds to the case where $\Gamma_{FV}(q) = 1 - \Gamma_{UF}(1-q)$.

(1994) and Rotttenstreich and Tversky (1997). The AA-monotone and the dual ambiguity models suggested in the current paper fully agree with that principle. They only diverge in the way it is implemented.

A last point to be mentioned, is that specifying the function G_x in the dual model is not necessarily needed to discuss the impact of ambiguity aversion. The section below provides an example where a general result is obtained without having to opt for a particular dual specification.

8 An application to saving behaviors

We consider the impact of an increase in ambiguity aversion in a standard two-period saving problem. Denote y_1 income in period 1 and $\tilde{y}_2$ the ambiguous uncertain period 2 income. The rate of interest is r, not random. Agents have to chose their saving $s(y_1, \tilde{y}_2, r)$ which implies consumption $c_1 = y_1 - s(y_1, \tilde{y}_2, r)$ in period 1 and $c_2 = y_2 + (1+r)s(y_1, \tilde{y}_2, r)$ in period 2, when y_2 realizes. To be consistent with our theoretical setting we assume that $\tilde{y}_2$ has a finite support. We denote again by $S = \{1, ...\Omega\}$ the states of the world. For any $y \in \mathbb{R}_+$ we denote by $Prob(\tilde{y}_2 = y) \in [0,1]^\Omega$ the vector of probabilities given by:

$$Prob(\tilde{y}_2 = y) = (Prob(\tilde{y}_2 = y|1), ..., Prob(\tilde{y}_2 = y|\Omega))$$

where $Prob(\tilde{y}_2 = y|\omega)$, is the probability to have a second period income equal to y if state of the world ω occurs. The question we are interested in is the following one: how does an increase in ambiguity aversion impact savings?

In this setting the set of consequences is the set of feasible two periods consumptions paths $X \subset \mathbb{R}^2_+$. We consider two agents A and B with preferences fulfilling Axioms 1 to 7 and who only differ by their degree of ambiguity aversion. That means that A and B's preferences can be represented as in (8) with the same function u, but possibly different functions G, denoted by G^A and G^B. We assume that preferences over constant acts are additively separable and exhibit risk aversion: $u(x_1, x_2) = u_1(x_1) + u_2(x_2)$ where u_1 and u_2 are increasing and concave.

Proposition 6 (Ambiguity aversion and saving behaviors) *Under the above assumptions, if A is more ambiguity averse than B in the sense of Definition 3, then A saves at least as much as B.*

Proof. For simplicity sake we set $r = 0$ in the proof. The case where $r \neq 0$ is totally similar. Denote by $\{y_2^1, ...y_2^n\}$ the support of $\tilde{y}_2$, where the y_2^j are ranked in an increasing order. For $i = A, B$, denote by $U^i(s)$ the utility of agent i when saving s. We have

$$U^i(s) = \sum_{j=1}^{n} u(y_1 - s, y_2^i + s)\left(G^i(\sum_{k=j}^{n} Prob(\tilde{y}_2 = y_2^k)) - G^i(\sum_{k=j+1}^{n} Prob(\tilde{y}_2 = y_2^k))\right)$$

where the convention $\sum_{k=n+1}^{n} = 0$ is used. Reordering the sum we have:

$$U^i(s) = u_1(y_1 - s) + u_2(y_2^1 + s) + \sum_{k=2}^{n} G^i(\sum_{k=j}^{n} Prob(\tilde{y}_2 = y_2^k))\left(u_2(y_2^j + s) - u_2(y_2^{j-1} + s)\right)$$

Denote by s_A the optimal saving of A and by s_B that of B. Let us assume that $s_A < s_B$ and show that it leads to a contradiction. By the optimality of A's saving, one must have:

$$u_1(y_1 - s_A) + u_2(y_2^1 + s_A) + \sum_{j=2}^{n} G^A(\sum_{k=j}^{n} Prob(\tilde{y}_2 = y_2^k))\left(u_2(y_2^j + s_A) - u_2(y_2^{j-1} + s_A)\right) \quad >$$
$$u_1(y_1 - s_B) + u_2(y_2^1 + s_B) + \sum_{j=2}^{n} G^A(\sum_{k=j}^{n} Prob(\tilde{y}_2 = y_2^k))\left(u(, y_2^j + s_B) - u_2(y_2^{j-1} + s_B)\right)$$

A symmetric inequality can be written associated to the optimality of $B's$ saving. By addition of these two inequalities one gets, after a simple factorization:

$$\sum_{j=2}^{n} H_j(s_A, s_B)\Delta G(\sum_{k=j}^{n} Prob(\tilde{y}_2 = y_2^k)) > 0$$

where

$$H_j(s_A, s_B) = \left(u_2(y_2^j + s_B) - u_2(y_2^{j-1} + s_B)\right) - \left(u_2(y_2^j + s_A) - u_2(y_2^{j-1} + s_A)\right)$$

and $\Delta G = G^B - G^A$. The contradiction follows from the fact that $\Delta G(\sum_{k=j}^{n} Prob(\tilde{y}_2 = y_2^k)) \geq 0$ because A is more ambiguity averse than B and $H_j(s_A, s_B) \leq 0$ because s_B was assumed to be larger than s_A and u_2 is concave. ■

Proposition 6 provides a simple and intuitive result. Ambiguity aversion increases savings. It is interesting to compare this result with what Osaki and Schlesinger (2014) derive using the model of Klibanoff, Marinacci and Mukerji (2005). Osaki and Schlesinger (2014) find that ambiguity aversion may have a positive or a negative impact on savings. They are able to establish that there exists a positive link between ambiguity aversion and precautionary savings only in one very particular case: that of agents with constant absolute ambiguity aversion and beliefs that can be ranked in terms of second order stochastic dominance. Addressing the same problem with the max-min model of Gilboa and Schmeidler (1989) would provide a similar message: that of ambiguity aversion having an ambiguous impact on savings, unless some strong assumptions are introduced regarding how the state-specific income distributions compare. In contradistinction, Proposition 6 says that greater ambiguity aversion should generate greater savings. The result is obtained without assuming specific parametric forms for the functions G^A and G^B and, more importantly, without making assumptions on how the state-specific income distributions compare.

9 Conclusion

Ellsberg's experiment created a significant challenge for economists: that of formulating mathematical models that would explain an apparently illogical behavior. Some of the assumptions of Savage or of AA had to be relaxed. While the "the sure-thing principle" and the "independence axiom" became the usual suspects, other "natural assumptions", like that of AA-monotonicity remained mostly unquestioned. In most cases, the assumption of monotonicity in the sense of AA is not really commented or just viewed as obvious. For example Gilboa, Maccheroni, Marinacci, and Schmeidler (2010) interpret AA-monotonicity as an "axiom of rationality", like transitivity of the preference relation. In other cases, the assumption is justified for being "quite common" (i.e., in Gilboa and Schmeidler, 1989) or "standard" (i.e., in Siniscalchi 2009 and in Maccheroni, Marinacci, and Rustichini 2006).

AA-monotonicity implies however a property of separability of preferences across mutually exclusive events. Thus, it precisely embeds the kind of assumptions that Machina (2009) sees as contradicted by Ellsberg's experiment. Why then maintaining it when trying to model ambiguity aversion? As explained in Section 6.1, AA-monotone ambiguity models rely on an asymmetric approach that involves maintaining separability of preferences across mutually exclusive "subjective events" (i.e. events that occur with an unknown probabilities), but relaxing the assumption of separability of preferences across mutually exclusive "objective events".

The current paper explores the possibility to replace AA's monotonicity assumption by a weaker one (monotonicity with respect to first-order stochastic dominance) which does not imply state separability. Combined with the comonotonic sure-thing principle, that provides a simple model where ambiguity aversion looks like a very intuitive form of pessimism in beliefs aggregation.

There is no doubt that the suggested dual approach suffers for potential shortcomings. Firstly, it relies on the horse-roulette lottery setting, which by itself is disputable.[17] Secondly, just like AA-monotonicity, the comonotonic sure-thing principle could be viewed as imposing

[17]See for example the discussion in Eichberger and Kelsey (1996).

excessive structure to the model. Thirdly, when applied to simple two-outcome experiments, it provides models that are formally indistinguishable with the standard AA-monotone models, and as a consequence could be subject to the same criticisms.

In my view, the merit of dual approach should not be searched in some "alleged superiority" in the modeling of ambiguity aversion. There is -so far- no piece of empirical evidence that would support such a statement, or the opposite one. Instead, I prefer to look at it as an intriguing contribution that shows how the very same empirical puzzle (that provided by Ellsberg's experiment) could have led to develop different frameworks, equally good in responding to this puzzle, but providing at the end very different conclusions about the impact of ambiguity.

References

Aczél, J. (2006): *Lectures on functional equations and their applications.* Courier Dover Publications.

Anscombe, F. J., and R. J. Aumann (1963): "A Definition of Subjective Probability," *Annals of Mathematical Statistics*, 34(1), 199–205.

Bordley, R. F., and R. W. Wolff (1981): "Note: On the Aggregation of Individual Probability Estimates," *Management Science*, 27(8), 959–964.

Cerreia-Vioglio, S., P. Ghirardato, F. Maccheroni, M. Marinacci, and M. Siniscalchi (2011): "Rational Preferences under Ambiguity," *Economic Theory*, 48(2-3), 341–375.

Chateauneuf, A. (1999): "Comonotonicity Axioms and Rank-Dependent Expected Utility Theory for Arbitrary Consequences," *Journal of Mathematical Economics*, 32(1), 21–45.

Chew, S.-H., and P. Wakker (1996): "The Comonotonic Sure-Thing Principle," *Journal of risk and uncertainty*, 12(1), 5–27.

Eichberger, J., and D. Kelsey (1996): "Uncertainty Aversion and Preference for Randomisation," *Journal of Economic Theory*, 71(1), 31–43.

Fishburn, P. C. (1970): *Utility Theory for Decision Making.* John Wiley & Sons.

Gajdos, T., T. Hayashi, J.-M. Tallon, and J.-C. Vergnaud (2008): "Attitude Toward Imprecise Information," *Journal of Economic Theory*, 140(1), 27–65.

Gilboa, I., F. Maccheroni, M. Marinacci, and D. Schmeidler (2010): “Objective and Subjective Rationality in a Multiple Prior Model,” *Econometrica*, 78(2), 755–770.

Gilboa, I., and D. Schmeidler (1989): “Maxmin Expected Utility with a Non-Unique Prior,” *Journal of Mathematical Economics*, 18(2), 141–153.

Gollier, C. (2004): *The Economics of Risk and Time.* MIT press.

Green, J. R., and B. Jullien (1988): “Ordinal Independence in Nonlinear Utility Theory,” *Journal of Risk and Uncertainty*, 1(4), 355–387.

Guetlein, M.-C. (2014): “Comparative Risk Aversion in the Presence of Ambiguity,” paper presented at the 2014 meeting of the european econometric society.

Izhakian, Y. (2012): “Ambiguity Measurement,” Working Paper EC-12-01, Stern School of Business, N.Y.U.

——— (2014a): “Expected Utility with Uncertain Probabilities Theory,” Paper downloadable on http://people.stern.nyu.edu/yizhakia/papers/euup.pdf.

——— (2014b): “A Theoretical Foundation of Ambiguity Measurement,” http://people.stern.nyu.edu/yizhakia/papers/ambgmsr.pdf.

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

Maccheroni, F., M. Marinacci, and A. Rustichini (2006): “Ambiguity Aversion, Robustness, and the Variational Representation of Preferences,” *Econometrica*, 74(6), 1447–1498.

Machina, M. J. (2009): “Risk, Ambiguity, and the Rank-Dependence Axioms,” *The American Economic Review*, 99(1), 385–392.

——— (2014): “Ambiguity Aversion With Three or More Outcomes,” *The American Economic Review*, 104(12), 3814–3840.

McConway, K. J. (1981): “Marginalization and Linear Opinion Pools,” *Journal of the American Statistical Association*, 76(374), 410–414.

Osaki, Y., and H. Schlesinger (2014): “Precautionary Saving and Ambiguity,” paper presented at the cear/mric behavioral insurance conference in atlanta.

Rottenstreich, Y., and A. Tversky (1997): “Unpacking, Repacking, and Anchoring: Advances in Support Theory.,” *Psychological review*, 104(2), 406.

SAVAGE, L. J. (1954): *The Foundations of Statistics.* John Wiley & Sons.

SCHMEIDLER, D. (1989): "Subjective Probability and Expected Utility Without Additivity," *Econometrica*, 57(3), 571–587.

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

TVERSKY, A., AND D. J. KOEHLER (1994): "Support Theory: a Nonextensional Representation of Subjective Probability.," *Psychological Review*, 101(4), 547.

WAGNER, C. (1982): "Allocation, Lehrer Models, and the Consensus of Probabilities," *Theory and Decision*, 14(2), 207–220.

WAKKER, P. (1993): "Additive Representations on Rank-Ordered Sets: II. The Topological Approach," *Journal of Mathematical Economics*, 22(1), 1–26.

A Proof of Proposition 1

The proof may be attributed to Chateauneuf (1999) as only minor variations have to introduced to adapt his proof to the current setting. We shall moreover note that Chateauneuf's proof uses a key result of Wakker (1993) showing the consequences of assuming separability on rank-ordered sets.

This being said, let us rewrite and reshape the proof of Chateauneuf to conform with the current setting. To simplify the proof we will focus on the set of acts for which all probabilities are in $\mathbb{Q}$ the set of rational numbers. The result then extends to the case where probabilities are not in $\mathbb{Q}$ by continuity. For simplicity sake we will use the notation $(0,1)_{\mathbb{Q}}$ and $[0,1]_{\mathbb{Q}}$ to mean, respectively, $(0,1) \cap \mathbb{Q}$ and $[0,1] \cap \mathbb{Q}$.

For any integer $n > 0$, let us denote by $\mathcal{M}_n$ the set of matrices of size $n \times \Omega$ (i.e., with n rows and Ω columns) where each column is a probability vector. Formally:

$$\mathcal{M}_n = \{(p_{i,j})_{\substack{1 \le i \le n \\ 1 \le j \le \Omega}} \in [0,1]_{\mathbb{Q}}^{n\Omega} | \forall j : \sum_{i=1}^{n} p_{ij} = 1\}$$

And denote by $\mathcal{M}_n^*$ the subset of $\mathcal{M}_n$ of such matrices with no raw equal to $\vec{0}$.

To use notation similar to those of Chateauneuf (1999), let us introduce the set $X_\uparrow^n$ defined by:

$$X_\uparrow^n = \{(x_i)_{1 \le i \le n} \in X^n | x_n \ge x_{n-1} \ge ... \ge x_1\}$$

For any $p = (p_{i,j}) \in \mathcal{M}_n$ and any $x = (x_i) \in X^n_\uparrow$ corresponds a single a act

$$(\sum_{i=1}^{n} p_{i1}\delta_{x_i}, ..., \sum_{i=1}^{n} p_{i\Omega}\delta_{x_i}) \in \mathcal{F}$$

For simplicity sake we denote by $p.x$ such an act.

For any $p \in \mathcal{M}^*_n$ we define the binary relation $\succeq_p$ over $X^n_\uparrow$ by:

$$x \succeq_p y \Leftrightarrow p.x \succeq p.y$$

The preference relation $\succeq_p$ is completely analogous to that found in Chateaueuf (1999), at the difference that in Chateauneuf (1999) p is a probability vector, while here it is a matrix that has Ω columns, with each column being a probability vector. This has no impact on the Lemmas 1 to 3 and the Proposition 1 of Chateaueuf (1999) that still apply. It follows for any $n \geq 3$ and any $p \in \mathcal{M}^*_n$ the relation of preferences $\succeq_p$ admit an additive representation. Formally there exist continuous jointly cardinal monotonic value functions $V_i(.,p) : X \to \overline{\mathbb{R}}$ such that

$$x \succeq_p y \Leftrightarrow \sum V_i(x_i, p) \geq \sum V_i(y_i, p) \tag{18}$$

Moreover the functions $V_i(., p)$ are finite, except possibly $V_1(., p)$ on $max(X)$ and $V_n(., p)$ on $min(X)$. For all $x = (x_i) \in X^n_\uparrow$, let us denote $V_p(x) = \sum V_i(x_i, p)$.

As in Chateaueuf (1999) we can notice that, since any two-outcome act can be seen as a three-outcome act, the existence of an additive representation extends to the case where $n = 2$.

Using non-triviality, on the one hand and the connectedness of X, on the other hand, we know that there are two elements a and b in X such that $a > b$. As additive representation are defined up to a positive affine transformation it can be assumed that, for all p a normalization was chosen such that $V_p(a) = 1$ and $V_p(b) = 0$, where a and b stands here for the constant vectors $(a, ..., a)$ and $(b, ..., b)$.

As we focus on simple acts, and that $\mathbb{Q}$ is dense into $\mathbb{R}$, the set of acts $\mathcal{F}$ is homoeorphic to the closure of $\cup_{n\geq 1}\mathcal{F}_n$ where:

$$\mathcal{F}_n = \{p.x | p \in \mathcal{M}^*_n \text{ and } x \in X^n_\uparrow\}.$$

The relation of preferences $\succeq$ generates preferences relation on $\cup_{n\geq 1}\mathcal{F}_n$ and thus, a fortiori, on each $\mathcal{F}_n$. Let us show, as in Chateauneuf's paper, that the restriction of $\succeq$ to $\mathcal{F}_n$ is represented $p.x \to V_p(x)$.

First, for any p and p' in $\mathcal{M}^*_n$ there exist a $n_0 \in \mathbb{N}$ such that $n_0 p$ and $n_0 p'$ only include integer numbers. Denote q_0 the matrix of size $n_0 \times \Omega$ where all elements are equal to $\frac{1}{n_0}$. Any element $p.x$ can be then be viewed as an element $q_0.\pi_p(x)$ for some $\pi_p(x) \in X_\uparrow^{n_0}$. Similarly, any $p'.x$ can be viewed as an element $q_0.\pi_{p'}(x)$ for for some $\pi_{p'}(x) \in X_\uparrow^{n_0}$. Moreover if x is a constant vector equal to $(x, ..., x)$, so must be $\pi_p(x)$ and $\pi_{p'}(x)$. Since V_{q_0} was normalized so that $V_{q_0}(a) = 1$ and $V_{q_0}(b) = 0$ we must have for constant vectors $V_{q_0}(\pi_p(x)) = V_p(x)$ and $V_{q_0}(\pi_{p'}(x)) = V_{p'}(x)$. It follows that $V_p(x) = V_{p'}(x)$ for all constant vectors x.

Now consider two elements $p.x$ and $p'.x'$ in $\mathcal{F}_n$. By continuity we now that there exists x_0 and x'_0 in X such that $\Delta_{x_0} \sim p.x$ and $\Delta_{x'_0} \sim p'.x'$. This implies that $V_p(x) = V_p(x_0)$ and $V_{p'}(x') = V_{p'}(x'_0)$ where x_0 and x'_0 denote here the constant vectors $(x_0, ..., x_0)$ and $(x'_0, ..., x'_0)$. Since these are constant vectors we have:

$$p.x \succeq p'.x' \Leftrightarrow \Delta_{x_0} \succeq \Delta_{x'_0} \Leftrightarrow V_{p'}(x_0) \geq V_{p'}(x'_0) \Leftrightarrow V_p(x_0) \geq V_{p'}(x'_0) \Leftrightarrow V_p(x) \geq V_{p'}(x')$$

which shows that the restriction of $\succeq$ to $\mathcal{F}_n$ is represented $p.x \rightarrow V_p(x)$.

It remains to show that $V_p(x)$ can be written as in (2).

For any vector of probabilities $\pi = (\pi_j) \in [0,1]_{\mathbb{Q}}^{\Omega}$, that is different from the constant vectors $\vec{0}$ and $\vec{1}$ we associate $p^\pi = (p^\pi_{ij}) \in \mathcal{M}_2^*$ defined by $p^\pi_{1j} = \pi_j$ and $p^\pi_{2j} = \vec{1} - \pi_j$. Define then the function $X \times [0,1]_{\mathbb{Q}}^{\Omega} \rightarrow \mathbb{R}$ by:

$$v(x, \pi) = V_2(x, p^\pi)$$

We then extend this function to the cases where $\pi = \vec{0}$ and $\pi = \vec{1}$ by setting:

$$\begin{aligned} v(x, \vec{0}) &= 0 \\ v(x, \vec{1}) &= V_1(x, p) + V_2(x, p) \text{ for any } p \in \mathcal{M}_2^* \end{aligned}$$

Remark that in the above definition the notation x is used to mean both an element of $x \in X$ (in $v(x, \pi)$) or to mean the constant vector with all coordinates equal to x (in $V_i(x, p)$).

We now show that the relation of preferences $\succeq$ can be represented by equation (2). By continuity, we have to show that it is the that case for its restriction to any of the $\mathcal{F}_n$. We we can proceed by induction, following exactly the proof of Chateauneuf (1999). For $n = 2$ and any $p \in \mathcal{M}_2^*$ and any $(x_1, x_2) \in X_\uparrow^2$ with $x_1 \neq x_2$.

$$\begin{aligned} V_p(x) &= V_1(x_1, p) + V_2(x_2, p) = v(x_1, \vec{1}) + v(x_2, p^\pi) - v(x_1, p^\pi). \\ &= \sum_{i=1}^{2} v(x_i, Prob(p.x \geq x_i)) - v(x_i, Prob(p.x > x_i)) \end{aligned}$$

where we recognize representation (2). In the case where $x_1 \sim x_2$ we get $V_p(x) = V_1(x_1, p) +$

$V_2(x_2, p) = v(x_1, \vec{1})$ where we also recognize representation (2).

Now let us assume that representation (2) was shown to hold on all $\mathcal{F}_k$ for $k \leq n$ and show that it holds on $\mathcal{F}_{n+1}$. Consider $p \in \mathcal{M}_n^*$ and $x = (x_1, ..., x_{n+1}) \in X_{\uparrow}^{n+1}$. If $x_j \sim x_{j+1}$ for some $j \in \{1, ..., n\}$ we can view $p.x$ as an element of $\mathcal{F}_n$ for which the result is assumed to hold. Let us therefore assume that $x_{j+1} > x_j$, for all $j \in \{1, ..., n\}$. By continuity there exists $\hat{x}_n$ such that:

$$p.(x_1, ...x_{n-1}, x_n, x_{n+1}) \sim p.(x_1, ..., x_{n-1}, \hat{x}_n, \hat{x}_n) \tag{19}$$

Using the comonotonic sure-thing principle we can replace x_{n-1} by x_n in the above equivalence to get:

$$p.(x_1, ..., x_n, x_n, x_{n+1}) \sim p.(x_1, ..., x_n, \hat{x}_n, \hat{x}_n)$$

This equivalence relates elements $\mathcal{F}_{n+1}$ that can also be viewed as elements of $\mathcal{F}_n$. We can then use representation specification (2) which provides:

$$v(x_n, p_n + p_{n+1}) - v(x_n, p_{n+1}) + v(x_{n+1}, p_{n+1}) = v(\hat{x}_n, p_n + p_{n+1}) \tag{20}$$

From (19) we know that:

$$U(p.(x_1, ...x_{n-1}, x_n, x_{n+1})) = U(p.(x_1, ..., x_{n-1}, \hat{x}_n, \hat{x}_n)) \tag{21}$$

Applying (2) to compute $U(p.(x_1, ..., x_{n-1}, \hat{x}_n, \hat{x}_n))$ we get:

$$\begin{aligned} U(p.(x_1, ..., x_{n-1}, \hat{x}_n, \hat{x}_n)) &= \sum_{i=1}^{n-1} v(x_i, Prob(p.(x_1, ..., x_{n-1}, \hat{x}_n, \hat{x}_n) \geq x_i)) \\ &\quad - \sum_{i=1}^{n-1} v(x_i, Prob(p.(x_1, ..., x_{n-1}, \hat{x}_n, \hat{x}_n) > x_i)) \\ &+ \quad v(\hat{x}_n, p_n + p_{n+1}) \end{aligned}$$

With (20) and (21) that provides:

$$U(p.x) = \sum_{i=1}^{n+1} \left(v(x_i, Prob(p.x \geq x_i)) - v(x_i, Prob(p.x > x_i))\right)$$

which show that representation (2) is obtained. The properties of the function v that are listed in Definition 1 are totally similar to those listed in Chateauneuf (1999). These properties can be shown to hold exactly as is done in Chateauneuf's paper.