Attitude toward imprecise information ⋆

T. Gajdos [a], T. Hayashi [b], J.-M. Tallon [a,*], J.-C. Vergnaud [a]

[a] *Université Paris I, Panthéon-Sorbonne & Ecole d'Economie de Paris, 106 Bd de l'Hôpital, 75013 Paris, France*

[b] *Department of Economics, University of Texas at Austin, Austin, TX 78712, USA*

Abstract

This paper presents an axiomatic model of decision making under uncertainty which incorporates objective but imprecise information. Information is assumed to take the form of a probability-possibility set, that is, a set P of probability measures on the state space. The decision maker is told that the true probability law lies in P and is assumed to rank pairs of the form (P, f) where f is an act mapping states into outcomes. The key representation result delivers maxmin expected utility where the min operator ranges over a set of probability priors –just as in the maxmin expected utility (MEU) representation result of Gilboa and Schmeidler (1989). However, unlike the MEU representation, the representation here also delivers a mapping, φ, which links the probability-possibility set, describing the available information, to the set of revealed priors. The mapping φ is shown to represent the decision maker's attitude to imprecise information: under our axioms, the set of representation priors is constituted as a selection from the probability-possibility set. This allows both expected utility when the selected set is a singleton and extreme pessimism when the selected set is the same as the probability-possibility set, i.e. , φ is the identity mapping. We define a notion of comparative imprecision aversion and show it is characterized by inclusion of the sets of revealed probability distributions, irrespective of the utility functions that capture risk attitude. We also identify an explicit attitude toward imprecision that underlies usual hedging axioms. Finally, we characterize, under extra axioms, a more specific functional form, in which the set of selected probability distributions is obtained by (i) solving for the "mean value" of the probability-possibility set, and (ii) shrinking the probability-possibility set toward the mean value to a degree determined by preferences.

Key words: Imprecise information, imprecision aversion, multiple priors, Steiner point.
JEL: D81.

1 Introduction

In many problems of choice under uncertainty, some information is available to the decision maker. Yet, this information is often far from being sufficiently precise to allow the decision maker to come up with an estimate of a probability distribution over the relevant states of nature. The archetypical example of such a situation is the so-called Ellsberg paradox (Ellsberg (1961)), in which subjects are given some imprecise information concerning the composition of an urn and are then asked to choose among various bets on the color of a ball drawn from that urn. The usual findings in this experiment is that most subjects' choices can not be rationalized if one assumes they hold probabilistic beliefs. Rather, one should assume that subjects have "multiple beliefs". In a very influential paper, Gilboa and Schmeidler (1989) axiomatized the following decision criterion, that allows to "solve" Ellsberg paradox: an act f is preferred to an act g if and only if there exists a set of priors $\mathcal{C}$ and a utility function u such that

$$\min_{p\in\mathcal{C}} \int u \circ f dp \geq \min_{p\in\mathcal{C}} \int u \circ g dp.$$

In this functional form, $\mathcal{C}$ is usually interpreted as the decision maker's set of beliefs. To go back to Ellsberg experiment, *if* one takes as the set of beliefs the distributions of the balls compatible with the information given to the subjects, then the functional with that set captures *via* the min operator an extreme form of pessimism.

However, nothing in the theoretical construct of Gilboa and Schmeidler (1989) supports this interpretation since the information the decision maker possesses

⋆ This paper is the result of merging two papers that were developed independently. One by Hayashi, "Information, Subjective Belief and Preference" available at http://www.eco.utexas.edu/∼th925/info_belief11.pdf and one by Gajdos, Tallon, and Vergnaud, "Coping with Imprecise Information: a Decision Theoretic Approach" available at http://eurequa.univ-paris1.fr/membres/tallon/GTV-coping-revision.pdf. We thank D. Ahn, N. Al Najjar, M. Cohen, E. Dekel, L. Epstein, T. Gilboa, R. Giraud, F. Gul, J.Y. Jaffray, P. Klibanoff, B. Lipman, M. Machina, M. Marinacci, S. Mukerji, M. Scarsini, M. Stinchcombe, and B. Walliser for useful comments. Comments by seminar audience at Rochester, Concordia, Texas-Austin, INSEAD, Toulouse University, University Paris I, the CERMSEM symposium, FUR XI, RUD 2004, Roy seminar, GREQAM, and the Cowles foundation workshop on decision under uncertainty are gratefully acknowledged. Comments by an anonymous referee and the associate editor have been useful to improve the paper. Financial support from the French Ministry of Research (Action concertée incitative) and the Ministry of Environment (S3E) is gratefully acknowledged.

* Corresponding author.
Email addresses: thibault.gajdos@univ-paris1.fr (T. Gajdos), th925@eco.utexas.edu (T. Hayashi), jean-marc.tallon@univ-paris1.fr (J.-M. Tallon), jean-christophe.vergnaud@univ-paris1.fr (J.-C. Vergnaud).

is not part of the model (although it is part of the motivation as, to the best of our knowledge, in all papers on ambiguity aversion.) In this paper, we explicitly incorporate the information the decision maker has as a primitive of the model [1] in order to further explore properties of Gilboa and Schmeidler's "set of priors".

The decision maker is assumed to have preferences over couples (P, f) where P is a set of probability distributions over the state space (hereafter *probability-possibility sets*) and f is an act. In an Anscombe and Aumann (1963) framework, we axiomatize the following functional form: for two probability-possibility sets P and Q and two acts f and g, (P, f) is preferred to (Q, g) if and only if

$$\min_{p \in \varphi(P)} \int u \circ f dp \geq \min_{p \in \varphi(Q)} \int u \circ g dp.$$

For a fixed probability-possibility set, this is identical to maxmin expected utility à la Gilboa and Schmeidler (1989): in their setting, the (un-modeled) prior information that the decision maker has is given. The novel aspect here is that the link between the information possessed by the decision maker and the "set of priors" is made explicit, through the mapping denoted by φ. We will call $\varphi(P)$ the selected probability-possibility set. The properties of the function φ can be further specified, as it embeds the decision maker's attitude toward imprecision.

The objects (P, f) are not standard (although see the discussion of related literature below). That the decision maker has preferences on such pairs means that, at least conceptually, we allow decision makers to compare the same act in different informational settings. The motivation [2] for this formalization can be best understood going back to Ellsberg's two urns example. In urn 1 there is a known proportion of black and white balls (50-50) while in urn 2, the composition is unknown. The decision maker has the choice to bet on black in urn 1 or on black in urn 2. The action (bet on black) itself is "the same" in the two cases but the information has changed from a given probability distribution $(\frac{1}{2}, \frac{1}{2})$ (urn 1) to the simplex (urn 2).

We also believe that our model can be used to think of situations besides laboratory experiments. Imagine a firm in the agro business contemplating investing in various crops in different countries. Then, P and Q would capture information relative to (long term) weather forecast in different parts of the world, while f and g would capture the act of investing in a particular crop. One could also consider the example of investing in some stock in one's home country in which information is supposedly easier to acquire than in a similar

[1] We do not provide a theory of how such information can be generated: see for instance Al Najjar (2007) and examples in Lehrer (2007).
[2] We will develop these arguments in Section 2.1.

stock in some exotic country. The widespread preference to invest in home country stocks (the so-called "home bias") can thus find an illustration in our model. Finally, a particular case of our model, arguably not the most interesting one as will be discussed in section 3.3, is one of choice over sets of lotteries in which (P, f) is evaluated by the induced distributions on outcomes.

Next, we study the properties of our decision criterion. We define a notion of *comparative imprecision aversion* with the feature that it can be completely separated from risk attitudes. Loosely speaking, we say that a decision maker b is more imprecision averse than a decision maker a if whenever a prefers to *bet* on an event when the information is given by a (precise) probability distribution rather than some imprecise information, b prefers the bet with the precise information as well. This notion captures in rather natural terms a preference for precise information, which does not require the two decision makers that are compared to have the same risk attitudes. Our result states that two decision makers can be compared according to that notion if and only if the selected probability-possibility set of one of them is included in the other's.

In our representation theorem, we use Gilboa and Schmeidler's axiom of uncertainty aversion which states that mixing two indifferent acts can be strictly preferred to any of these acts, for hedging reasons. We can however provide a more direct way of modeling the decision maker's attitude toward imprecision, which also provides an easy way of experimentally testing the axiom. We show in particular that uncertainty aversion is implied by an axiom of aversion toward imprecision which compares the same act under two different probability-possibility sets. Aversion toward imprecision states, loosely speaking, that the decision maker always prefers to act in a setting in which he possesses more information, i.e., the decision maker is averse toward a "garbling" of the available information. At this stage, we simply remark that the notion we adopt of what it means for a probability-possibility set to be more imprecise than another one is rather weak and partial in the sense that it does not enable one to compare many sets (this will be discussed in Section 3.) One of the advantages of our setting is that it allows a clean separation between imprecision neutrality and the absence of imprecision. The latter is a feature of the information the decision maker possesses, while imprecision neutrality is characterized by the fact that the decision maker's selected probability-possibility set is reduced to a singleton.

The next step in the paper is to provide a more specific functional form. This is done under extra axioms that capture some invariance properties, which will be discussed in Section 4. The selected probability-possibility set is obtained by (i) solving for the "mean value" of the probability-possibility set, and (ii) shrinking the set toward its mean value according to a degree given by preference. The mean value is the *Steiner point* of the set (see Schneider

(1993)). For cores of belief functions for instance, it coincides with the Shapley value of the belief function. Denoting $s(P)$ the Steiner point of P, we obtain that $\varphi(P) = (1-\varepsilon)\{s(P)\} + \varepsilon P$ and hence (P, f) is preferred to (Q, g) if and only if

$$\varepsilon \min_{p \in P} \int u \circ f dp + (1-\varepsilon) \int u \circ f ds(P) \geq \varepsilon \min_{p \in Q} \int u \circ g dp + (1-\varepsilon) \int u \circ g ds(Q).$$

This functional form, already suggested in Ellsberg (1961), consists of taking the convex combination of two quantities: first, the minimum expected utility with respect to all the probability-possibility set and, second, the expected utility with respect to a particular probability distribution in this set. The parameter ε is obtained as part of the representation result and can be interpreted as a degree of imprecision aversion. When $\varepsilon = 0$, we obtain expected utility. When $\varepsilon = 1$ the functional form expresses the extreme case where the decision maker takes the worst case scenario in the entire probability-possibility set.

A portfolio choice example illustrates the mechanics of the model. In particular, it shows how the distinction between imprecision and attitude toward imprecision can give rise to different comparative statics exercises.

Relationship with the literature

Our model incorporates information as an object on which the decision maker has well defined preferences. To the best of our knowledge, Jaffray (1989) is the first to axiomatize a decision criterion that takes into account "objective information" in a setting that is more general than risk. In his model, preferences are defined over belief functions. The criterion he axiomatizes is a weighted sum of the minimum and of the maximum expected utility. This criterion prevents a decision maker from behaving as an expected utility maximizer, contrary to ours, which obtains as a limit case the expected utility criterion. Interest in this approach has been renewed recently, in a series of papers that have in common that objects of choice are sets of lotteries. Olszewski (2007) also characterizes a version of the α-maxmin expected utility in which the decision maker puts weights both on the best-case and the worst-case scenarios. Stinchcombe (2003) characterizes a general class of expected utility for sets of lotteries. Ahn (2005) also studies preferences over sets of lotteries, and characterizes a conditional subjective expected utility in which the decision maker has a prior probability over lotteries and updates it according to each objective set. A limitation of his analysis is that the sets of probabilities considered are "regular", i.e., have the same dimension as the simplex on the space of outcome.

Notice that our model *does not* reduce to one of choice over sets of lotteries (for more on this, see the end of Section 3). Consider for instance a decision maker who is asked whether he prefers betting on drawing a black ball in an urn filled with 100 balls that could be black or white in unknown proportion on the one hand or betting on drawing a red ball in an urn filled with 100 balls that could be yellow, blue, green, or red in unknown proportion. Nothing in our framework imposes that the decision maker is indifferent between these two bets. Such an indifference would follow if we were to impose that the decision makers only takes into account the induced distributions on outcomes when evaluating a pair (information,act). In this case, our framework would be comparable with, say, Olszewski (2007). Our criterion would then reduce to maxmin expected utility over the set of lotteries as we retain an axiom of uncertainty aversion. [3]

Viero (2007) derives a representation of preferences for a choice theory where the agent does not know the precise probability distributions over outcomes conditional on states. Instead, he knows only a possible set of these distributions for each state, thus generalizing the usual Anscombe and Aumann (1963) setting . The criterion obtained, called Optimism-Weighted Subjective Expected Utility takes in each state a weighted average of expected utility of the best and worst lottery and aggregate these numbers through an expectation over the states. The beliefs over the states are subjective.

Wang (2003) is closely related to our analysis. In his approach the available information is also explicitly incorporated in the decision model. Information takes the form of a set of probability distributions together with an anchor, i.e., a probability distribution that has particular salience. As in our analysis, he assumes that decision makers have preferences over couples (information,act). However, his axiom of ambiguity aversion is much stronger than ours and forces the decision maker to be a maximizer of the minimum expected utility taken over the entire information set. There is no scope in his model for less extreme attitude toward ambiguity. Following Wang's approach, Gajdos et al. (2004) proposed a weaker version of aversion toward imprecision still assuming that information was coming as a set of distributions together with an anchor. The notion of aversion toward imprecision developed in Section 3 is based on the one analyzed in Gajdos et al. (2004).

The notion of comparative imprecision aversion could itself be compared to the one found in Epstein (1999) and Ghirardato and Marinacci (2002). The latter define comparative ambiguity aversion using constant acts. They therefore need to control for risk attitudes in a separate manner and in the end, can compare (with respect to their ambiguity attitudes) only decision makers

[3] Olszewski (2007) uses instead an axiom of Set Betweenness, which is in general not satisfied by our criterion.

that have the same utility functions. [4] Epstein (1999) uses in place of our bets in the definition of comparative uncertainty aversion, acts that are measurable with respect to an exogenously defined set of unambiguous events. As a consequence, in order to be compared, preferences of two decision makers have to coincide on the set of unambiguous events. If the latter is rich enough, utility functions then coincide. Our notion of comparative imprecision aversion, based on the comparison of bets under precise and imprecise information does not require utility functions to be the same when comparing two decision makers. Said differently, risk attitudes are simply irrelevant to the imprecision aversion comparison.

The functional form axiomatized in Section 4 appears in some previous work (Gajdos et al. (2004) and Tapking (2004)), based on a rather different set of axioms and in a more limited setting. Kopylov (2006) also axiomatizes this functional form, for a fixed information-possibility set. In a setting similar to ours, Giraud (2006) axiomatizes a model in which the decision maker has non additive second order beliefs.

Klibanoff et al. (2005) provide a fully subjective model of ambiguity aversion, in which attitude toward ambiguity is captured by a smooth function over the expected utilities associated with a set of priors. The latter, as in Gilboa and Schmeidler (1989) is subjective. Hence, although their model allows for a flexible and explicit modeling of ambiguity attitudes, there is no link between the set of priors and the available information. Interestingly, part of Klibanoff et al. (2005)'s motivation is similar to ours, that is disentangling ambiguity attitude from the information the decision maker has. Formally, however, this separation holds in their model only if one makes the extra assumption that revealed beliefs coincide with the objective information available. In particular, comparative statics are more transparent in our model, as information can be exogenously changed.

Lehrer (2007) also provides an "information-based" model. He axiomatizes particular cases of the Choquet expected utility model and the multiple prior model in which revealed beliefs have the form of partially-specified probabilities (i.e., the decision maker "knows" the probability of some, though not all, events or the expectation of some, though not all, of the random variables.) His construction is entirely subjective: it does give some structure to the set of revealed beliefs but the link with prior information is not made explicit.

[4] They actually mention that if one wants to compare two decision makers with different utility functions, one has first to completely elicit them.

2 Extended multiple-priors model

We start with a benchmark model that extends the multiple-priors model by Gilboa and Schmeidler (1989) into the variable information setting.

2.1 Representation theorem

2.1.1 Set up

Let $\Omega = \mathbb{N}$ be the countable set of all the *potential* states of the world.[5] Let $\mathcal{S}$ be the family of nonempty and finite subsets of Ω. For each $S \in \mathcal{S}$, denote the set of probability measures over S by $\Delta(S)$. Let $\mathcal{P}(S)$ be the family of compact and convex subsets of $\Delta(S)$, where compactness is defined with regard to the Euclidian space $\mathbb{R}^S$. Let $\mathcal{P}$ be the family of *probability-possibility sets*, that is defined by $\mathcal{P} = \bigcup_{S \in \mathcal{S}} \mathcal{P}(S)$. For each $P \in \mathcal{P}$, its support is denoted by $supp(P)$ and is equal to the union over $p \in P$ of the support of p.[6]

The space of probability-possibility sets $\mathcal{P}$ is a mixture space under the operation defined by

$$\lambda P + (1-\lambda)Q = \{\lambda p + (1-\lambda)q : p \in P, q \in Q\}.$$

The set of pure outcomes is denoted by X. Let $\Delta^*(X)$ be the set of simple lotteries (probability measures with finite supports) over X. Let $\mathcal{F} = \{f : \Omega \to \Delta^*(X)\}$ be the set of *lottery acts*. Without loss, any lottery is viewed as a constant act which delivers that lottery regardless of the states.

The domain of objects of choice is $\mathcal{P} \times \mathcal{F}$. The decision maker has a preference relation over $\mathcal{P} \times \mathcal{F}$, which is denoted by $\succsim$. The decision maker compares pairs of probability-possibility sets and acts. When

$$(P, f) \succsim (Q, g),$$

the decision maker prefers choosing f under P to choosing g under Q. When $Q = P$, the preference relation represents the ranking of acts given the information embodied by P. When $g = f$, the preference relation represents

[5] We assume that the state space is countably infinite. This assumption is not needed for Theorems 1, 2, 3, and 6 which remain true in a finite setting. We use it to prove Theorems 4, 5, 7, and 8. As we explain there, we could actually do the analysis in a finite setting, at the cost of cumbersome notation, and chose rather to use the modeling device of an infinite state space.

[6] For the sake of notational simplicity, we will consider that all probability distributions are defined over Ω and that $p(\omega) = 0 \ \forall \omega \in \Omega \setminus supp(p)$.

the ranking of probability-possibility sets given the action embodied by f. For $E \in \mathcal{S}$, denote $f_E g$ the act that yields $f(\omega)$ if $\omega \in E$ and $g(\omega)$ if not.

2.1.2 Discussion

The object of choice should be understood in the following manner. Consider Ellsberg's two urns example. In the first urn, there are 50 red and 50 blue balls. In the second urn, there are 100 yellow and green balls in unknown proportion. Two representations are possible. The first one is to construct the "grand state space" with four states according to which color is drawn and construct the probability-possibility set on these four states. The second, which we adopt in the paper, is to consider two different "situations" (known urn or unknown urn), each one with its probability-possibility set. In each situation, the decision maker can take a decision, which is represented by a mapping from some set of integers (that encode the situation) to a space of outcome. One can then compare betting in one situation versus betting in the other situation.

This example enlightens a few things. First, information is given at the outset. The decision maker can choose to act in a situation in which there is more (or less) information but cannot modify the information. Second, information is not contradictory among two situations. It simply bears on different objects. Since each set bears on a different part of the uncertainty, it does not make sense to take the intersection of these sets. Third, the labeling of "states" within each situation is arbitrary and is a mere encoding: hence one can name in the same way two acts that actually represent quite distinct choices. Betting on red or betting on yellow in the example above can be represented by a couple $(1, 0)$ in each situation and be given the label "f". The preference $(\{1/2, 1/2\}, f) \succ (\Delta(\{1, 2\}), f)$ represents the fact that the decision maker prefers to bet in a known urn rather than in an urn whose composition is unknown.

This introduces in the model some invariance properties, since "states" are not really meaningful and are simply represented by some integers: we assume that the decision maker is only sensitive to the information he has on the states and not to the manner states are encoded. Hence, in the Ellsberg urn, $(\{1/2, 1/2\}, f)$ has to be indifferent to $(\{1/2, 1/2\}, g)$ where $g = (0, 1)$. Similarly, $(\Delta(\{1, 2\}), f)$ has to be indifferent to $(\Delta(\{1, 2\}), g)$.

Consider finally a decision maker who chooses according to the induced set of distributions on outcome, as in Olszewski (2007). In this case, an act $f = (1, 0, 0, \dots)$ would be evaluated in the same way whether the information is $\Delta(\{1, 2\})$ or $\Delta(\{1, 2, 3\})$: $(\Delta(\{1, 2\}), f) \sim (\Delta(\{1, 2, 3\}), f)$. Our model does not force this equivalence.

2.1.3 Axioms

We now introduce the axioms. The first two axioms are quite standard.

Axiom 1 *(Order) The preference relation $\succsim$ is complete and transitive.*

Axiom 2 *(Act Continuity) For every $P \in \mathcal{P}$ and $f, g, h \in \mathcal{F}$, if $(P, f) \succ (P, g) \succ (P, h)$, then there exist α and β in $(0, 1)$ such that*

$$(P, \alpha f + (1-\alpha)h) \succ (P, g) \succ (P, \beta f + (1-\beta)h).$$

The third axiom states that the preference over lotteries is independent of information sets and is nondegenerate. When a lottery is given regardless of the states of the world, information about their likelihood is irrelevant. Also, we exclude the case in which the decision maker is indifferent between all lotteries.

Axiom 3 *(Outcome Preference) (i) For every $P, Q \in \mathcal{P}$ and $l \in \Delta^*(X)$, $(P, l) \sim (Q, l)$, (ii) there exist $P \in \mathcal{P}$ and $l, m \in \Delta^*(X)$ such that $(P, l) \succ (P, m)$.*

The following three axioms are parallel to those in Gilboa and Schmeidler (1989).

Axiom 4 *(c-Independence) For every $f, g \in \mathcal{F}, l \in \Delta^*(X)$, $\lambda \in (0, 1)$ and $P \in \mathcal{P}$,*

$$(P, f) \succsim (P, g) \implies (P, \lambda f + (1-\lambda)l) \succsim (P, \lambda g + (1-\lambda)l).$$

Axiom 5 *(Uncertainty aversion) For every $f, g \in \mathcal{F}$, $\lambda \in (0, 1)$ and $P \in \mathcal{P}$,*

$$(P, f) \sim (P, g) \implies (P, \lambda f + (1-\lambda)g) \succsim (P, f).$$

Axiom 6 *(Monotonicity) For every $f, g \in \mathcal{F}$ and $P \in \mathcal{P}$,*

$$(P, f(\omega)) \succsim (P, g(\omega)) \ \ \forall \omega \in supp(P) \implies (P, f) \succsim (P, g).$$

2.1.4 Gilboa-Schmeidler extended

Theorem 1 *The preference relation $\succsim$ satisfies Axioms 1 to 6 if and only if there exist a function $U : \mathcal{P} \times \mathcal{F} \to R$ which represents $\succsim$, a mixture-linear function $u : \Delta(X) \to R$ and a mapping $\varphi : \mathcal{P} \to \mathcal{P}$ such that*

$$U(P, f) = \min_{p \in \varphi(P)} \sum_{\omega \in \Omega} u(f(\omega))\, p(\omega).$$

Moreover, u is unique up to positive linear transformations and φ is unique and has the property $supp(\varphi(P)) \subset supp(P)$.

We purposely kept as close as possible to the original axioms of Gilboa and Schmeidler (1989). In particular, we kept their two key axioms (c-Independence) and (Uncertainty Aversion). We will argue later that the latter can be replaced by a more explicit representation of the agent's attitude toward imprecision of the available information.

At this stage, the only property of φ is that $supp(\varphi(P)) \subset supp(P)$, meaning that the decision maker considers the information represented by P to be credible in the sense that states that are not given any weight by any of the relevant probability distributions are simply irrelevant. Note that this implies that the support of $\varphi(P)$ is finite.

2.2 The link between P and $\varphi(P)$

In our setup $\varphi(P)$ embodies attitude toward uncertainty as well as the processing of some given information. Call it the selected probability-possibility set. Our approach is meant to shed some light as to what is behind this set. In the representation theorem above, P is objective information, while the function φ is subjective in the same way the utility function u is. We can however specify further the property of φ that will constrain the link between the selected probability-possibility set and the initial probability-possibility set. This link is left rather unclear in Theorem 1.

For probability $\{p\} \in \mathcal{P}$ and act $f \in \mathcal{F}$, define the induced distribution over outcomes by

$$l(p, f) = \sum_{\omega \in \Omega} p(\omega) f(\omega).$$

The next axiom states that the evaluation of an act under *precise* information –that is, when the probability-possibility set is given as a singleton– depends only on its induced distribution.

Axiom 7 *(Reduction under Precise Information) For every $\{p\} \in \mathcal{P}$ and $f \in \mathcal{F}$,*

$$(\{p\}, f) \sim (\{p\}, l(p, f)).$$

If one adds this axiom to the ones in Theorem 1, φ has the further property that $\varphi(\{p\}) = \{p\}$ for all $\{p\} \in \mathcal{P}$. Thus, when told a precise probabilistic information the decision maker is a vNM decision maker.

The next axiom states that if one act is preferable to another under every

element of the information set, the ranking is unchanged under the whole set.

Axiom 8 *(Dominance) For every $f, g \in \mathcal{F}$ and $P \in \mathcal{P}$,*

$$(\{p\}, f) \succsim (\{p\}, g) \text{ for every } p \in P \implies (P, f) \succsim (P, g).$$

Another way to interpret the axiom is by saying that if $(P, g) \succ (P, f)$, then it has to be the case that there is some (precise) information in P under which g is preferred to f. Thus, the decision maker does not contemplate information outside of the set P as relevant for comparing the two acts.

If one adds (Dominance) to Axioms 1 to 6, then Theorem 1 holds with the property that $\varphi(P) \subset co\,(\cup_{p \in P} \varphi(\{p\}))$.

The following theorem combines these axioms and provides a stronger condition on φ.[7]

Theorem 2 *The preference relation $\succsim$ satisfies Axioms 1 to 5, 7 and 8 if and only if there exist a function $U : \mathcal{P} \times \mathcal{F} \to R$ which represents $\succsim$, a mixture-linear function $u : \Delta(X) \to R$ and a mapping $\varphi : \mathcal{P} \to \mathcal{P}$ such that*

$$U(P, f) = \min_{p \in \varphi(P)} \sum_{\omega \in \Omega} u(f(\omega))\, p(\omega).$$

Moreover, u is unique up to positive linear transformations and φ is unique, has the property

(Selection): $\varphi(P) \subset P$ for every $P \in \mathcal{P}$.

Hence, when one combines (Reduction under Precise Information) and (Dominance), one gets that $\varphi(P)$ is a selection from P (thus explaining the terminology adopted of "selected probability-possibility set".) If P is a singleton, then the constraint imposed by the (Selection) property rules out any subjectivity in the selected probability-possibility set of the agent, which has to coincide with the given singleton.

Theorem 2 is not very specific as to which form φ could take. The first example that comes to mind is when φ is the identity mapping: $\varphi(P) = P$. A second example, that will be developed in Section 6 is to consider $\varphi(P) = (1-\varepsilon)e(P) + \varepsilon\ P$ where $e(.)$ is a mapping from $\mathcal{P}$ to $\mathcal{P}$ with the property that $e(P)$ is a singleton included in P, which gives the "reference distribution" for each set P. It corresponds to what is known in the literature as the ε-contamination case. Another class of examples is to take $\varphi(P)$ to be the set of maximizers

[7] Introducing Axioms 7 and 8 entail some redundancies : under Axioms 1 to 4, (Reduction under Precise Information) and (Dominance) imply (Monotonicity).

(or minimizers) of some function: $\varphi(P) = argmax_{p \in P} F(p)$. One can consider for instance entropy, where $F(p) = -\sum_{\omega \in supp(p)} p(\omega) \log_2(p(\omega))$.

In some applications, it is not farfetched to assume that there is a salient probability distribution in the set P. In that case, $F(p)$ could be taken to be the relative entropy (see Maccheroni et al. (2006)) with respect to that distribution. Such types of selection might be useful in some contexts, and they are allowed by our general theorem.

2.3 Comparative imprecision aversion

Say that a decision maker b is more averse toward imprecision than a decision maker a if whenever he prefers an act under a singleton probability-possibility set over the same act under a general probability-possibility set, so does a. Furthermore, one would like to separate out this attitude toward imprecision from the traditional attitude toward risk. In order to do that, one has to define carefully the set of acts for which the definition applies. For two prizes $x, y \in X$ and an event E denote $x_E y$ the act giving the degenerate lottery yielding x for sure when event E realizes and the degenerate lottery yielding y for sure if E does not realize. Such an act will be called a bet in the following.

Definition 1 *Let $\succsim_a$ and $\succsim_b$ be two preference relations defined on $\mathcal{P} \times \mathcal{F}$. Suppose there exist two prizes, $\bar{x}$ and $\underline{x}$ in X such that both a and b strictly prefer $\bar{x}$ to $\underline{x}$. We say that $\succsim_b$ is* more averse to bet imprecision than *$\succsim_a$ if for all $E \in \mathcal{S}$, $P \in \mathcal{P}$, and $\{p\} \in \mathcal{P}$,*

$$(\{p\}, \bar{x}_E \underline{x}) \succsim_a [\succ_a](P, \bar{x}_E \underline{x}) \Rightarrow (\{p\}, \bar{x}_E \underline{x}) \succsim_b [\succ_b](P, \bar{x}_E \underline{x})$$

Theorem 3 *Let $\succsim_a$ and $\succsim_b$ be two preference relations defined on $\mathcal{P} \times \mathcal{F}$ satisfying Axioms 1 to 7. Then, the following assertions are equivalent:*

(i) $\succsim_b$ is more averse to bet imprecision than $\succsim_a$,

(ii) for all $P \in \mathcal{P}$,

$$\left\{ q \in P | \forall E \subset supp(P), q(E) \geq \min_{p \in \varphi_a(P)} p(E) \right\} \subset \left\{ q \in P | \forall E \subset supp(P), q(E) \geq \min_{p \in \varphi_b(P)} p(E) \right\}.$$

Thus, the core of the "smallest" (convex) capacity that contains $\varphi_a(P)$ is included in the core of the "smallest" (convex) capacity that contains $\varphi_b(P)$. [8]

[8] $\overline{\nu(E)} = \min_{p \in \varphi_a(P)} p(E)$ defines a capacity, that is, a set function from the set of all subsets of $supp(P)$ to $[0, 1]$, which is monotone with respect to set inclusion. This capacity has the property of being a belief function (see Chateauneuf and Jaffray

Theorem 4 below provides a stronger characterization of comparative imprecision aversion (in terms of set inclusion of the selected probability possibility sets) that comes at the cost of imposing a stronger invariance property.

This notion of comparative aversion to imprecision ranks preferences that do not necessarily have the same attitudes toward risk. This is of particular interest in applications if one wants to study the effects of risk aversion and imprecision aversion separately. For instance, one might want to compare portfolio choices of two agents, one being less risk averse but more imprecision averse than the other. This type of comparison cannot be done if imprecision attitudes can be compared only among preferences that have the same risk attitude, represented by the utility function. To the best of our knowledge, there is no available result in the literature that achieves this separation.

In our model, decision makers are von Neumann-Morgenstern under precise information, so that aversion toward risk is captured by the concavity of the utility index. Now, one could imagine that decision makers could distort probabilities as in rank dependent expected utility. We conjecture that Theorem 3 would remain valid in this more general setting, as the definition of comparative imprecision aversion amounts only to compare the probability of the good event, a comparison that should not be affected by increasing distortion function.

3 Imprecision aversion

In this Section, we propose a definition of an (incomplete) order on sets of distributions that could arguably be appropriate to model imprecision of a set of distributions and show that this notion is behind Gilboa and Schmeidler's definition of uncertainty aversion. But before dealing with this issue, we need to introduce an axiom that links the usual notion of mixture with information possibility sets that have a particular structure.

3.1 Preliminaries: mixing

Take the usual notion of mixing, that is at the heart of Gilboa and Schmeidler's approach to uncertainty aversion: $\alpha f + (1-\alpha)g$ is the act that yields the lottery $\alpha f(\omega) + (1-\alpha)g(\omega)$ in state ω. To the extent that mixing acts is seen as being equivalent to tossing a coin, we can express an idea similar to mixing by playing on the information structure rather than on the acts.

(1989) for instance).

Take the following example: consider the set of distributions given by

$$Q = \{(\alpha p, (1-\alpha)p, \alpha(1-p), (1-\alpha)(1-p)) | p \in [0,1]\}$$

where α is some number in $(0,1)$. This set has the particular feature that for any $q \in Q$, $\frac{q(1)}{q(2)} = \frac{\alpha}{1-\alpha} = \frac{q(3)}{q(4)}$. To pursue, consider now an act h together with Q. Axiom 9 below assimilates (Q,h) with the following mixture: $(\Delta(\{1,2\}), \alpha f + (1-\alpha)g)$ where $f(1) = h(1)$, $g(1) = h(2)$, $f(2) = h(3)$, $g(2) = h(4)$. This new pair represents a situation in which, when state 1 occurs –which happens with any probability $p \in [0,1]$– the decision maker is faced with the lottery $\alpha h(1) + (1-\alpha)h(2)$. In state 2, he is faced with the lottery $\alpha h(3) + (1-\alpha)h(4)$. Note again that in this operation, the relative probability of being faced with $h(1)$ compared to $h(2)$ is given by $\frac{\alpha}{1-\alpha}$.

More generally, an information/act pair $(P, \alpha f + (1-\alpha)g)$ can be interpreted by saying that a state ω is determined, according to some unknown probability $p(\omega)$ that belongs to P. Then, once the state is realized, a roulette lottery or a coin flip takes place with odds $\alpha : (1-\alpha)$. As illustrated in Figure 1, (Q,h) where $h(\omega) = f(\frac{\omega+1}{2})$ if ω is odd, and $h(\omega) = g(\frac{\omega}{2})$ if ω is even while $Q = \{q | \exists p \in P \text{ s.th. } q(\omega) = \alpha p(\frac{\omega+1}{2})$ if ω is odd and $q(\omega) = (1-\alpha)p(\frac{\omega}{2})$ if ω is even$\}$ can then be seen as the result of "collapsing" the roulette lottery in the probability distribution that determines the state. Said differently, the state now incorporates whether the coin toss ended heads or tails. Each state is now split in two: state ω is split into (state ω, heads) and (state ω, tails) and, conditionally on being in state ω, there is a given probability distribution that is fixed across states, according to which heads or tails is determined. The next axiom states that the decision maker sees the two objects as the same. Thus, the axiom says, the operation described above is neutral for the decision maker as it does not modify the timing of the process: uncertainty first and then risk. In spirit, this axiom is very similar to the usual reduction of compound lottery axiom. [9]

Axiom 9 *(Decomposition Indifference) Let $f, g, h \in \mathcal{F}$ and $P, Q \in \mathcal{P}$. If*

- *$h(\omega) = f(\frac{\omega+1}{2})$ if ω is odd, and $h(\omega) = g(\frac{\omega}{2})$ if ω is even and,*

[9] In this axiom we need the state space to be infinite, although we could develop a model with a finite state space, in which a randomizing device would be explicitly part of the probability-possibility set. In that case, we would assume a finite state space S and append to it $\{0,1\}^{\#S}$, representing independent objective randomizing devices. We would then consider only sets of distributions on this product space that have a specific structure, i.e., the product of a set of distributions on S and of independent distributions on $\{0,1\}$. Some invariance properties would have to be assumed and Axiom 9 rewritten. Hence, although not necessary, assuming that the state space is infinite is, in our view, more convenient.

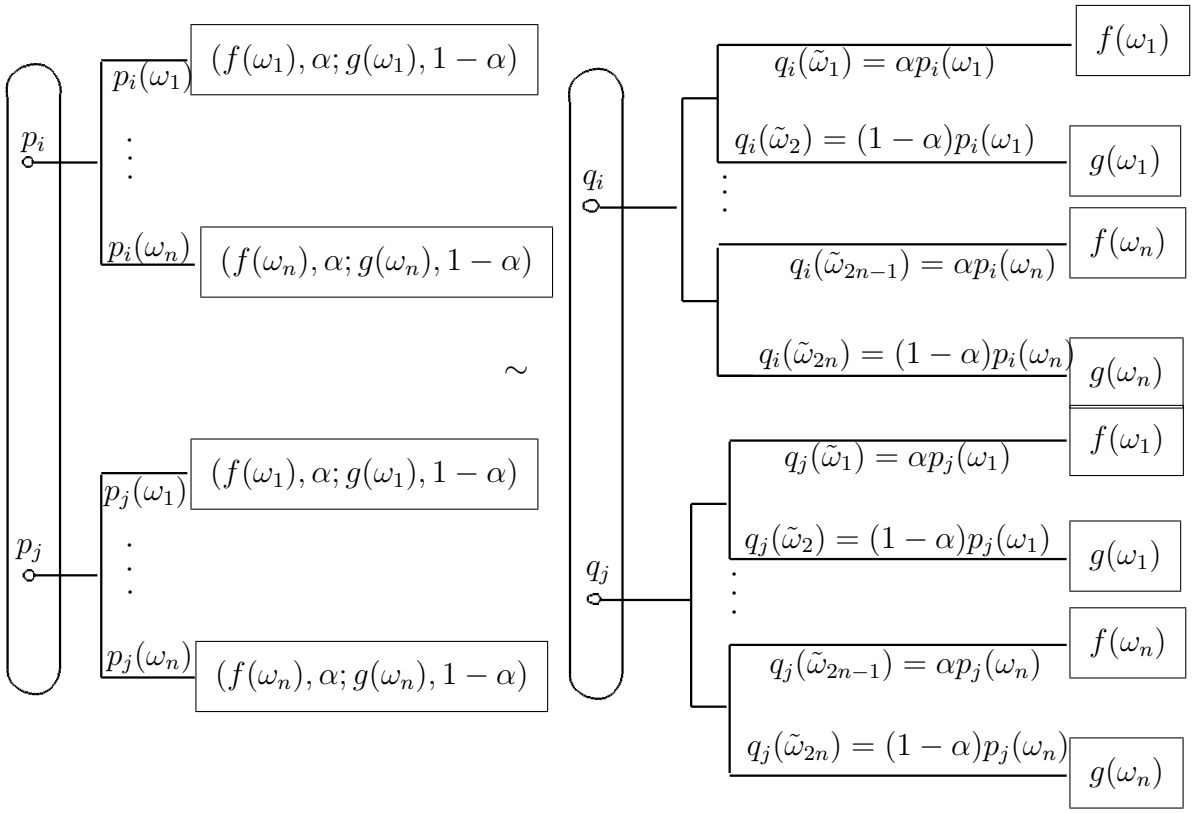

Fig. 1. Decomposition Indifference

- $Q = \{q | \exists p \in P \text{ s.th. } q(\omega) = \alpha p(\frac{\omega+1}{2}) \text{ if } \omega \text{ is odd and } q(\omega) = (1-\alpha)p(\frac{\omega}{2}) \text{ if } \omega \text{ is even}\}$ *for some* $\alpha \in [0,1]$

then $(P, \alpha f + (1-\alpha)g) \sim (Q, h)$.

The information embedded in the set Q is viewed as being equivalent to the information provided by the mixture operation. It expresses the fact that states should here be viewed as a mere encoding device. (Decomposition Indifference) implies some further properties of the selection function φ.

Proposition 1 *Let* $\succsim$ *satisfies Axioms 1 to 6. Then, the following two assertions are equivalent:*

(i) Axiom 9 holds,

(ii) the mapping $\varphi : \mathcal{P} \to \mathcal{P}$ *in Theorem 1 has the property that*

$$\varphi(Q) = \{q | \exists p \in \varphi(P) \text{ s.th. } q(\omega) = \alpha p\left(\frac{\omega+1}{2}\right) \ \omega \text{ odd }, q(\omega) = (1-\alpha)p\left(\frac{\omega}{2}\right) \ \omega \text{ even}\} \tag{1}$$

$\forall P, Q \in \mathcal{P}$, *such that* $Q = \{q | \exists p \in P \text{ s.th. } q(\omega) = \alpha p(\frac{\omega+1}{2}) \text{ if } \omega \text{ is odd and } q(\omega) = (1-\alpha)p(\frac{\omega}{2}) \text{ if } \omega \text{ is even}\}$ *for some* $\alpha \in (0,1)$.

Example 1 *Assume that $\varphi(P)$ is the maximum entropy of P. Given that*

$$\min_{p\in argmax\{-\sum_\omega p(\omega)\log_2 p(\omega)\}} \sum_\omega p(\omega)[\alpha f(\omega)+(1-\alpha)g(\omega)]$$

$$=$$

$$\min_{p\in R}\left[\sum_{\omega\ odd}\ p(\tfrac{\omega+1}{2})f(\tfrac{\omega+1}{2})+\sum_{\omega\ even}\ p(\tfrac{\omega}{2})g(\tfrac{\omega}{2})\right]$$

where $R = argmax\{-\sum_{\omega\ odd}\ \alpha p(\frac{\omega+1}{2})\log_2 \alpha p(\frac{\omega+1}{2})-\sum_{\omega\ even}(1-\alpha)p(\frac{\omega}{2})\log_2(1-\alpha)p(\frac{\omega}{2})\}$*, Axiom 9 is satisfied for this particular case.*$\diamondsuit$

Example 2 *Axiom 9 (and as a consequence Property (1) in Proposition (1)) will not be satisfied if, for instance* $\varphi(P) = argmin_{p\in P}\sum_\omega(p(\omega)-\bar{p}(\omega))^2$ *or* $\varphi(P)=\{p\in P|d(p)-d(\bar{p})\leq\zeta\}$ *where* $\bar{p}$ *is some given probability distribution,* d *some distance function and* ζ *a positive number. This is because keeping* $\bar{p}$ *fix while "spreading" the* p*'s affects the distance between these objects.* $\diamondsuit$

Property (1) of the selection function will be used in all the results of this section.

3.2 Imprecision aversion

We first provide a characterization of comparative bet imprecision aversion under Axiom 9, which gives a tighter result than Theorem 3.

Theorem 4 *Let* $\succsim_a$ *and* $\succsim_b$ *be two preference relations defined on* $\mathcal{P}\times\mathcal{F}$ *satisfying Axioms 1 to 7, and Axiom 9. Then, the following assertions are equivalent:*

(i) $\succsim_b$ *is more averse to bet imprecision than* $\succsim_a$*,*

(ii) for all $P\in\mathcal{P}$*,* $\varphi_a(P)\subset\varphi_b(P)$*.*

We now give a new foundation for the uncertainty aversion axiom, showing that it is implied by an axiom of aversion toward imprecision. The latter compares an act under two different probability-possibility settings and states that the decision maker always prefers the more precise information. We therefore have to define a notion of imprecision on sets of probability distributions. The most natural definition would be that P is more precise than Q whenever $P\subset Q$. This is actually the definition proposed by Wang (2003). However, this definition turns out to be much too strong. Indeed, the idea behind the notion of aversion toward imprecision is that an imprecision averse decision maker should *always* prefer a more precise information, whatever the act under consideration. Consider an act f for which the worst outcome is obtained, say, in state 1. Then, Wang's notion of precision would force the decision maker

to prefer $(\{(1,0)\}, f)$, that is, putting probability one on the worst outcome to $(\Delta(\{1,2\}), f)$, that is, being totally uncertain about the state; a feature of the axiom which is very unlikely and unappealing.

On the other hand, it is clear that a set being more precise than another has something to do with set inclusion. The following definition restricts the inclusion condition to some sets of probability distributions that are comparable in some sense, exactly as the comparison of two distributions in terms of risk focusses on distributions that have the same mean.

Definition 2 *Let $P, Q \in \mathcal{P}$. Say that P is* conditionally more precise *than Q if*

- *$P \subset Q$ and,*
- *there exists a partition $(E_1, \ldots, E_n)$ of Ω such that*

(i) $\forall p \in P$, $\forall q \in Q$, $p(E_i) = q(E_i)$ for all $i = 1, \ldots, n$,
(ii) $co\{p(.|E_i); p \in P\} = co\{q(.|E_i); q \in Q\}$ for all i such that $E_i \in supp(Q)$.

Note that this notion is very weak in the sense that it is very incomplete. For instance, an n-dimensional simplex cannot be compared through this definition with any of its subsets. Indeed, two sets P and Q, ordered by set inclusion, can be compared only if there exists a partition of the state space on which they agree *and* have precise probabilities (item (i) of the definition), and furthermore, conditionally on each cell of this partition, they give the same information (item (ii) of the definition). This means that the extra information contained in P is about some correlation between what happens in one cell E_i with what happens in another cell E_j. Said differently, the extra information is orthogonal to the "initial" probabilistic information, reflected in the fact that the cells of the partition have precise probabilities attached to them. The way this is expressed is via conditional probabilities: however, it should be underlined that these are simply a means to express properties of the probability-possibility sets that are compared. The use of conditional probabilities in this definition is not linked to any subjective assessment of the decision maker. Theorem 5 below provides the link between attitude toward this type of information and attitude toward hedging via mixing which is the basis for Gilboa and Schmeidler's Uncertainty Aversion axiom.

Example 3 *Consider the family*

$$P_\alpha = \left\{ \left(p, \frac{1}{2} - p, q, \frac{1}{2} - q \right) | p \in \left[0, \frac{1}{2} \right], q \in \left[0, \frac{1}{2} \right], |q - p| \leq \alpha \right\}$$

where $\alpha \in [0, \frac{1}{2}]$. One obviously has $P_\alpha \subset P_{\alpha'}$ for all $\alpha' \geq \alpha$. Now fix $\alpha < \frac{1}{2}$ and let $Q \equiv P_{1/2}$. $\{\{1,2\}, \{3,4\}, \{5, \ldots\}\}$ is a partition of the state space such that $\forall p \in P_\alpha$, $\forall q \in Q$, $p(E_i) = q(E_i)$. The set of probabilities conditional on $\{1,2\}$ is the same when computed starting from P_α and from Q. The same is true for

conditionals with respect to $\{3,4\}$. *Thus,* P_α *is conditionally more precise than* Q. *The nature of the extra information that is present in* P_α *is maybe clearest for* $\alpha = 0$. *In that case, one has* $q = p$ *and the extra information that is present in* P_0 *is a strong correlation between the different cells of the partition. More generally, we can look at upper and lower probabilities for events according to* P_α *and* Q. *We know they agree on the partition* $\{\{1,2\},\{3,4\},\{5,\dots\}\}$. *One can also check that the upper and lower probabilities on the events* $\{1,3\}$ *and* $\{2,4\}$ *are the same for the two sets (0 and 1 respectively). However, the lower and upper probability of events* $\{2,3\}$ *and* $\{1,4\}$ *do differ for the two sets. One has, with obvious notation,* $\underline{p}_\alpha(\{2,3\}) = 1/2-\alpha$ *and* $\bar{p}_\alpha(\{2,3\}) = 1/2+\alpha$ *while* $\underline{q}(\{2,3\}) = 0$ *and* $\bar{q}(\{2,3\}) = 1$, *and similarly for event* $\{1,4\}$. *The fact that* $\underline{p}_\alpha > \underline{q}$ *and* $\bar{p}_\alpha < \bar{q}$, *is another way to see that* P_α *is more precise than* Q.$\diamondsuit$

We can now state our axiom.

Axiom 10 *(Aversion toward Imprecision) Let* $P, Q \in \mathcal{P}$ *be such that* P *is conditionally more precise than* Q, *then for all* $f \in \mathcal{F}$, $(P,f) \succsim (Q,f)$.

Remark 1 *Assume Theorem 1 and (Aversion toward Imprecision) hold. Then, for any* $P, Q \in \mathcal{P}$ *such that* P *is conditionally more precise than* Q, $\varphi(P) \subseteq \varphi(Q)$.

For the following theorem, we need a weak form of information independence (a stronger form will be introduced and discussed in the next Section.)

Axiom 11 *(Weak Information Independence) For every* $P, Q \in \mathcal{P}$, $f \in \mathcal{F}$, *and* $\lambda \in (0,1)$,

$$(P,f) \sim (Q,f) \implies (\lambda P + (1-\lambda)Q, f) \sim (P,f).$$

Theorem 5 *Suppose the Axioms (1), (9) and (11) hold. Then Axiom (10) implies Axiom (5).*

Thus, through this theorem, we identify characteristics of the objective information that the decision maker dislikes when he prefers the mixture of two indifferent acts to either act. The following example shows that equivalence does not hold in the previous theorem.

Example 4 *Let* $\alpha \in (0,1)$ *and consider*

$$Q = \{(\alpha p, (1-\alpha)p, \alpha(1-p), (1-\alpha)(1-p)); p \in [0,1]\}$$

and

$$Q' = \{(\alpha p, (1-\alpha)q, \alpha(1-p), (1-\alpha)(1-q)); p,q \in [0,1]\} = \alpha\Delta(\{1,3\}) + (1-\alpha)\Delta(\{2,4\})$$

and let $U(P,f) = \min_{p\in\varphi(P)} \sum_{\omega\in\Omega} u(f(\omega))\, p(\omega)$. *By construction, this satisfies (Uncertainty Aversion). Furthermore, it can be checked that* Q *is conditionally more precise than* Q'. *Hence, if* $\varphi(Q)$ *is not included in* $\varphi(Q')$, *(Aversion toward Imprecision) is violated.*

By (Decomposition Indifference) and Proposition 1,

$$\varphi(Q) = \{(\alpha p, (1-\alpha)p, \alpha(1-p), (1-\alpha)(1-p)); (p, 1-p) \in \varphi(\Delta(\{1,2\}))\}$$

Assuming (with slight abuse of notation) $\varphi(\Delta(\{1,3\})) = (\frac{1}{2}, 0, \frac{1}{2}, 0)$ *and* $\varphi(\Delta(\{2,4\})) = (0, \frac{1}{2}, 0, \frac{1}{2})$, *one gets by (Weak Independence) that* $\varphi(Q') = (\alpha\frac{1}{2}, (1-\alpha)\frac{1}{2}, \alpha\frac{1}{2}, (1-\alpha)\frac{1}{2})$. *Assuming* $\varphi(\Delta(\{1,2\})) = \Delta(\{1,2\}$ *yields the violation of (Aversion toward Imprecision).* $\diamondsuit$

3.3 On sets of lotteries

Olszewski (2007) considers a setting in which the objects of choice are sets of lotteries and derives an α-maxmin representation: a set is evaluated by the weighted average of the expected utilities of the best and the worst lottery in the set, with the weights interpreted as a measure of (comparative) attitude to objective ambiguity.

A particular case of our model is when the decision maker evaluates a pair (P,f) by the set of its induced distributions on outcomes. Take outcomes to be real numbers and consider $f = (1,0,0,\dots)$. Then, since the distributions on outcomes induced by $(\Delta(\{1,2\}), f)$ and $(\Delta(\{1,2,3\}), f)$ are the same, such a decision maker must satisfy $(\Delta(\{1,2\}), f) \sim (\Delta(\{1,2,3\}), f)$.

This indifference, in and by itself, rules out, in our model, the possibility for the decision maker to be imprecision neutral.[10] Indeed, an imprecision neutral decision maker would satisfy $(\Delta(\{1,2\}), f) \sim (\{(\frac{1}{2}, \frac{1}{2})\}, f)$ and $(\Delta(\{1,2,3\}), f) \sim (\{(\frac{1}{3}, \frac{1}{3}, \frac{1}{3})\}, f)$. But since $(\{(\frac{1}{2}, \frac{1}{2})\}, f) \succ (\{(\frac{1}{3}, \frac{1}{3}, \frac{1}{3})\}, f)$, this yields a contradiction to $(\Delta(\{1,2,3\}), f) \sim (\{(\frac{1}{3}, \frac{1}{3}, \frac{1}{3})\}, f)$.

Actually, when evaluating a pair (P,f) through the induced set of distributions on outcomes, our model yields maxmin expected utility with respect to the set of lotteries itself. More precisely, $\varphi(P) = P$ in this case. This can be seen on the previous example. Say that $U(\Delta(\{1,2,3\}), (1,0,0,0,\dots)) = au(1)+(1-a)u(0)$ and normalize $u(1) = 1$ and $u(0) = 0$. This means that $\forall p \in \varphi(\Delta(\{1,2,3\}))$, $p(1) \geq a$ and there exists $\bar{p} \in \varphi(\Delta(\{1,2,3\}))$, such that $\bar{p}(1) = a$. Since

$$(\Delta(\{1,2,3\}), (1,0,0,0,\dots)) \sim (\Delta(\{1,2,3\}), (0,1,0,0,\dots)) \sim (\Delta(\{1,2,3\}), (1,1,0,0\dots)),$$

[10] Imprecision neutrality is naturally defined by replacing $\succsim$ by $\sim$ in Axiom 10.

we get first that $\forall p \in \varphi(\Delta(\{1,2,3\}))$, $p(2) \geq a$, and, second, that there exists $\bar{p}' \in \varphi(\Delta(\{1,2,3\}))$ such that $\bar{p}'(1)+\bar{p}'(2) = a$. But this yields a contradiction unless $a = 0$. Hence, the distribution $\bar{p}' = (0,0,1,0,\dots) \in \varphi(\Delta(\{1,2,3\}))$. A similar reasoning, shows that all the extreme points of the simplex are also in $\varphi(\Delta(\{1,2,3\}))$ and therefore we can conclude that $\varphi(\Delta(\{1,2,3\})) = \Delta(\{1,2,3\})$.

Hence, adopting a framework in which the decision maker reduces (P,f) to its set of induced distribution prevents modeling a decision maker that would behave as an expected utility maximizer. This is actually also true in Olszewski (2007): with the α-maxmin functional he axiomatizes, one has $U(P,f) = \alpha \max_{p\in P} E_p[u\circ f] + (1-\alpha)\min_{p\in P} E_p[u\circ f]$. Hence, it is not possible to have that $(\Delta(\{1,2\}),f) \sim (\{(\frac{1}{2},\frac{1}{2})\},f)$ and $(\Delta(\{1,2,3\}),f) \sim (\{(\frac{1}{3},\frac{1}{3},\frac{1}{3})\},f)$: the first indifference would require $\alpha = \frac{1}{2}$ while the second would require $\alpha = \frac{1}{3}$.

4 Contraction representation: Axiomatic foundation

In this Section we provide an axiomatic characterization of the contraction representation, i.e., $\varphi(P) = (1-\varepsilon)s(P) + \varepsilon P$, where $s(P)$ is the Steiner point of P. The contraction representation is characterized by three additional axioms: an independence axiom on mixtures of probability-possibility sets, an invariance axiom with regard to certain transformations of probability measures, and a stronger continuity axiom.

4.1 Information independence

We start by introducing the independence axiom. It states that the ranking of probability-possibility sets given an act is unchanged under taking mixtures with a common probability-possibility set. It is a natural extension of von-Neumann-Morgenstern's independence axiom to the setting of imprecise information.

Axiom 12 *(Information Independence) For every $P, P', Q \in \mathcal{P}$, $f \in \mathcal{F}$, and $\lambda \in (0,1)$,*

$$(P,f) \sim (P',f) \implies (\lambda P + (1-\lambda)Q, f) \sim (\lambda P' + (1-\lambda)Q, f).$$

This axiom seems related to the c-Independence axiom. However, it might be worth mentioning that (Information Independence) and (c-Independence) are

orthogonal to each other. [11]

Addition of (Information Independence) restricts our φ function to be linear in mixtures of probability-possibility sets.

Proposition 2 *The preference relation $\succsim$ satisfies Axioms 1 to 6, and 12 if and only if we have the representation as in Theorem 1 with the additional property,*

(Mixture-linearity): $\varphi(\lambda P + (1-\lambda)Q) = \lambda\varphi(P) + (1-\lambda)\varphi(Q)$ $\forall P, Q \in \mathcal{P}$, $\forall \lambda \in [0,1]$.

The same claim holds for Axioms 1 to 5, 7, 8, and 12, with the further property that $\varphi(P) \subset P$ for all $P \in \mathcal{P}$.

4.2 Invariance and continuity

Next we introduce the invariance axiom. It roughly says that the decision maker's attitude toward information should not change under transformations of the state space (and probability simplex) that do not change attitude toward any precise information. This is interpreted as a requirement for a sophisticated attitude toward imprecise information.

First we give an informal presentation of the invariance axiom. For each $S \in \mathcal{S}$, let $\psi : \Delta(S) \to \Delta(S)$ denote a transformation on the simplex, and let $\widetilde{\psi} : \mathcal{F} \to \mathcal{F}$ be the transformation of acts associated with ψ. We will consider a class of transformations that do not change the ranking of precise information. That is, the transformations to be considered ψ should satisfy

$$(\{p\}, f) \succsim (\{q\}, f) \Rightarrow (\{\psi(p)\}, \widetilde{\psi}(f)) \succsim (\{\psi(q)\}, \widetilde{\psi}(f))$$

for every $p, q \in \Delta(S)$ and $f \in \mathcal{F}$.

For the transformations that satisfy the above property, we will impose an axiom in the form that for every $S \in \mathcal{S}$, every $P, Q \in \mathcal{P}(S)$ and $f \in \mathcal{F}$,

$$(P, f) \succsim (Q, f) \Rightarrow (\psi(P), \widetilde{\psi}(f)) \succsim (\psi(Q), \widetilde{\psi}(f)).$$

[11] A version of α-maximin representation in the form

$$U(P, f) = \alpha(f) \max_{p \in P} E_p[u \circ f] + (1 - \alpha(f)) \min_{p \in P} E_p[u \circ f]$$

where $\alpha(\cdot)$ is not constant in c-mixtures satisfies (Information Independence) but violates (c-Independence). Representation as in Theorem 1 where φ is not mixture-linear satisfies (c-Independence) but violates (Information Independence).

Below we formally identify the class of transformations that is appropriate for our argument. Here we restrict attention to a class of bistochastic matrices, that are stochastic generalization of permutations. An $|S| \times |S|$-matrix Π is *S-bistochastic* if it is nonnegative and $\sum_{\omega \in S} \Pi_{\omega\omega'} = 1$ for each $\omega' \in S$, and $\sum_{\omega' \in S} \Pi_{\omega\omega'} = 1$ for each $\omega \in S$. For an S-bistochastic matrix Π and $f \in \mathcal{F}$, define the transformed act $\Pi f \in \mathcal{F}$ by $(\Pi f)(\omega) = \sum_{\omega' \in S} \Pi_{\omega\omega'} f(\omega')$ for each $\omega \in S$, and $(\Pi f)(\omega) = f(\omega)$ for each $\omega \notin S$.

Any bistochastic matrix may be expressed as a convex combination of permutation matrices (see Birkoff (1946)). In that sense, it is a stochastic generalization of permutation. We consider a subclass of bistochastic matrices that do not change attitude toward any precise information.

Definition 3 *A bistochastic transformation Π is S-unitary if for every $p, q \in \Delta(S)$ and $f \in \mathcal{F}$,*

$$(\{p\}, f) \succsim (\{q\}, f) \Longrightarrow (\{\Pi p\}, \Pi f) \succsim (\{\Pi q\}, \Pi f).$$

Denote the set of all S-unitary transformations by $\mathcal{T}(S)$.

We note that unitary transformations include permutations as a special case.

The following lemma shows that the class of unitary transformations is nonempty, and is a natural correspondence of the standard unitary transformation on the Euclidian space, in the probability simplex.

Lemma 1 *Assume Axioms 1 to 4 and 7.*[12] *Then, any bistochastic transformation Π is S-unitary if and only if there exists $\lambda \in [0, 1]$ such that*

$$\Pi^t \Pi = \lambda I + \frac{1 - \lambda}{|S|} E,$$

where I denotes the identity matrix and E denotes the matrix with all entries being 1.

We state the axiom.

Axiom 13 *(Invariance to Unitary Transformations) For every $S \in \mathcal{S}$, any $\Pi \in \mathcal{T}(S)$, $f \in \mathcal{F}$ and $P, Q \in \mathcal{P}$,*

$$(P, f) \succsim (Q, f) \quad \Longrightarrow \quad (\Pi P, \Pi f) \succsim (\Pi Q, \Pi f).$$

To interpret, suppose that the decision maker prefers P to Q given f. Then under the above-noted axioms, Πf induces the same ranking of probabilities

[12] Alternatively, one can assume Axioms 1 to 3, 7, and 12.

as f. That is, for any $p \in P$ and $q \in Q$,

$$(\{p\}, f) \succsim (\{q\}, f) \Longrightarrow (\{\Pi p\}, \Pi f) \succsim (\{\Pi q\}, \Pi f),$$

and

$$(\{q\}, f) \succsim (\{p\}, f) \Longrightarrow (\{\Pi q\}, \Pi f) \succsim (\{\Pi p\}, \Pi f).$$

Thus after the transformation, given Πf, ΠP and ΠQ play the same roles as P and Q do in the original situation given f. Therefore, it is intuitive that the ranking of information sets is unchanged, which leads to the ranking $(\Pi P, \Pi f) \succsim (\Pi Q, \Pi f)$.

The implication of the above invariance axiom is that the decision maker does not consider that a probability measure at a certain location or direction is more likely to be true. At the level of how the decision maker selects the probability distributions, it says that the selection depends only on the shape of the set, and is independent of its location and direction. We should note that the notion of unitary invariance presumes that preference follows the expected utility hypothesis under precise information, which is characterized either by the conjunction of (Reduction under Precise Information) and (c-Independence), or by the conjunction of (Reduction under Precise Information) and (Information Independence). Unitary transformation is a stochastic extension of permutation, and it is assumed that attitude toward precise information is neutral to such stochastic extensions. This is the ground assumption under which the unitary-invariance axiom makes sense.

Finally, we consider a continuity axiom with regard to probability-possibility sets. [13] Recall that for each $S \in \mathcal{S}$, $\Delta(S)$ is a compact subset of the Euclidian space $\mathbb{R}^{|S|}$, and $\mathcal{P}(S)$ is a compact metric space with regard to the Hausdorff metric.

Axiom 14 *(Information Continuity): For every $S \in \mathcal{S}$, $f \in \mathcal{F}$ and $P \in \mathcal{P}(S)$, the sets $\{Q \in \mathcal{P}(S) : (Q, f) \succsim (P, f)\}$ and $\{Q \in \mathcal{P}(S) : (P, f) \succsim (Q, f)\}$ are closed with regard to the Hausdorff metric.*

Examples below illustrates the roles of the above axioms.

Example 5 *Fix $S \in \mathcal{S}$, and let $\mathcal{P}^*(S)$ be the set consisting of full-dimensional compact convex subsets, and singleton points of $\Delta(S)$. Let α be the uniform distribution over $\Delta(S)$. For $P \in \mathcal{P}^*(S)$ which is non-singleton, its center of gravity is defined by*

$$c(P) = \frac{1}{\alpha(P)} \int_P p\alpha(dp).$$

[13] This axiom has been criticized by Olszewski (2007). However, his criticism does not apply here as we only consider convex sets.

When the set is a singleton, its center of gravity is itself. Consider a mapping $\varphi : \mathcal{P}^*(S) \to \mathcal{P}^*(S)$ *defined by*

$$\varphi(P) = (1-\varepsilon)\{c(P)\} + \varepsilon P,$$

where $\varepsilon \in [0,1]$ *is a fixed parameter.*
Representation as in Theorem 1 which is restricted to $\mathcal{P}^*(S) \times \mathcal{F}$ *and has* φ *in the above form satisfies (Invariance to Unitary Transformations), since the selection depends only on the shape of the set and is invariant to location and scale. It also satisfies (Information Continuity) on* $\mathcal{P}^*(S) \times \mathcal{F}$*. However, it fails to satisfy (Information Independence), and also it fails to extend to* $\mathcal{P}(S) \times \mathcal{F}$ *so as to maintain (Information Continuity).*$\diamondsuit$

Example 6 *Fix* $S \in \mathcal{S}$ *and let* $e = (\frac{1}{|S|}, \cdots, \frac{1}{|S|})$ *and* $V = \{v \in \mathbb{R}^S : \langle v, e\rangle = 0, \|v\| = 1\}$ *be the* $|S|-2$ *dimensional unit sphere orthogonal to* e*. Let* μ *be a non-atomic probability measure over* V*. For* $P \in \mathcal{P}(S)$*, its generalized Steiner point with respect to* μ *is defined by*

$$s_\mu(P) = \int_V \arg\max_{p \in P} \langle p, v\rangle \; \mu(dv).$$

Consider a mapping $\varphi : \mathcal{P}(S) \to \mathcal{P}(S)$ *defined by*

$$\varphi(P) = (1-\varepsilon)\{s_\mu(P)\} + \varepsilon P,$$

where $\varepsilon \in [0,1]$ *is a fixed parameter.*
Representation as in Theorem 1 with φ *taking the above form satisfies (Information Independence) and (Information Continuity), but in general fails to satisfy (Invariance to Unitary Transformations) unless* μ *is uniform. When a permutation invariance condition is imposed on* φ*, we obtain that* μ *is permutation-invariant, i.e.,* $\mu(\pi \circ E) = \mu(E)$ *for every Borel subset* E *of* V *and every permutation* π*. However, the class of permutation-invariant measures is still very large.*
Whether we obtain the above class of selection mappings if we drop or weaken (Invariance to Unitary Transformations) is an open problem.$\diamondsuit$

4.3 Contraction representation result

We now provide the contraction representation in which the selected probability-possibility set is obtained by (i) solving for the 'center' of the probability-possibility set, and (ii) shrinking the set toward the center to a degree given by preferences. The 'center' is the *Steiner point*.

Imagine that a vector v is drawn from the unit sphere around the origin according to the uniform distribution. Then the Steiner point of set P, denoted

by $s(P)$, is the expected maximizer of pv over $p \in P$. Here the uncertainty is restricted to be about 'which direction is the right one to go.' Given set P and direction v, one goes within the set along this direction until she reaches an extreme point of the set that maximizes the inner product. However, since direction v is uncertain, she takes the expectation of such extreme points, where the distribution is uniform. This contrasts to another notion of center, the center of gravity, where the decision maker has a prior over the probability simplex (that is, the uncertainty is about 'which probability measure is the right one'), and takes the conditional expectation of measures given a probability-possibility set.

More formally, fix $S \in \mathcal{S}$ and let $e = (\frac{1}{|S|}, \cdots, \frac{1}{|S|})$ and $V = \{v \in \mathbb{R}^{|S|} : \langle v, e\rangle = 0, \|v\| = 1\}$ be the $|S| - 2$ dimensional unit sphere orthogonal to e. For $P \in \mathcal{P}(S)$, its Steiner point is defined by

$$s(P) = \int_V \arg\max_{p \in P} \langle p, v\rangle \nu(dv),$$

where ν is the uniform distribution over V.[14]

A characterization of the Steiner point goes as follows (see Schneider (1993), Theorem 3.4.2). The only way to select a point from a set according to a selection process that is 1) continuous, 2) linear, and 3) satisfies an invariance property with respect to isometries,[15] is to take its Steiner point (also called Steiner curvature centroid).

Example 7 *Steiner point of a segment is its midpoint.*

Example 8 *Steiner point of a polytope is the weighted average of its vertices, in which the weight for each vertex is proportional to its outer angle. Take for instance the set* $\{p|p(1)+p(2)+p(3) = 1$ *and* $p(2) \geq p(3)\}$*. Then, the Steiner point is obtained by putting weight* $150/360$ *on* $(1, 0, 0)$*,* $120/360$ *on* $(0, 1, 0)$ *and* $90/360$ *on* $(0, 1/2, 1/2)$*.*

Example 9 *When a probability-possibility set is given as the core of a lower probability (convex capacity), its Steiner point coincides with the Shapley value of the lower probability. This is not surprising since in the domain of convex capacities the Shapley value is the unique single-valued selection of the core that satisfies mixture independence and permutation invariance.*$\diamondsuit$

[14] Multiplicity of maximizers inside the integral does not matter since uniform distribution is non-atomic.

[15] More specifically, if one transforms a set via an isometry –a combination of rotations and translations– then the point selected from the transformed set should be the same as the one obtained by applying the isometry to the point selected initially.

We now state the contraction representation result.

Theorem 6 *The preference relation $\succsim$ satisfies Axioms 1 to 5, 7, 8, and 12 to 14 if and only if we have the representation as in Theorem 1 with the additional property that for every $S \in \mathcal{S}$, and $P \in \mathcal{P}(S)$,*

$$\varphi(P) = (1-\varepsilon)\{s(P)\} + \varepsilon\, P$$

with $\varepsilon \in [0,1]$ that is unique.

Notice that the rate ε is constant for every probability-possibility set with finite support.

Remark 2 *Under the representation of Theorem 6, a decision maker b who is more averse to bet imprecision than a decision maker a will have $\varepsilon^b > \varepsilon^a$.*

4.4 Imprecision premium

We define here a notion of imprecision premium which captures how much an agent is "willing to lose" when betting on an event in order to act in a setting that has no imprecision. Consider a preference relation $\succsim$ and let $\bar{x}$ and $\underline{x}$ be two prizes in X such that $\bar{x} \succ \underline{x}$. For any event $E \in \mathcal{S}$, let q^E be a probability distribution such that $(P, \bar{x}_E\underline{x}) \sim (\left\{q^E\right\}, \bar{x}_E\underline{x})$. Under Axioms 1 to 6, such a probability distribution exists and is independent of $\bar{x}$ and $\underline{x}$ since $(P, \bar{x}_E\underline{x}) \sim (\left\{q^E\right\}, \bar{x}_E\underline{x})$ if and only if $q^E(E) = \min_{p\in\varphi(P)} p(E)$.

Definition 4 *For any $P \in \mathcal{P}(S)$, any event $E \in \mathcal{S}$, let*

- *the* absolute imprecision premium, *$\pi^A(E,P)$ be defined by $s(P)(E) - q^E(E)$,*
- *the* relative imprecision premium, *$\pi^R(E,P)$ be defined by $\frac{\pi^A(E,P)}{s(P)(E) - Min_{p\in P} p(E)}$ whenever $s(P)(E) \neq Min_{p\in P} p(E)$.*

The absolute premium is thus the mass of probability on the good event that the agent is willing to forego in order to act in a precise situation represented by $s(P)$ rather than with the imprecise probability-possibility set P.

Theorem 7 *Let $\succsim_a$ and $\succsim_b$ be two preference relations defined on $\mathcal{P} \times \mathcal{F}$, satisfying Axioms 1 to 7, and Axiom 9. Then, the following assertions are equivalent:*

(i) $\succsim_b$ is more averse to bet imprecision than $\succsim_a$,

(ii) for all $P \in \mathcal{P}(S)$ and all event $E \in \mathcal{S}$, $\pi_b^A(E,P) \geq \pi_a^A(E,P)$.

Below, we show that the constancy of the relative imprecision premium provides an alternative characterization (under Axiom 9) of the contraction representation.

Definition 5 *A decision maker is said to have constant relative imprecision premium θ if for any $P \in \mathcal{P}$ and any $E \in \mathcal{S}$ such that $s(P)(E) \neq Min_{p \in P} p(E)$, $\pi^R(E, P) = \theta$.*

Theorem 8 *Consider a decision maker satisfying Axioms 1 to 7, and Axiom 9. The following assertions are equivalent:*

(*i*) *the decision maker has constant relative imprecision premium, equal to ε,*
(*ii*) *for all $P \in \mathcal{P}$, $\varphi(P) = (1 - \varepsilon)\{s(P)\} + \varepsilon P$.*

5 Example

We develop in this section a simple application of our analysis to portfolio choice that is similar in spirit to Klibanoff et al. (2005)'s. There are three assets, a, b, and c. The following table gives the payoff matrix

s	1	2	3	4
a	k	k	k	k
b	$\bar{b}$	$\bar{b}$	1	1
c	$\bar{c}$	1	1	$\bar{c}$

We put the following restrictions on the parameters: $\bar{c} > \bar{b} > k > 1$. The information available is given by the set

$$P_\alpha = \left\{ \left(p, \frac{1}{2} - p, q, \frac{1}{2} - q \right) | p \in \left[0, \frac{1}{2} \right], q \in \left[0, \frac{1}{2} \right], |q - p| \leq \alpha \right\}$$

where $\alpha \in [0, \frac{1}{2}]$. Hence, the probability of $\{1, 2\}$ is precise, equal to $1/2$ and similarly for $\{3, 4\}$. α is a measure of how "imprecise the set is": a higher α corresponds to a higher degree of imprecision. Taken with this information, the assets have a natural interpretation: asset a is the safe asset, b is the "risky" asset as its payoffs are measurable with respect to the partition $\{\{1, 2\}, \{3, 4\}\}$, and asset c is the "imprecise" asset.

We consider a decision maker with CARA utility function $u(w) = -e^{-\gamma w}$, where γ is the coefficient of absolute risk aversion. The selected set is given

by:

$$\varphi\left(P_\alpha\right)=\left\{\left(p,\frac{1}{2}-p,q,\frac{1}{2}-q\right)|p\in\left[\frac{1}{4}-\theta,\frac{1}{4}+\theta\right],q\in\left[\frac{1}{4}-\theta,\frac{1}{4}+\theta\right],|q-p|\leq\alpha\right\}$$

θ is the parameter of imprecision aversion: note that $\varphi(P_\alpha)$ is of the contraction type with $\theta=1-\varepsilon$. To make things interesting, we assume that $\theta\geq\alpha/2$, so that the constraint $|q-p|\leq\alpha$ is effective in the computation of the optimal portfolio (although see Remark 3 below.)

The decision maker has one unit of wealth that he has to allocate among the three assets. We allow for short sales. We consider successively three cases depending on which assets are actually available, the first case being the benchmark situation of choice between the safe and the risky asset.

Case 1: choice between safe and risky asset.

This case is the usual one and one gets that $b^\star=\frac{1}{\gamma(1-\bar{b})}log\left(\frac{k-1}{\bar{b}-k}\right)$, which is naturally independent from the parameters θ and α. Under the parameter restrictions, it is easy to see that increasing risk aversion decreases holding of the risky asset.

Case 2: choice between safe and imprecise asset.

The problem to be solved here is to find the optimal amount of the imprecise asset, i.e., the solution to:

$$\max_c \min_{\pi\in\varphi(P_\alpha)}-\left[(\pi(1)+\pi(4))e^{-\gamma((1-c)k+c\bar{c})}+(\pi(2)+\pi(3))e^{-\gamma((1-c)k+c)}\right],$$

or rewritten in terms of p and q:

$$\max_c \min_{\varphi(P_\alpha)}-\left[(p+1/2-q))e^{-\gamma((1-c)k+c\bar{c})}+(1/2-p+q)e^{-\gamma((1-c)k+c)}\right]$$

As long as $c>0$, $-e^{-\gamma((1-c)k+c\bar{c})}>-e^{-\gamma((1-c)k+c)}$ and hence the decision maker will "use" the probability in $\varphi(P_\alpha)$ that put the highest weight on the event $\{2,3\}$ and lowest weight on $\{1,4\}$. Hence, one wants to minimize $p-q$. Let therefore $q=1/4+\theta$ and $p=1/4+\theta-\alpha$.[16] Solving for the optimal solution

[16] Actually, it is easy to see that this is not the only possible choice of a minimizing probability. $q=1/4-\theta+\alpha$ and $p=1/4-\theta$ would also minimize $p-q$. The optimal solution however does not depend on which one of these probability distributions is used, as the objective function depends only on $p-q$.

yields

$$c^\star = \frac{1}{\gamma(\bar{c}-1)} log\left(\frac{(\bar{c}-k)(1/2-\alpha)}{(k-1)(1/2+\alpha)}\right)$$

One can check that $c^\star$ is positive as conjectured if $(k-1)/(\bar{c}-k) < (1/2-\alpha)/(1/2+\alpha)$. Here, the comparative statics with respect to γ works as in the single risky asset case. What is more interesting, although intuitive, is that the imprecise asset holding is decreasing in α: an increase in imprecision of the information provided reduces the amount of asset the decision maker wants to hold. Note also that imprecise asset holding does not depend, in this example, on the imprecision aversion parameter θ (as long as $\theta \geq \alpha/2$).

Case 3: choice among all three assets.

This is the more general case and is a bit more tedious to study. Let's write u_s the utility of the portfolio in state s. As long as $b > 0$ and $c > 0$, one has that $u_1 > u_2$ and $u_4 > u_3$ and furthermore, $u_4 - u_3 > u_1 - u_2$. Hence, the minimizing probability that belongs to $\varphi(P_\alpha)$ is given by $p = 1/4 + \theta - \alpha$ and $q = 1/4 + \alpha$.

Let $K = \frac{(\bar{c}-k)(\bar{b}-1)}{(\bar{c}-\bar{b})(k-1)}$. Under our assumption, $K > 1$. Then, the optimal solution can be written:

$$b^\star = \frac{1}{\gamma(\bar{b}-1)} log\left[(K-1)\frac{1/4-\theta+\alpha}{1/4+\theta}\right]$$
$$c^\star = \frac{1}{\gamma(\bar{c}-1)} log\left[\frac{\bar{c}-\bar{b}}{\bar{b}-1}\left((K-1)\frac{1/4-\theta}{1/4+\theta} + \frac{1/4+\theta-\alpha}{1/4-\theta+\alpha}\right)\right]$$

Under some further (uninteresting) restrictions on the parameters, one can check that $b^\star > 0$ and $c^\star > 0$ as conjectured when picking the minimizing probability distribution.

One can thus perform comparative statics exercises. As α increases, that is as the information available is less precise, the decision maker will hold *more* of the risky asset and *less* of the imprecise asset. Thus, there is some form of substitution among assets as imprecision increases. This suggests that the observed under-diversification of decision makers' portfolio might be a consequence of how imprecision affects different assets. More specifically, consider parameter values such that $b^\star > c^\star$ (in our toy example this is the case for a large range of parameter values.) Note that if one were to ignore the effect of uncertainty on asset holding by wrongly setting $\alpha = 0$, the predicted holding of the risky asset would be lower than $b^\star$ while the predicted holding of the imprecise asset would be higher than $c^\star$, i.e., the predicted holdings would appear to be more diversified. Thus if one fails to identify which assets are

affected by imprecision, one could overestimate the predicted weight of these assets in the optimal portfolio.

Finally, it is also easy to show that the holdings of the risky as well as the imprecise assets are decreasing in the risk aversion parameter γ, as well as with the imprecision aversion parameter θ. This might help explaining phenomena like the equity premium puzzle, as imprecision aversion essentially reinforces the effect of risk aversion. Interestingly, these two very tentative hints as to how to account for the under-diversification puzzle *and* the equity premium puzzle in our model are linked to two different parameters (imprecision and imprecision aversion) and could therefore be incorporated in the *same* model.

Remark 3 The comparative static exercises performed were done under the assumption that $\theta \geq \alpha/2$. If this were not the case, then one can show that the minimizing probability used to evaluate the portfolio returns does not depend on α (when looking at the choice among all three assets.) Hence, over the full range of parameters there is a discontinuity in how imprecision affects holding of the risky and imprecise assets.

Remark 4 Note that all the action in this example does not take place because of the non-differentiability introduced by the min operator, as for instance in Epstein and Wang (1994) or Mukerji and Tallon (2001). Rather, the comparative statics were done at points where, locally, the decision maker behaves like an expected utility maximizer. More precisely, in usual maxmin expected utility models, decision makers look like expected utility maximizers away from the 45 degree line and there is no sense in which one can "change the set of priors" as there is no explicit link with the available information. In our model, there is some leverage in that respect even away from the kinks, as we have a way to link changes in the set of revealed probability distributions to changes in available information and to changes in imprecision attitudes. Thus, although non smooth, our model remains tractable in applications.

Appendix

Proof of Theorem 1

We only show sufficiency. Fix $P \in \mathcal{P}$. Then, Gilboa and Schmeidler (1989) result yields that Axioms 1 to 6 hold if and only if there exists a function $U_P : \mathcal{P} \times \mathcal{F} \to R$ such that $U_P(P, f) \geq U_P(P, g)$ if and only if $(P, f) \succsim (P, g)$ and a mixture-linear function $u_P : \Delta(X) \to \mathbb{R}$ and a unique set $\bar{P} \in \mathcal{P}$ such that

$$U_P(P, f) = \min_{p \in \bar{P}} \sum_{\omega \in \Omega} u_P(f(\omega))\, p(\omega).$$

Moreover, u_P is unique up to positive linear transformations. For each P, define $\varphi(P) = \bar{P}$ as obtained from Gilboa and Schmeidler's theorem. The latter also implies that the decision maker is an expected utility maximizer over constant acts. Axiom 3 implies that, for any $P, Q \in \mathcal{P}$, u_P and u_Q represent the same expected utility over constant acts. Hence, they can be taken so that $u_P = u_Q = u$.

To show that the representation can be extended to the entire domain $\mathcal{P} \times \mathcal{F}$, let $(P, f) \succsim (Q, g)$. Since $S(P)$ and $S(Q)$ are finite and $f(\omega)$ and $g(\omega)$ have finite support, using Axiom (3), there exist $\overline{x}$ and $\underline{x}$ in X such that for all $\omega \in S(P) \cup S(Q)$, for all $x \in Supp(f(\omega)) \cup Supp(g(\omega))$, $(P, k_{\overline{x}}) \succsim (P, k_x) \succsim \left(P, k_{\underline{x}}\right)$ where $k_{\overline{x}}$ (resp. k_x and $k_{\underline{x}}$) is the constant act giving the degenerate lottery $\delta_{\bar{x}}$ (resp. δ_x and $\delta_{\underline{x}}$) yielding $\overline{x}$ (resp. x and $\underline{x}$) for sure. Hence, by Axiom 6, we know that $(P, k_{\overline{x}}) \succsim (P, f) \succsim \left(P, k_{\underline{x}}\right)$ and $(P, k_{\overline{x}}) \succsim (P, g) \succsim \left(P, k_{\underline{x}}\right)$

By Axioms 1 and 2, there exists λ such that $(P, f) \sim (P, \lambda k_{\overline{x}} + (1 - \lambda) k_{\underline{x}})$. Similarly, there exists μ such that $(Q, g) \sim (Q, \mu k_{\overline{x}} + (1 - \mu) k_{\underline{x}})$. Thus,

$$(P, f) \succsim (Q, g) \Leftrightarrow (P, \lambda k_{\overline{x}} + (1 - \lambda) k_{\underline{x}}) \succsim (Q, \mu k_{\overline{x}} + (1 - \mu) k_{\underline{x}})$$

Now, $(P, f) \sim (P, \lambda k_{\overline{x}} + (1 - \lambda) k_{\underline{x}})$ implies that $\min_{p \in \varphi(P)} \int u \circ f dp = u(\lambda \delta_{\overline{x}} + (1 - \lambda)\delta_{\underline{x}})$. We also have that $\min_{p \in \varphi(Q)} \int u \circ g dp = u(\mu \delta_{\overline{x}} + (1 - \mu)\delta_{\underline{x}})$ and $u(\lambda \delta_{\overline{x}} + (1 - \lambda)\delta_{\underline{x}}) \geq u(\mu \delta_{\overline{x}} + (1 - \mu)\delta_{\underline{x}})$, which implies that

$$\min_{p \in \varphi(P)} \sum_{\omega \in \Omega} u_P(f(\omega))\, p(\omega) \geq \min_{p \in \varphi(Q)} \sum_{\omega \in \Omega} u_P(g(\omega))\, p(\omega).$$

We end by proving that $supp(\varphi(P)) \subset supp(P)$. Assume to the contrary that $supp(\varphi(P))$ is not included in $supp(P)$. Then there exists $p^* \in \varphi(P)$ and $\omega \in \Omega$ such that $\omega \in supp(p^*)$ and $\omega \notin supp(P)$. Consider $\overline{x}$ and $\underline{x}$ in X such that $u(\delta_{\overline{x}}) > u(\delta_{\underline{x}})$ and let f be defined by $f(\omega) = \delta_{\overline{x}}$ for all $\omega \in supp(P)$, $f(\omega) = \delta_{\underline{x}}$ otherwise. Thus, $\sum_{\omega \in \Omega} u(f(\omega))\, p^*(\omega) < u(\overline{x})$. Consider also the act

g defined by $g(\omega) = \delta_{\bar{x}}$ for all $\omega \in \Omega$. Then,

$$\min_{p\in\varphi(P)} \sum_{\omega\in\Omega} u(f(\omega))\, p(\omega) \leq \sum_{\omega\in\Omega} u(f(\omega))\, p^*(\omega) < u(\bar{x}) = \min_{p\in\varphi(P)} \sum_{\omega\in\Omega} u(g(\omega))\, p(\omega).$$

Hence, $(P, g) \succ (P, f)$, a violation of Axiom 6 since $g = f_{supp(P)}g$.

Proof of Theorem 2

We prove the selection property in two independent steps: the first one gives the consequences of Axiom 7, the second gives the consequences of Axiom 8.

Property 1 *Assume Theorem 1. Then, under Axiom 7,* $\varphi(\{p\}) = \{p\}$.

Proof: $(\{p\}, f) \sim (\{p\}, l(p, f))$ implies by Theorem 1 that $\min_{q\in\varphi\{p\}} \sum_{\omega\in\Omega} u(f(\omega))\, p(\omega) = \sum_{\omega\in\Omega} p(\omega)u(f(\omega))$. This is true for any f. By normalization, let u be such that $u(\delta_{\bar{x}}) = 1$ and $u(\delta_{\underline{x}}) = 0$ for some $\bar{x}, \underline{x} \in X$. Let f be an act that yields $(\delta_{\bar{x}}, \delta_{\underline{x}} \ldots, \delta_{\underline{x}})$, a vector of dimension $\#supp(p)$, for state in $supp(p)$ and consider permutations on the support. One then gets that $\min_{q\in\varphi\{p\}} q(\omega) = p(\omega)$ for all $\omega \in supp(\{p\})$. Hence, $\varphi(\{p\}) = \{p\}$.

Property 2 *Assume Theorem 1. Then, under Axiom 8,* $\varphi(P) \subset co\,(\cup_{p\in P}\varphi(\{p\}))$.

Proof: Let $P \in \mathcal{P}$ and assume that $\varphi(P) \not\subset co\,(\cup_{p\in P}\varphi(\{p\}))$, i.e., there exists $p^* \in \varphi(P)$ such that $p^* \notin co\,(\cup_{p\in P}\varphi(\{p\}))$. Using a separation argument, there exists a function $\phi : \Omega \to \mathbb{R}$ such that $\int \phi dp^* < \min_{p\in co\left(\cup_{p\in P}\varphi(\{p\})\right)} \int \phi dp$. Let $\bar{x}, \underline{x} \in X$ such that $u(\delta_{\bar{x}}) > u(\delta_{\underline{x}})$. Normalize u so that $u(\delta_{\bar{x}}) = 1$ and $u(\delta_{\underline{x}}) = 0$.

Since $supp(P) \cup supp\,(co\,(\cup_{p\in P}\varphi(\{p\})))$ is a finite set, there exist numbers $m > 0$ and ℓ, such that for all $\omega \in supp(P) \cup supp\,(co\,(\cup_{p\in P}\varphi(\{p\})))$, $m\phi(\omega) + \ell \in [0, 1]$.

Construct f as follows:

$$f(\omega) = \begin{cases} (m\phi(\omega) + \ell)\delta_{\bar{x}} + (1 - (m\phi(\omega) + \ell))\delta_{\underline{x}} & \forall\omega \in supp(P) \cup supp\,(co\,(\cup_{p\in P}\varphi(\{p\}))) \\ \delta_{\underline{x}} \text{ otherwise} \end{cases}$$

and let $\beta \equiv \sum_{\omega\in\Omega} u(f(\omega))p^*(\omega) \in [0, 1]$.

Let $g = \frac{1}{2}f + \frac{1}{2}k_{\beta\delta_{\bar{x}}+(1-\beta)\delta_{\underline{x}}}$. Since for all p,

$$\sum_{\omega\in\Omega} u(k_{\beta\delta_{\bar{x}}+(1-\beta)\delta_{\underline{x}}})p(\omega) = \beta$$

one gets that $\sum_{\omega\in\Omega} u(f(\omega))p^*(\omega) = \sum_{\omega\in\Omega} u(g(\omega))p^*(\omega)$.

Now, observe that for all p, $\sum_{\omega\in\Omega} u(g(\omega))p(\omega) = \frac{1}{2}\sum_{\omega\in\Omega} u(f(\omega))p(\omega) + \frac{1}{2}\sum_{\omega\in\Omega} u(f(\omega))p^*(\omega)$.

By construction,

$$\min_{q\in\varphi(\{p\})} \sum_{\omega\in\Omega} u(f(\omega))q(\omega) \geq \min_{q\in co\left(\cup_{p\in P}\varphi(\{p\})\right)} \sum_{\omega\in\Omega} u(f(\omega))q(\omega) > \sum_{\omega\in\Omega} u(f(\omega))p^*(\omega)$$

Thus, since $\min_{q\in\varphi(\{p\})} \sum_{\omega\in\Omega} u(g(\omega))q(\omega) = \frac{1}{2}\min_{q\in\varphi(\{p\})} \sum_{\omega\in\Omega} u(f(\omega))q(\omega) + \frac{1}{2}\sum_{\omega\in\Omega} u(f(\omega))p^*(\omega)$, we get

$$\min_{q\in\varphi(\{p\})} \sum_{\omega\in\Omega} u(f(\omega))q(\omega) > \min_{q\in\varphi(\{p\})} \sum_{\omega\in\Omega} u(g(\omega))q(\omega)$$

and therefore $(\{p\}, f) \succ (\{p\}, g)$ for all $p \in P$.

On the other hand, $\min_{q\in\varphi(P)} \sum_{\omega\in\Omega} u(f(\omega))q(\omega) \leq \sum_{\omega\in\Omega} u(f(\omega))p^*(\omega)$. Hence,

$$\min_{q\in\varphi(P)} \sum_{\omega\in\Omega} u(g(\omega))q(\omega) = \frac{1}{2}\min_{q\in\varphi(P)} \sum_{\omega\in\Omega} u(f(\omega))q(\omega) + \frac{1}{2}\sum_{\omega\in\Omega} u(f(\omega))p^*(\omega)$$
$$\geq$$
$$\min_{q\in\varphi(P)} \sum_{\omega\in\Omega} u(f(\omega))q(\omega).$$

Therefore $(P, g) \succsim (P, f)$, a contradiction.

Combining these two properties yields the (Selection) property. Finally, we prove that monotonicity is implied by the two extra axioms.

Property 3 *Assume Axioms 1 to 4 hold. Then, Axioms 7 and 8 imply Axiom 6.*

Proof: Take $P \in \mathcal{P}$ and $f, g \in \mathcal{F}$ such that $(P, f(\omega)) \succsim (P, g(\omega))$ for every $\omega \in supp(P)$. By Axiom 3, we can define, for fixed P, the preference over lottery outcomes $\succsim^\star$ by $l \succsim m$ if $(P, l) \succsim^\star (P, m)$. By Axioms 2 and 4, $\succsim^\star$ satisfies the vNM conditions for the existence of a mixture linear utility function u.

$$\begin{aligned}
(P, f(\omega)) \succsim (P, g(\omega)) \ \ \forall \omega \in supp(P) &\Leftrightarrow u(f(\omega) \geq u(g(\omega)) \ \ \forall \omega \in supp(P) \\
&\Leftrightarrow \sum_{\omega\in supp(P)} p(\omega)u(f(\omega) \geq \sum_{\omega\in supp(P)} p(\omega)u(g(\omega)) \ \ \forall p \in P \\
&\Leftrightarrow u(l(p, f)) \geq u(l(p, g)) \ \ \forall p \in P \\
&\Leftrightarrow (\{p\}, f) \succsim (\{p\}, g) \ \text{ by Axiom 7} \\
&\Leftrightarrow (P, f) \succsim (P, g) \ \text{ by Axiom 8.}
\end{aligned}$$

Proof of Theorem 3

$[(i) \Rightarrow (ii)]$ By (i), for any $q \in P$ and any $E \subset supp(P)$, $q(E) \geq \min_{p \in \varphi_a(P)} p(E)$ imply $q(E) \geq \min_{p \in \varphi_b(P)} p(E)$. Hence, $\min_{p \in \varphi_a(P)} p(E) \geq \min_{p \in \varphi_b(P)} p(E)$, proving (ii).

$[(ii) \Rightarrow (i)]$ Straightforward.

Proof of Proposition 1

$[(i) \Rightarrow (ii)]$ Let $P, Q \in \mathcal{P}$ and $\alpha \in [0,1]$ satisfy the assumptions of (ii) of the Proposition. We first prove that $\varphi(Q) \subseteq Q^*$ where

$$Q^* = \{q | \exists p \in \varphi(P) \text{ s.th. } q(\omega) = \alpha p(\frac{\omega+1}{2}) \text{ if } \omega \text{ odd and } q(\omega) = (1-\alpha)p(\frac{\omega}{2}) \text{ if } \omega \text{ even}\}$$

Assume that there exists $p^* \in \varphi(Q)$ such that $p^* \notin Q^*$. Since Q^* is a convex set, using a separation argument, there exists a function $\phi : \Omega \to \mathbb{R}$ such that $\int \phi dp^* < \min_{p \in Q^*} \int \phi dp$. Since $supp(Q)$ is a finite set, there exist numbers a, b with $a > 0$, such that $\forall \omega \in supp(Q)$, $(a\phi(\omega) + b) \in u(\Delta(X))$. Then, for all $\omega \in supp(Q)$ there exists $y(\omega) \in \Delta(X)$ such that $u(y(\omega)) = a\phi(\omega) + b$. Define $h \in \mathcal{F}$ by $h(\omega) = y(\omega)$ for all $\omega \in supp(Q)$, $h(\omega) = \delta_x$ for all $\omega \in \Omega \backslash supp(Q)$, where $x \in X$.

Then define $f, g \in \mathcal{F}$ by $f(\omega) = h(2\omega - 1)$ and $g(\omega) = h(2\omega)$. We have that $supp(Q^*) \subseteq supp(Q)$ and therefore

$$\min_{p \in Q^*} \sum_{\omega \in \Omega} u(h(\omega))\, p(\omega) = \min_{p \in \varphi(P)} \sum_{\omega \in \Omega} u(\alpha f(\omega) + (1-\alpha) g(w))\, p(\omega)$$

while

$$\min_{p \in Q^*} \sum_{\omega \in \Omega} u(h(\omega))\, p(\omega) > \sum_{\omega \in \Omega} u(h(\omega))\, p^*(\omega) \geq \min_{p \in \varphi(Q)} \sum_{\omega \in \Omega} u(h(\omega))\, p(\omega)$$

and thus

$$(P, \alpha f + (1-\alpha) g) \succ (Q, h)$$

which contradicts Axiom 9.

We can show that $Q^* \subseteq \varphi(Q)$ with the same kind of proof.

$[(ii) \Rightarrow (i)]$ Let f, g, h and P, Q be as in Axiom 9. Given the representation theorem and (ii), we have that:

$U(P, \alpha f + (1-\alpha) g) = \min_{p \in \varphi(P)} \sum_{\omega \in \Omega} p(\omega)[\alpha u(f(\omega)) + (1-\alpha) u(g(\omega))]$ and

$U(Q, h) = \min_{q \in \varphi(Q)} \sum_{\omega \in \Omega} q(\omega) u(h(\omega))$, which, given Property (1), is equal to

$$\min_{p\in\varphi(P)} \sum_{\omega\in\Omega,\ \omega \text{ odd}} \alpha p(\frac{\omega+1}{2})u(f(\frac{\omega+1}{2})) + \sum_{\omega\in\Omega,\ \omega \text{ even}} (1-\alpha)p(\frac{\omega}{2})u(g(\frac{\omega}{2}))$$

Collecting terms yields immediately that this is equal to $\min_{p\in\varphi(P)} \sum_{\omega\in\Omega} p(\omega)[\alpha u(f(\omega))+ (1-\alpha)u(g(\omega))]$ and hence $(P, \alpha f + (1-\alpha)g) \sim (Q, h)$.

Proof of Theorem 4

$[(i) \Rightarrow (ii)]$ Let $P \in \mathcal{P}$. Assume that $\varphi_a(P) \not\subset \varphi_b(P)$, i.e., there exists $p^* \in \varphi_a(P)$ such that $p^* \notin \varphi_b(P)$. Using a separation argument, there exists a function $\phi : \Omega \to \mathbb{R}$ such that $\int \phi dp^* < \min_{p\in\varphi_b(P)} \int \phi dp$. Let $\bar{x}$ and $\underline{x}$ in X be such that both a and b strictly prefer $\bar{x}$ to $\underline{x}$. Note that we can choose by normalization u_a and u_b so that $u_a(\bar{x}) = u_b(\bar{x}) = 1 > u_a(\underline{x}) = u_b(\underline{x}) = 0$. Since $supp(P)$ is a finite set, there exist numbers $m > 0$ and ℓ, such that for all $\omega \in supp(P)$, $m\phi(\omega) + \ell \in [0, 1]$. Denote $supp(P) = \{1, .., n\}$.[17] Let $\alpha_\omega = m\phi(\omega) + \ell$, $\omega \in supp(P)$.

Let $f^0 \in \mathcal{F}$ such that $f^0(\omega) = \alpha_\omega \delta_{\bar{x}} + (1-\alpha_\omega)\delta_{\underline{x}}$ for all $\omega = 1, ..., n$ and $f^0(\omega) = \delta_{\underline{x}}$ for all $\omega \in \Omega \backslash supp(P)$.

The act f^0 can be used to "separate" $p^\star$ from $\varphi_b(P)$. It has the feature that, in each state, it yields a lottery on the same outcomes $\bar{x}$ and $\underline{x}$. The rest of the proof consists in using Axiom 9 recursively to build an equivalent pair act/probability-possibility set in which the act is now a bet of the form $\bar{x}_E\underline{x}$.

At each stage of the recursion, each state ω is split into two, say $\bar{\omega}$ and $\underline{\omega}$. At each stage, build a new act by taking the appropriate state (defined by the recursion) and decomposing the lottery $\alpha\bar{x} + (1-\alpha)\underline{x}$ in state ω into an act that yields $\bar{x}$ for sure in "sub-state" $\bar{\omega}$ and $\underline{x}$ in "sub-state" $\underline{\omega}$. For all the other states, simply replicate the act (i.e., the value in each sub-state is the same.)

Formally, define $f^i \in \mathcal{F}$ for $i = 1, ..., n$ as follows:

- For all $\omega \in \{1, \cdots, 2^{i-1}n\}$,
  - If $f^{i-1}(\omega) \neq f^0(i)$, then $f^i(2\omega - 1) = f^i(2\omega) = f^{i-1}(\omega)$
  - If $f^{i-1}(\omega) = f^0(i)$, then $f^i(2\omega - 1) = \delta_{\bar{x}}$ and $f^i(2\omega) = \delta_{\underline{x}}$
- For all $\omega > 2^i n$, $f^i(\omega) = \delta_{\underline{x}}$.

[17] This is not without loss of generality. However, it is straightforward conceptually although notationally involved to generalize the proof to the case where $supp(P)$ is any finite set $\{\omega_1, \ldots, \omega_n\}$.

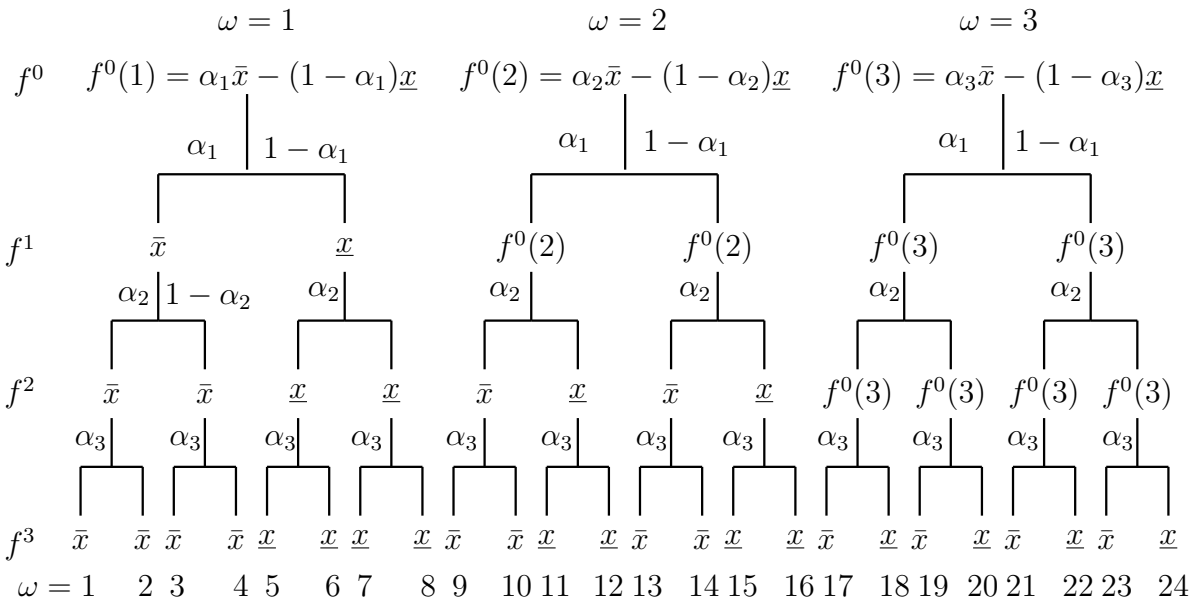

Fig. 2. An example of construction of a bet

Figure 2 illustrates the recursion for an act f^0 whose support is $\{1, 2, 3\}$, which is ultimately spread, via duplication of the states, on a bet that involves 24 states.

A similar operation has to be done for the probability-possibility sets, so as to maintain indifference throughout.

- $Q^1 = \{q | \exists p \in P \text{ s.th. } q(2\omega - 1) = \alpha_1 p(\omega) \text{ and } q(2\omega) = (1 - \alpha_1)p(\omega) \text{ for all } \omega \in \{1, \cdots, n\}\}$
- for $i = 2, ..., n$, $Q^i = \{q | \exists p \in Q^{i-1} \text{ s.th. } q(2\omega - 1) = \alpha_i p(\omega) \text{ and } q(2\omega) = (1 - \alpha_i)p(\omega) \text{ for all } \omega \in \{1, \cdots, 2^{i-1}n\}\}$

Finally, expand p^* in a similar way:

- p^{*1} be such that $p^{*1}(2\omega - 1) = \alpha_1 p^*(\omega)$ and $p^{*1}(2\omega) = (1 - \alpha_1)p^*(\omega)$ for all $\omega \in \{1, \cdots, n\}$
- for $i = 2, ..., n$, let p^{*i} be such that $p^{*i}(2\omega - 1) = \alpha_i p^{*i-1}(\omega)$ and $p^{*i}(2\omega) = (1 - \alpha_i)p^{*i-1}(2\omega)$ for all $\omega \in \{1, \cdots, 2^{i-1}n\}$.

By Proposition 1, one can check that:

- for $h = a, b$

$$\min_{p \in \varphi_h(P)} \sum_{\omega \in \Omega} u_h(f^0(\omega))\, p(\omega) = \min_{p \in \varphi_h(Q^1)} \sum_{\omega \in \Omega} u_h(f^1(\omega))\, p(\omega)$$
$$\text{and } \sum_{\omega \in \Omega} u_h(f^0(\omega))\, p^*(\omega) = \sum_{\omega \in \Omega} u_h(f^1(\omega))\, p^{*1}(\omega)$$

- for $h = a, b$, for $i = 2, ..., n$

$$\min_{p\in\varphi_h(Q^{i-1})}\sum_{\omega\in\Omega}u_h(f^{i-1}(\omega))\,p(\omega)=\min_{p\in\varphi_h(Q^{i})}\sum_{\omega\in\Omega}u_h(f^{i}(\omega))\,p(\omega)$$
$$\text{and }\sum_{\omega\in\Omega}u_h(f^{i-1}(\omega))\,p^{*i-1}(\omega)=\sum_{\omega\in\Omega}u_h(f^{i}(\omega))\,p^{*i}(\omega).$$

- for $i=1,...,n$, observe that $p^{*i}\in\varphi_a(Q^i)$. Hence,

$$\min_{p\in\varphi_b(P)}\sum_{\omega\in\Omega}u_b(f^{0}(\omega))\,p(\omega)=\min_{p\in\varphi_b(Q^n)}\sum_{\omega\in\Omega}u_b(f^{n}(\omega))\,p(\omega)>$$
$$\sum_{\omega\in\Omega}u_b(f^{0}(\omega))\,p^{*}(\omega)=\sum_{\omega\in\Omega}u_b(f^{n}(\omega))\,p^{*n}(\omega),$$

  while

$$\sum_{\omega\in\Omega}u_a(f^{0}(\omega))\,p^{*}(\omega)=\sum_{\omega\in\Omega}u_a(f^{n}(\omega))\,p^{*n}(\omega)\geq$$
$$\min_{p\in\varphi_a(P)}\sum_{\omega\in\Omega}u_a(f^{0}(\omega))\,p(\omega)=\min_{p\in\varphi_a(Q^n)}\sum_{\omega\in\Omega}u_a(f^{n}(\omega))\,p(\omega).$$

Observe finally that, as a result of the way the recursion works, f^n is of the form $\bar{x}_E\underline{x}$. Now, since by (Reduction under Precise Information) $\varphi(\{p^{*n}\})=\{p^{*n}\}$, we have

$$(\{p^{*n}\},\bar{x}_E\underline{x})\succeq_a(Q^n,\bar{x}_E\underline{x})$$

while

$$(Q^n,\bar{x}_E\underline{x})\succ_b(\{p^{*n}\},\bar{x}_E\underline{x}),$$

which contradicts the fact that $\succeq_b$ is more averse to bet imprecision than $\succeq_a$.

$[(ii)\Rightarrow(i)]$ Straightforward.

Proof of Theorem 5

Let $f,g\in\mathcal{F}$ and $P\in\mathcal{P}$ and assume $(P,f)\sim(P,g)$. Define h by $h(\omega)=f(\frac{\omega+1}{2})$ if ω is odd, and $h(\omega)=g(\frac{\omega}{2})$ if ω is even. For any $\alpha\in[0,1]$, define $Q(\alpha)=\{q|\exists p\in P$ s. th. $q(\omega)=\alpha p(\frac{\omega+1}{2})$ if ω is odd and $q(\omega)=(1-\alpha)p(\frac{\omega}{2})$ if ω is even$\}$.

By (Decomposition Indifference) $(Q(1),h)\sim(P,f)$ and $(Q(0),h)\sim(P,g)$ and hence, since by assumption $(P,g)\sim(P,f)$ and $\succsim$ is transitive, $(Q(0),h)\sim(Q(1),h)$. By (Weak Information Independence) for any $\lambda\in[0,1]$, $(\lambda Q(1)+(1-\lambda)Q(0),h)\sim(Q(0),h)$. Hence, $(\lambda Q(1)+(1-\lambda)Q(0),h)\sim(P,f)$.

Now, $Q(\lambda)$ is conditionally more precise than $\lambda Q(1)+(1-\lambda)Q(0)$. Indeed, (i) $Q(\lambda)\subset\lambda Q(1)+(1-\lambda)Q(0)$ and (ii) take as a partition of the state space $\{E_1,E_2,\dots\}$ where $E_n=\{2n-1,2n\}$ for $n=1,\dots,$ then the condition holds. Hence, $(Q(\lambda),h)\succ(P,f)$ and therefore, since (Decomposition Indifference) implies that $(Q(\lambda),h)\sim(P,\lambda f+(1-\lambda)g)$, we have: $(P,\lambda f+(1-\lambda)g)\succ(P,f)$.

Proofs for the contraction representation result

Proof of Proposition 2 (Mixture linearity)

Lemma 2 *Under (Information Independence), for every $P, Q \in \mathcal{P}$, $f \in \mathcal{F}$ and $\lambda \in [0,1]$,*

$$U(\lambda P + (1-\lambda)Q, f) = \lambda U(P, f) + (1-\lambda)U(Q, f).$$

Proof. Let $\{p\}, \{q\} \in \mathcal{P}$ be such that $(P, f) \sim (\{p\}, f)$ and $(Q, f) \sim (\{q\}, f)$, respectively. Then, repeated application of (Information Independence) delivers

$$(\lambda P + (1-\lambda)Q, f) \sim (\lambda\{p\} + (1-\lambda)Q, f) \sim (\lambda\{p\} + (1-\lambda)\{q\}, f).$$

Since $U(\lambda\{p\} + (1-\lambda)\{q\}, f) = \lambda U(\{p\}, f) + (1-\lambda)U(\{q\}, f)$ is true for precise information, we obtain the claim. ■

Lemma 3 *Assume Axioms 1 to 6, and 12. Then, for every $P, Q \in \mathcal{P}$ and $\lambda \in [0,1]$, $\varphi(\lambda P + (1-\lambda)Q) = \lambda\varphi(P) + (1-\lambda)\varphi(Q)$.*

Proof. By construction,

$$U(\lambda P + (1-\lambda)Q, f) = \min_{p \in \varphi(\lambda P + (1-\lambda)Q)} \sum_{\omega \in S} u(f(\omega))\, p(\omega).$$

for any $f \in \mathcal{F}$.

By mixture-linearity of U over $\mathcal{P}$, the above is equal to

$$\begin{aligned} \lambda U(P, f) + (1-\lambda)U(Q, f) &= \lambda \min_{p \in \varphi(P)} \sum_{\omega \in S} u(f(\omega))\, p(\omega) + (1-\lambda) \min_{p \in \varphi(Q)} \sum_{\omega \in S} u(f(\omega))\, p(\omega) \\ &= \min_{p \in \lambda\varphi(P) + (1-\lambda)\varphi(Q)} \sum_{\omega \in S} u(f(\omega))\, p(\omega) \end{aligned}$$

for any $f \in \mathcal{F}$. By uniqueness of $\varphi(\cdot)$, we obtain the result. ■

This also serves as a proof of Proposition 2.

Continuity

Lemma 4 *Assume Axioms 1 to 6, and 14. Then, the mapping $\varphi : \mathcal{P}(S) \to \mathcal{P}(S)$ is continuous with respect to the Hausdorff metric for each fixed $S \in \mathcal{S}$.*

Proof. Let $\{P^n\}$ be a sequence in $\mathcal{P}(S)$ converging to $P \in \mathcal{P}(S)$. Because $\mathcal{P}(S)$ is compact, without loss of generality we assume that $\{\varphi(P^n)\}$ is convergent. Suppose $\varphi^* \equiv \lim_{n\to\infty} \varphi(P^n) \neq \varphi(P)$. Then there exists $f \in \mathcal{F}$ such that

$$\begin{aligned} U(P,f) &= \min_{p\in\varphi(P)} \sum_{\omega\in S} u(f(\omega))\, p(\omega) \\ &> \min_{p\in\varphi^*} \sum_{\omega\in S} u(f(\omega))\, p(\omega) \\ &= \lim_{n\to\infty} \min_{p\in\varphi(P^n)} \sum_{\omega\in S} u(f(\omega))\, p(\omega) \\ &= \lim_{n\to\infty} U(P^n,f), \end{aligned}$$

which is a contradiction to (Information Continuity). ■

Proof of Lemma 1

We now prove Lemma 1. Assume Axioms 1 to 4 and 7. Or, assume Axioms 1 to 3, 7, and 12. Let $e = \left(\frac{1}{|S|}, \cdots, \frac{1}{|A|}\right) \in \mathbb{R}^{|S|}$.

Proof. Because of the above-noted axioms, we have expected utility representation when information is precise. Without loss of generality, we deal with payoff vectors.

($\Rightarrow$): In the payoff space, the unitary condition is equivalent to saying that

$$\langle p, x\rangle \geq \langle q, x\rangle \quad\Longrightarrow\quad \langle \Pi p, \Pi x\rangle \geq \langle \Pi q, \Pi x\rangle.$$

This implies that

$$(p-q)^t x = 0 \quad\Longrightarrow\quad (p-q)^t \Pi^t \Pi x = 0.$$

Take any $\omega \in S$ and let δ_ω be the vector which assigns 1 on the ω-th coordinate and 0 on the others. Then, $(p-q)^t\delta_\omega = 0$ is equivalent to $p_\omega = q_\omega$.

Pick any $\omega', \omega'' \neq \omega$, and take $p, q \in \Delta(S)$ such that $(p-q)_{\omega'} = \alpha > 0$, $(p-q)_{\omega''} = -\alpha$, and all the other coordinates of $p-q$ are zero. Then p and q satisfy the assumption of the above condition and we obtain

$$(p-q)^t \Pi^t \Pi \delta_\omega = \alpha(\Pi^t_{\omega'}\Pi_\omega - \Pi^t_{\omega''}\Pi_\omega) = 0,$$

which implies $\Pi^t_{\omega'}\Pi_\omega = \Pi^t_{\omega''}\Pi_\omega$. Since ω is arbitrary, this implies that all the off-diagonal entries of $\Pi^t\Pi$ are the same. Therefore, all the diagonal entries of $\Pi^t\Pi$ are the same too.

Remaining to show is that the diagonal entries of $\Pi^t\Pi$ cannot be smaller than the off-diagonal entries. Denote the diagonal entry by a, off-diagonal entry by

b. Let p be the vector in $\Delta(S)$ given by $p_\omega = 1-(|S|-1)c$ and $p_{\omega'} = c$ for all $\omega' \neq \omega$, where $1 \leq c \leq \frac{1}{|S|-1}$. Then we have $(\delta_\omega - p)^t \delta_\omega = (|S|-1)c \geq 0$. The unitary invariance condition implies that $(\delta_\omega - p)^t \Pi^t \Pi \delta_\omega \geq 0$. Since $(\delta_\omega - p)^t \Pi^t \Pi \delta_\omega = (|S|-1)c(a-b)$, we obtain the desired result.

($\Leftarrow$): Since

$$\begin{aligned}\langle \Pi p, \Pi x \rangle &= p^t \Pi^t \Pi x \\ &= p^t \left(\lambda I + \frac{1-\lambda}{|S|} E \right) x \\ &= \lambda \langle p, x \rangle + \frac{1-\lambda}{|S|} \langle e, x \rangle,\end{aligned}$$

we obtain the desired property

$$\langle p, x \rangle \geq \langle q, x \rangle \quad \Longrightarrow \quad \langle \Pi p, \Pi x \rangle \geq \langle \Pi q, \Pi x \rangle.$$

■

For later use, we show the following mathematical fact.

Lemma 5 *For any bistochastic matrix Π, the following two statements are equivalent:*
(i) The matrix $\Pi^t \Pi$ takes the form

$$\Pi^t \Pi = \lambda I + \frac{1-\lambda}{|S|} E,$$

for a given $\lambda \in [0,1]$, where I is the identity matrix and E is a matrix in which all the entries are 1.
(ii) The matrix Π satisfies $\|\Pi p - \Pi q\| = \sqrt{\lambda} \|p - q\|$ for any $p, q \in \Delta(S)$ for a given $\lambda \in [0,1]$, where $\| \cdot \|$ denotes the Euclidian norm;

Proof. For the proof, let $e = (\frac{1}{|S|}, \cdots, \frac{1}{|S|})$. Let δ_ω be the vector which assigns 1 on the ω-th coordinate and 0 on the others. Also let Π_ω be the ω-th column vector of Π.
(i) $\Rightarrow$ (ii): It follows from

$$\begin{aligned}\langle \Pi p - \Pi q, \Pi p - \Pi q \rangle &= (p-q)^t \Pi^t \Pi (p-q) \\ &= (p-q)^t \left(\lambda I + \frac{1-\lambda}{|S|} E \right) (p-q) \\ &= \lambda (p-q)^t I (p-q) + \frac{1-\lambda}{|S|} (p-q)^t E (p-q) \\ &= \lambda \langle p-q, p-q \rangle,\end{aligned}$$

where the last line follows from $E(p-q) = Ep - Eq = \mathbf{1} - \mathbf{1} = 0$.
(ii) $\Rightarrow$ (i): Condition (ii) is written as

$$\langle \Pi p - \Pi q, \Pi p - \Pi q \rangle = \lambda \langle p-q, p-q \rangle$$

for all $p, q \in \Delta(S)$. In particular, by taking $q = e$, we have

$$\langle \Pi p - \Pi e, \Pi p - \Pi e \rangle = \lambda \langle p-e, p-e \rangle$$

Since $\langle \Pi p, e \rangle = \langle p, e \rangle = \langle e, e \rangle = \frac{1}{|S|}$, the above condition reduces to

$$p^t \Pi^t \Pi p = \lambda p^t p + \frac{1-\lambda}{|S|}$$

By taking $p = \delta_\omega$, we have

$$\Pi_\omega^t \Pi_\omega = \lambda + \frac{1-\lambda}{|S|},$$

which is the (ω, ω) diagonal entry of $\Pi^t \Pi$. Since ω is arbitrary, all the diagonal entries are the same.

To show the claim for the off-diagonal entries, let $p = \delta_\omega$ and $q = \delta_{\omega'}$, where $\omega' \neq \omega$. Then the condition reduces to

$$\Pi_\omega^t \Pi_\omega - 2\Pi_\omega^t \Pi_{\omega'} + \Pi_{\omega'}^t \Pi_{\omega'} = 2\lambda.$$

Since $\Pi_\omega^t \Pi_\omega = \Pi_{\omega'}^t \Pi_{\omega'} = \lambda + \frac{1-\lambda}{|S|}$, we obtain

$$\Pi_\omega^t \Pi_{\omega'} = \frac{1-\lambda}{|S|}$$

which is the (ω, ω') off-diagonal entry of $\Pi^t \Pi$. Since ω, ω' are arbitrary, all the off-diagonal entries are the same. ■

Unitary invariance

Now we show that our selection mapping φ is unitary-invariant.

Lemma 6 *Assume Axioms 1 to 6, 7 and 13. Then, for any* $P \in \mathcal{P}(S)$ *and* $\Pi \in \mathcal{T}(S)$, $\varphi(\Pi P) = \Pi\varphi(P)$.

Proof. Suppose $\varphi(\Pi P) \nsubseteq \Pi\varphi(P)$. Then, there is $y \in R^S$ such that

$$\min_{p\in\Pi\varphi(P)} \sum_{\omega\in S} y(\omega)\ p(\omega) > \min_{p'\in\varphi(\Pi P)} \sum_{\omega\in S} y(\omega)\ p'(\omega)$$

By taking $y = \Pi x$, both sides are written as

$$\min_{p\in\varphi(P)} \sum_{\omega\in S} (\Pi x)(\omega)\ (\Pi p)(\omega) > \min_{p'\in\varphi(\Pi P)} \sum_{\omega\in S} (\Pi x)(\omega)\ p'(\omega) \qquad (*)$$

By homogeneity with respect to x, without loss of generality, we can set $x = u \circ f$ and for some $f \in \mathcal{F}$. Take $p^* \in \arg\min_{p\in\varphi(P)} \sum_\omega x(\omega)\ p(\omega)$. Since

$$\sum_{\omega\in S} (\Pi x)(\omega)\ (\Pi p)(\omega) = \lambda \sum_{\omega\in S} x(\omega)\ p(\omega) + \frac{1-\lambda}{|S|} \sum_{\omega\in S} x(\omega),$$

we have $p^* \in \arg\min_{p\in\varphi(P)} \sum_{\omega\in S} (\Pi x)(\omega)\ (\Pi p)(\omega)$. Thus, the left hand side of $(*)$ is equal to $U(\{\Pi p^*\}, \Pi f)$. On the other hand, the right hand side of $(*)$ is $U(\Pi P, \Pi f)$. Thus, $U(\{\Pi p^*\}, \Pi f) > U(\Pi P, \Pi f)$

By definition of p^*, we have $U(\{p^*\}, f) = U(P, f)$. This contradicts (Invariance to Unitary Transformations).

We similarly obtain a contradiction for the case $\varphi(\Pi P) \nsupseteq \Pi\varphi(P)$. ■

Constructing the additive invariant mapping

Below we translate the properties obtained above to the corresponding properties in the Euclidian space of dimension $|S| - 1$, in several steps. Recall the notation $e = (\frac{1}{|S|}, \cdots, \frac{1}{|S|})$. Define $\varphi^* : \mathcal{P}(S) - \{e\} \to \mathcal{P}(S) - \{e\}$ by

$$\varphi^*(K) = \varphi(K + \{e\}) - \{e\}$$

Lemma 7 *Assume Axioms 1 to 6, 7 , 12 and 13. Then, for any* $K \in \mathcal{P}(S) - \{e\}$ *and* $\lambda \geq 0$ *with* $\lambda K \in \mathcal{P}(S) - \{e\}$, $\varphi^*(\lambda K) = \lambda\varphi^*(K)$.

Proof. The case of $\lambda = 0$ or 1 is obvious. Let $\lambda \in (0, 1)$. Then,

$$\begin{aligned}\varphi^*(\lambda K) &= \varphi(\lambda K + \{e\}) - \{e\} \\ &= \varphi(\lambda(K + \{e\}) + (1-\lambda)\{e\}) - \{e\} \\ &= \lambda\varphi(K + \{e\}) + (1-\lambda)\varphi(\{e\}) - \{e\} \\ &= \lambda\varphi(K + \{e\}) + (1-\lambda)\{e\} - \{e\} \\ &= \lambda(\varphi(K + \{e\}) - \{e\}) \\ &= \lambda\varphi^*(K).\end{aligned}$$

The case of $\lambda > 1$ is immediate from the above. ■

Let H_e be the $|S| - 1$ dimensional linear subspace of $\mathbb{R}^S$ which is orthogonal to e. Let $\mathcal{K}_e$ be the family of compact convex subsets of H_e, endowed with the Hausdorff metric. By the above lemma, we extend φ^* to $\mathcal{K}_e$. Also it preserves continuity in the Hausdorff metric. The claim below shows that φ^* satisfies additivity and translation invariance.

Lemma 8 *Assume Axioms 1 to 6, 7 , 12 and 13. Then, for any $K, K' \in \mathcal{K}_e$, $\varphi^*(K + K') = \varphi^*(K) + \varphi^*(K')$. In particular, $\varphi^*(K + \{z\}) = \varphi^*(K) + \{z\}$.*

Proof. Take sufficiently small $\lambda > 0$, then $\lambda K, \lambda K' \in \mathcal{P}(S) - \{e\}$. By homogeneity shown in the previous lemma,

$$\varphi^*(K + K') = \frac{2}{\lambda}\varphi^*\left(\frac{\lambda K + \lambda K'}{2}\right)$$

Then, we have

$$\begin{aligned}\varphi^*\left(\frac{\lambda K + \lambda K'}{2}\right) &= \varphi\left(\frac{\lambda K + \{e\}}{2} + \frac{\lambda K' + \{e\}}{2}\right) - \{e\} \\ &= \frac{1}{2}\varphi(\lambda K + \{e\}) + \frac{1}{2}\varphi(\lambda K' + \{e\}) - \{e\} \\ &= \frac{\varphi(\lambda K + \{e\}) - \{e\}}{2} + \frac{\varphi(\lambda K' + \{e\}) - \{e\}}{2} \\ &= \frac{1}{2}\varphi^*(\lambda K) + \frac{1}{2}\varphi^*(\lambda K') \\ &= \frac{\lambda}{2}\varphi^*(K) + \frac{\lambda}{2}\varphi^*(K'),\end{aligned}$$

where the second line follows from mixture linearity of φ. ■

For later use, we show the following mathematical fact.

Lemma 9 *Let $F : \Delta(S) \to \Delta(S)$ be a mixture-linear mapping satisfying $F(e) = e$. Then there is a unique bistochastic matrix Π such that $F(p) = \Pi p$ for every $p \in \Delta(S)$.*

Proof. Given such F, define Π by $\Pi_{ij} = F_i(\delta_j)$ where δ_j is a probability which assigns unit mass on state $j \in S$. By mixture linearity, Π represents F.

Suppose there are two matrices Π and Π' which represent F. If $\Pi_{ij} \neq \Pi'_{ij}$ for some $i, j \in \Omega$, this leads to $F_i(\delta_j) = \Pi_{ij} \neq \Pi'_{ij} = F_i(\delta_j)$, a contradiction. Thus Π is unique.

If $\Pi_{ij} < 0$ for some $i, j \in S$, this leads to $F_i(\delta_j) < 0$, which is a contradiction.

For any $j \in S$, $\Pi\delta_j = (\Pi_{ij})_{i \in S} \in \Delta(S)$. Therefore, $\sum_{i \in S} \Pi_{ij} = 1$ for each $j \in S$.

Since $\Pi e = e$, for each $i \in N$, $\frac{1}{|S|} \sum_{j \in S} \Pi_{ij} = \frac{1}{|S|}$. Therefore, $\sum_{j \in S} \Pi_{ij} = 1$ for each $i \in S$. ■

Lemma 10 *Assume Axioms 1 to 6, 7 , 12 and 13. Let $G : H_e \to H_e$ be a linear transformation such that $G(\Delta(S) - \{e\}) \subset \Delta(S) - \{e\}$ and there exists $\lambda_G \in (0, 1]$ such that $\|G(x)\| = \lambda_G \|x\|$ for any $x \in H_e$.*
Then, $\varphi^*(GK) = G\varphi^*(K)$ for all $K \in \mathcal{K}_e$.

Proof. Given G, define $F_G : \Delta(S) \to \Delta(S)$ by

$$F_G(p) = G(p - e) + e.$$

Then, it is easy to see that F_G takes values in $\Delta(S)$ and is mixture linear and $F_G(e) = e$. By Lemma 9, it has a representation by a doubly stochastic matrix Π_G and $F_G(p) = \Pi_G p$. Since F_G satisfies the unitary property, Π_G is in $\mathcal{T}(S)$.

By homogeneity of φ^*, without loss of generality we can take $K \in \mathcal{P}(S) - \{e\}$. By Lemma 9, G has a corresponding unitary transformation Π_G and $G(x) = \Pi_G(x + e) - e$ for any $x \in \Delta(S) - \{e\}$.

Then, from the unitary invariance property of φ we have

$$\begin{aligned}
\varphi^*(G(K)) &= \varphi(G(K) + \{e\}) - \{e\} \\
&= \varphi(\Pi_G(K + \{e\}) - \{e\} + \{e\}) - \{e\} \\
&= \varphi(\Pi_G(K + \{e\})) - \{e\} \\
&= \Pi_G \varphi(K + \{e\}) - \{e\} \\
&= G(\varphi^*(K)).
\end{aligned}$$

■

A linear transformation $I : H_e \to H_e$ is called *isometry* if $\|I(x)\| = \|x\|$. Let $\mathcal{I}$ be the set of isometries. For any isometry $I \in \mathcal{I}$, one can take $\lambda > 0$ so that λI satisfies the assumption of Lemma 10. Conversely, any isometry is obtained from a matrix satisfying the assumption of Lemma 10.

By homogeneity of φ^*, we obtain

Lemma 11 *Assume Axioms 1 to 6, 7 , 12 and 13. Then, the mapping φ^* is equivariant in isometries. That is, for any isometry $I \in \mathcal{I}$, $\varphi^*(I(K)) = I(\varphi^*(K))$.*

The $|S|-1$ dimensional Euclidian space $\mathbb{R}^{|S|-1}$ is the image of the linear subspace H_e by some isometry. Let $J : H_e \to \mathbb{R}^{|S|-1}$ be such isometry. All the relevant operations are preserved under isometry. Let $\mathcal{K}^{|S|-1}$ be the space of compact convex subsets of $\mathbb{R}^{|S|-1}$. The space $\mathcal{K}^{|S|-1}$ is also the image of $\mathcal{K}_e$ by the isometry.

Define $\varphi^{**} : \mathcal{K}^{|S|-1} \to \mathcal{K}^{|S|-1}$ by

$$\varphi^{**}(K) = J(\varphi^*(J^{-1}(K))).$$

Then, φ^{**} is continuous, additive and equivariant in isometries in $\mathbb{R}^{|S|-1}$ and translations, and satisfies $\varphi^{**}(K) \subset K$ for any $K \in \mathcal{K}^{|S|-1}$. Continuity of φ^{**} easily follows from that of φ.

Let $W = \{w \in \mathbb{R}^{|S|-1} : \|w\| = 1\}$ be the $|S|-2$ dimensional unit sphere. For a compact convex set $K \in \mathcal{K}^{|S|-1}$, its *Steiner point* is defined by

$$s^{**}(K) = \int_W \arg\max_{p \in K} \langle p, w \rangle \, \nu(dw)$$

where ν is the uniform distribution over W. [18]

Lemma 12 *Assume Axioms 1 to 5, 7, 8, and 12 to 14. Then, there exist $\varepsilon \geq 0$ and $\delta \geq 0$ such that*

$$\varphi^{**}(K) = \varepsilon \left[K - \{s^{**}(K)\}\right] + \delta \left[-K + \{s^{**}(K)\}\right] + \{s^{**}(K)\}.$$

for every $K \in \mathcal{K}^{|S|-1}$.

Proof. Case 1 $|S| = 1, 2$: Obvious.

Case 2 $|S| = 3$: Since the image of a segment is its subsegment, we can apply Theorem 1.8 (b) in Schneider (1974) so that we obtain

$$\varphi^{**}(K) = \varepsilon T_1 \left[K - \{s^{**}(K)\}\right] + \delta T_2 \left[-K + \{s^{**}(K)\}\right] + \{s^{**}(K)\}$$

with $\varepsilon \geq 0$, $\delta \geq 0$ and T_1, T_2 being some two dimensional rotation matrices.

[18] Schneider (1974) has adopted a different definition of Steiner point, but it is equivalent to the current definition, which follows from Theorem 9.4.1 in Aubin and Frankowska (1990).

Consider a segment with midpoint 0. Since its image is its subsegment, (T_1 and T_2) must be the identity or the symmetry with respect to the origin respectively. Thus, without loss of generality

$$\varphi^{**}(K) = \varepsilon\,[K - \{s^{**}(K)\}] + \delta\,[-K + \{s^{**}(K)\}] + \{s^{**}(K)\}$$

Case 3 $|S| \geq 4$: Since $\varphi^{**}(K) \subset K$ for any $K \in \mathcal{K}^{|\Omega|-1}$, the image of any segment is its subsegment. Thus we can apply Theorem 1.8 (b) in Schneider (1974) so that we obtain

$$\varphi^{**}(K) = \varepsilon\,[K - \{s^{**}(K)\}] + \delta\,[-K + \{s^{**}(K)\}] + \{s^{**}(K)\}$$

with $\varepsilon \geq 0$, $\delta \geq 0$. ■

Proof of Theorem 6

Finally, assume Axioms 1 to 5, 7, 8, and 12 to 14. Remember that Axioms 7 and 8 together imply Axiom 6, under Axioms 1 to 4. Since $\varphi(P) \subset P$ holds for all $P \in \mathcal{P}$ here, we have $\varphi^{**}(K) \subset K$ for all $K \in \mathcal{K}^{|\Omega|-1}$.

We show $\varepsilon \in [0, 1]$ and $\delta = 0$. Since $\varphi^{*}(K) \subset K$ for any K, ε cannot exceed 1. Now consider a family of triangles

$$K_\theta = \left\{(x_1, x_2, 0, \cdots, 0) \in \mathbb{R}^{|S|-1} : x_2 \leq \frac{\cos\theta}{\sin\theta} x_1,\ x_2 \geq -\frac{\cos\theta}{\sin\theta} x_1,\ x_1 \leq \sin\theta\right\}$$

indexed by $0 < \theta < \frac{\pi}{2}$. Then we have $s^{**}(K_\theta) = (\frac{\pi-\theta}{\pi}\sin\theta, 0, 0, \cdots, 0)$. Let $\overline{x}_1(K_\theta) = \max_{x \in \varphi^{**}(K_\theta)} x_1$, then we have $\overline{x}_1(K_\theta) = \frac{\pi-\theta}{\pi}\sin\theta + \varepsilon\frac{\theta}{\pi}\sin\theta + \delta\frac{\pi-\theta}{\pi}\sin\theta$. Since $\varphi^{**}(K_\theta) \subset K_\theta$, this cannot exceed $\sin\theta$. Since $\sin\theta$ is positive, we can divide both sides of $\overline{x}_1(K_\theta) \leq \sin\theta$ by $\sin\theta$ and by arranging we get

$$\delta \leq \frac{\frac{\theta}{\pi}}{1 - \frac{\theta}{\pi}}(1 - \varepsilon).$$

Since this is true for any $\theta \in (0, \frac{\pi}{2})$, we obtain $\delta = 0$.

Thus

$$\varphi^{**}(K) = \varepsilon\,[K - \{s^{**}(K)\}] + \{s^{**}(K)\}$$

with $\varepsilon \in [0, 1]$. Since Steiner point and every relevant operation are preserved by isometry, we obtain

$$\varphi(P) = \varepsilon\,[P - \{s(P)\}] + s(P).$$

Constancy of ε with regard to S

Let ε_S be the rate corresponding to S. When $S \subset S'$, since $P \in \mathcal{P}(S)$ implies $P \in \mathcal{P}(S')$, we must have $\varepsilon_S = \varepsilon_{S'}$. For every S, S', since $\varepsilon_S = \varepsilon_{S \cup S'}$, and $\varepsilon_{S'} = \varepsilon_{S \cup S'}$, we obtain the desired claim.

Proof of Theorem 7

Given the equivalence proved in Theorem 4, we show that $\varphi_a(P) \subset \varphi_b(P)$ implies (ii) and then show that (ii) implies (i).

$[\varphi_a(P) \subset \varphi_b(P) \Rightarrow (ii)]$ Since $\pi_a^A(E,P) = s(P)(E) - Min_{p \in \varphi_a(P)} p(E)$ and $\pi_b^A(E,P) = s(P)(E) - Min_{p \in \varphi_b(P)} p(E)$, $\varphi_a(P) \subset \varphi_b(P)$ implies that $\pi_b^A(E,P) \geq \pi_a^A(E,P)$.

$[(ii) \Rightarrow (i)]$ Consider prizes $\bar{x}$ and $\underline{x}$ in X such that both a and b strictly prefer $\bar{x}$ to $\underline{x}$, and let $P \in \mathcal{P}$, and $E \subset \Omega$. For any p, for any agent $i = a, b$, $(\{p\}, \bar{x}_E \underline{x}) \succsim_i [\succ_i](P, \bar{x}_E \underline{x})$ if, and only if, $\pi_i^A(E,P) \geq [\succ] s(P)(E) - p(E)$. Therefore since $\pi_b^A(E,P) \geq \pi_a^A(E,P)$, we have

$$(\{p\}, \bar{x}_E \underline{x}) \succsim_a [\succ_a](\bar{x}_E \underline{x}, P) \Rightarrow (\{p\}, \bar{x}_E \underline{x}) \succsim_b [\succ_b](P, \bar{x}_E \underline{x})$$

which completes the proof that $\succsim_b$ is more averse to bet imprecision than $\succsim_a$.

Proof of Theorem 8

$[(i) \Rightarrow (ii)]$ Let $P \in \mathcal{P}$, and p be a boundary point of P. Define:

$$\overline{\varepsilon} = Sup \{\varepsilon' | \varepsilon' \in [0,1] \text{ s.th. } (\varepsilon' p + (1 - \varepsilon') s(P)) \in \varphi(P)\}.$$

Then $\overline{p} = \overline{\varepsilon} p + (1 - \overline{\varepsilon}) s(P)$ is a boundary point of $\varphi(P)$ since $\varphi(P)$ is closed. Since it is convex as well, there exists a function $\phi : S \to \mathbb{R}$ such that $\int \phi d\overline{p} = \min_{p \in \varphi(P)} \int \phi dp$.

Using the notation and definitions introduced in the proof of Theorem 4 in order to define $f^n = \bar{x}_E \underline{x}$, $p^n, \overline{p^n}$ and Q^n, we have that $(f^n, \{\overline{p^n}\}) \sim (f^n, Q^n)$. Note that $\overline{p^n} = \overline{\varepsilon} p^n + (1 - \overline{\varepsilon}) s(Q^n)$. Thus

$$\pi^R(E, Q^n) = \frac{s(Q^n)(E) - \overline{p^n}(E)}{s(Q^n)(E) - Min_{q \in Q^n} q(E)} \leq \frac{s(Q^n)(E) - \overline{p^n}(E)}{s(Q^n)(E) - p^n(E)} = \overline{\varepsilon}.$$

If $\varepsilon > \overline{\varepsilon}$ we get a contradiction with the fact that $\pi^R(E, Q^n) = \varepsilon$. Therefore, for any boundary point p of P, $\overline{\varepsilon}(p) = Sup \{\varepsilon' | \varepsilon' \in [0,1] \text{ s.th. } (\varepsilon' p + (1 - \varepsilon) s(P)) \in \varphi(P)\}$ is such that $\overline{\varepsilon}(p) \geq \varepsilon$. Let p^* be a boundary point of P such that $\overline{\varepsilon}(p^*) \geq \overline{\varepsilon}(p)$

for all boundary point p of P. Then, there exists a function $\phi : S \to \mathbb{R}$ such that $\int \phi dp^* = \min_{p \in P} \int \phi dp$. Define $\overline{p^*} = \overline{\varepsilon}(p^*)p^* + (1 - \overline{\varepsilon}(p^*))s(P)$ and consider now $p' \in \varphi(P)$. There exists a boundary point p of P and $\varepsilon' < \overline{\varepsilon}(p)$ such that $p' = \varepsilon' p + (1 - \varepsilon')s(P)$.

Let us use again the notation and definition introduced in the proof of Theorem 4 . Since $\int u \circ (\bar{x}_E \underline{x})\, dp^{*n} \leq \int u \circ (\bar{x}_E \underline{x})\, dp^n$ and $\int u \circ (\bar{x}_E \underline{x})\, dp^{*n} \leq \int u \circ (\bar{x}_E \underline{x})\, ds(Q^n)$, we have that $\int u \circ (\bar{x}_E \underline{x})\, d\overline{p^{*n}} \leq \int u \circ (\bar{x}_E \underline{x})\, dp'^n$. Thus $\int u \circ (\bar{x}_E \underline{x})\, d\overline{p^{*n}} = \min_{r \in \varphi(Q^n)} \int u \circ (\bar{x}_E \underline{x})\, dr$ while $\int u \circ (\bar{x}_E \underline{x})\, d\overline{p^{*n}} = \min_{r \in \varphi(Q^n)} \int u \circ (\bar{x}_E \underline{x})\, dr$. Therefore

$$\pi^R(E, Q^n) = \frac{s(Q^n)(E) - \overline{q^*}(E)}{s(Q^n)(E) - Min_{q \in Q^n} q(E)} = \overline{\varepsilon}\,(p^*)\,,$$

and thus $\overline{\varepsilon}\,(p^*) = \varepsilon$. Hence, for all boundary point p of P, $\overline{\varepsilon}\,(p) = \varepsilon$ which proves that $\varphi(P) = \varepsilon P + (1 - \varepsilon)s(P)$.

$[(ii) \Rightarrow (i)]$ Consider $P \in \mathcal{P}$, and $E \subset \Omega$ such that $s(P)(E) \neq Min_{p \in P} p(E)$. We have

$$\min_{p \in \varphi(P)} p(E) = \varepsilon \min_{p \in P} p(E) + (1 - \varepsilon)s(P)(E),$$

and therefore

$$\pi^R(E, P) = \frac{s(P)(E) - Min_{p \in \varphi(P)} p(E)}{s(P)(E) - Min_{p \in P} p(E)} = \varepsilon.$$

References

Ahn, D., 2005. Ambiguity without a state space. mimeo, University of California at Berkeley, forthcoming *Rev. Econ. Stud.*.

Al Najjar, N., 2007. Decision makers as statisticians: diversity, ambiguity and robustness. Technical report. Northwestern University.

Anscombe, F., Aumann, R., 1963. A definition of subjective probability. Annals of Mathematical Statistics 34, 199–205.

Aubin, J.-P., Frankowska, H., 1990. Set-Valued Analysis. Birkhauser, Boston.

Birkoff, G., 1946. Three observations on linear algebra. Univ. Nat. Tucuman Ser A5, 147–151.

Chateauneuf, A., Jaffray, J., 1989. Some characterizations of lower probabilities and other monotone capacities through the use of möbius inversion. Math. Soc. Sci. 17, 263–283.

Ellsberg, D., 1961. Risk, ambiguity, and the Savage axioms. Quart. J. Econ. 75, 643–669.

Epstein, L., 1999. A definition of uncertainty aversion. Rev. Econ. Stud. 66, 579–608.

Epstein, L., Wang, T., 1994. Intertemporal asset pricing under Knightian uncertainty. Econometrica 62 (3), 283–322.

Gajdos, T., Tallon, J.-M., Vergnaud, J.-C., 2004. Decision making with imprecise probabilistic information. J. Math. Econ. 40 (6), 647–681.

Ghirardato, P., Marinacci, M., 2002. Ambiguity aversion made precise: a comparative foundation and some implications. J. Econ. Theory 102, 251–282.

Gilboa, I., Schmeidler, D., 1989. Maxmin expected utility with a non-unique prior. J. Math. Econ. 18, 141–153.

Giraud, R., 2006. Objective imprecise probabilistic information, second order beliefs and ambiguity aversion: an axiomatization. Technical report, Université de Franche Comté.

Jaffray, J.-Y., 1989. Linear utility for belief functions. Operations Research Letters 8, 107–112.

Klibanoff, P., Marinacci, M., Mukerji, S., 2005. A smooth model of decision making under uncertainy. Econometrica (6), 1849–1892.

Kopylov, I., 2006. A parametric model of ambiguity hedging. Technical Report. University of California at Irvine.

Lehrer, E., 2007. Partially-specified probabilities: decisions and games. Technical report. Tel-Aviv University.

Maccheroni, F., Marinacci, M., Rustichini, A., 2006. Ambiguity aversion, robustness, and the variational representation of preferences. Econometrica 74, 1447–1498.

Mukerji, S., Tallon, J.-M., 2001. Ambiguity aversion and incompleteness of financial markets. Rev. Econ. Stud.(68), 883–904.

Olszewski, W., 2007. Preferences over sets of lotteries. Review of Economic Studies 74, 567–595.

Schneider, R., 1974. Equivariant endomorphisms of the space of convex bodies. Transactions of the American Mathmatical Society 194, 53–78.

Schneider, R., 1993. Convex Bodies: the Brunn-Minkowski Theory. Cambridge University Press, Cambridge, Mass.

Stinchcombe, M., 2003. Choice with ambiguity as sets of probabilities. mimeo, University of Texas, Austin.

Tapking, J., 2004. Axioms for preferences revealing subjective uncertainty and uncertainty aversion. J. Math. Econ. 40, 771–797.

Viero, M.-L., 2007. Exactly what happens after the anscombe-aumann race? representing preferences in vague environments. mimeo, Queen's University.

Wang, T., 2003. A class of multi-prior preferences. Technical report, University British Columbia.