Queen's Economics Department Working Paper No. 1094

Exactly What Happens After the Anscombe-Aumann Race? Representing Preferences in Vague Environments

Marie-Louise Vierø
Department of Economics, Queen

Department of Economics
Queen's University
94 University Avenue
Kingston, Ontario, Canada
K7L 3N6

10-2006

Exactly What Happens After the Anscombe-Aumann Race?

Representing Preferences in Vague Environments*

Marie-Louise Vierø†
Queen's University

April 21, 2008

Abstract

This paper derives a representation of preferences for a choice theory with vague environments; vague in the sense that the agent does not know the precise lotteries over outcomes conditional on states. Instead, he knows only a possible set of these lotteries for each state. Thus, this paper's main departure from the standard subjective expected utility model is to relax an assumption about the environment, rather than weakening the axiomatic structure. My model is consistent with the behavior observed in the Ellsberg experiment. It can capture the same type of behavior as the multiple priors models, but can also result in behavior that is different from both the behavior implied by standard subjective expected utility models and the behavior implied by the multiple priors models.

Keywords: Decision Theory, Vagueness, Utility, Optimism
JEL classifications: D800, D810, D000

*This paper is a revised chapter of my Ph.D. dissertation at Cornell University. I am very grateful to David Easley for extensive comments and suggestions. I also thank Larry Blume, Ani Guerdjikova, Edi Karni, Ted O'Donoghue, Maureen O'Hara, Jacob Sagi, seminar participants at the University of Copenhagen, Cornell University, ESEM 2006, Fuqua - Duke University, FUR 2006, Johns Hopkins University, Purdue University, Queen's University, the University of Warwick, the associate editor, and an anonymous referee for comments and suggestions.

†Please address correspondence to: Marie-Louise Vierø, Department of Economics, Dunning Hall Room 316, Queens University, 94 University Avenue, Kingston, Ontario K7L 3N6, Canada; phone: (+1) 613-533-2292; e-mail: viero@econ.queensu.ca

1 Introduction

This paper derives a representation of preferences for a choice theory with vague environments; vague in the sense that the decision maker does not know the precise lotteries over outcomes conditional on states. Instead, he knows only a possible set of these lotteries for each state. Taking the standard Anscombe and Aumann (1963) model as a benchmark, this paper relaxes an assumption about the environment by allowing for acts that are functions from states into sets of lotteries rather than into singleton lotteries. Thus, a vague environment can be thought of as a hybrid between the Anscombe-Aumann environment and an environment considered by Ahn (2008) and Olszewski (2007), who both consider choice between sets of objective lotteries. Anscombe and Aumann's environment is the special case of this paper's environment where all sets are singleton.[1]

As an example, consider a situation where the decision maker is a webhost who has to buy a server for his company. What he cares about is how much he spends on the server and whether or not it will break down so he loses data. The probability that the server will break down depends on whether the server has a high load or a low load. The decision maker can choose between two different servers: either a brand-name server or a recently introduced generic no-brand-name server. For the brand-name server he can go on the web and look up what the probability of failure is if it has a high load, and what the probability of failure is if it has a low load. However, for the new no-brand-name product he cannot find these probabilities, so his information about that server is much less precise. He does know the characteristics of the very best servers and the very worst servers in the market, but he is not sure where the generic server under consideration fits in. This means that he only knows the possible ranges of the probabilities that the server will fail under a high respectively low load, rather than the precise probabilities.

In this example, the states of the world are 'High load' and 'Low load'. The outcomes the decision maker directly cares about are the possible combinations of money left after buying the server and having a working versus a broken server. The environment is vague because the decision maker only knows the possible range of the failure-probability in each state for the generic server.

It seems intuitive that the decision maker will treat the brand-name server and the generic server differently, because he has precise objective information about the proba-

[1] Anscombe and Aumann's model is often explained as a horse race followed by a spin of a roulette, where which roulette is spun depends on which horse wins the race. The present paper's title refers to the fact that it considers decision makers who do not know the exact probabilities on the roulettes.

bilities in each state for one, but only imprecise objective information about the other. This suggests that vagueness is an important feature of the environment, which should be taken into account. As I will demonstrate later, allowing for these vague environments has important economic implications.

The two main results in the paper (Theorems 2 and 3) provide axioms necessary and sufficient for modeling an agent as if he evaluates an act by computing for each state the von Neumann-Morgenstern utility of the best and worst lotteries within the set returned by the act, then weighs these together, assigns unique subjective probabilities to the states, and uses these and the weighted utility for each state to compute his overall utility of the act. The weight on the best lottery has a natural interpretation as a measure of the agent's optimism, or the weight on the worst as pessimism. In the Optimism-Weighted Subjective Expected Utility representation (Theorem 2), these weights, i.e. the decision maker's optimism, are state-independent, while in the Asymmetric Optimism-Weighted Subjective Expected Utility representation (Theorem 3) they are state-dependent.

My axiomatic structure consists of the standard Anscombe-Aumann axioms, properly expanded to my more general class of acts, plus two mild additional axioms. The first of these roughly says that a decision maker does not mind more vagueness if the additional vagueness is caused by adding better possibilities, and that he would never like more vagueness if it is caused by including worse possibilities. The second non-standard axiom is a dominance axiom, which roughly says that a vague act is strictly worse than a precise act if the vagueness is caused by adding only strictly worse lotteries to the precise act.

The framework presented in this paper would be a natural model of a world where there is a continuum of underlying states, which are grouped into coarse discrete states in the decision maker's perception of the world. Say, for example, that the probability of an agent being able to do his job efficiently depends not only on whether his workplace is well-functioning or dysfunctional, but varies depending on how dysfunctional it is. The continuum of states corresponding to different degrees of dysfunctionality can be grouped into the coarse state 'dysfunctional workplace' on which the probability of being efficient ranges from what it is when the workers literally sabotage each other to what it is when the workers just do not eat lunch together. Such a situation is exactly captured by the proposed model. I would like to emphasize that this example of decision makers having a coarse perception of the state space is intended as a motivation. The present paper's model is one of a fixed, known set of states and acts that map these states into sets of probabilities. For a direct treatment of coarse contingencies, see Epstein, Marinacci, and

Seo (2007).

The departure from the standard model that I make in this paper is in a different dimension than the departure made in the literature on ambiguity aversion, which includes, for example, Gilboa and Schmeidler (1989), Ghirardato, Maccheroni, and Marinacci (2004), and Klibanoff, Marinacci, and Mukerji (2005). That literature does not change the environment, but instead relaxes the independence axiom. Other examples of papers that consider non-singleton priors over states are Bewley (1986), Mukerji (1997), Epstein and Schneider (2003), and Gajdos, Hayashi, Tallon, and Vergnaud (forthcoming).

Kreps (1979), Kreps (1992), Nehring (1999), Dekel, Lipman, and Rustichini (2001), and Ozdenoren (2002) derive the decision maker's subjective state space from his preferences. Although they involve sets, the objects of choice are fundamentally different from those in the present paper. In the subjective state space literature the decision maker has a future choice among the alternatives in the set, while in the present paper it is nature that makes the future choice.

Other related papers where the objects of choice involve sets are Ghirardato (2001), Olszewski (2007), and Ahn (2008). Ghirardato (2001) generalizes the Savage framework by allowing the acts to be mappings from states into sets of consequences rather than into unique consequences, while I generalize the Anscombe and Aumann framework and have acts map into sets of lotteries. In addition to the difference in domains, there are also important differences in the representations obtained in Ghirardato (2001) and here, which I will discuss after the main representation theorems in section 3.

Olszewski (2007) and Ahn (2008) provide axioms and representation theorems for a decision maker who chooses between sets of lotteries. Their environments can, adjusting for technical differences, be viewed as one-state versions of the environment I consider, and hence they do not consider the problem of assigning subjective probabilities to states. Conceptually, if we interpret Olszewski's model as a model of 'objective ambiguity', the model in the present paper allows for asymmetric objective ambiguity, where the asymmetry can be both across states and across acts. The representations in my Theorems 1 and 3 then allow for the decision maker's attitude towards objective ambiguity, i.e. his optimism, to be asymmetric across states as well, while Theorem 2 gives axioms necessary and sufficient for his optimism to be state-independent. The relationship to Olszewski will be discussed in detail after Theorem 1.

Finally, a number of recent papers in econometrics have been concerned with set valued random variables, see for example Manski and Tamer (2002). A vague act is exactly a set-valued random variable, thus the present paper provides a choice theory associated with these.

The paper is organized as follows: Section 2 presents the model. In section 3 the axioms on preferences are introduced and the representation theorems are derived. In section 4 the model is applied to a simple contracting problem, and it is shown how the introduction of vagueness changes the optimal contract. Section 5 concludes and discusses directions for further research.

2 A Model of Vague Environments

Let $S = \{1, \ldots, n\}$ be a finite set of states and let $X = \{x_1, \ldots, x_m\}$ be a finite set of consequences, i.e., the outcomes the decision maker directly cares about. Let Δ be the set of all probability distributions, or lotteries, over X. Compound lotteries from Δ are identified by their reduced form lotteries. Finally, let $\mathscr{P}$ be the space of non-empty, compact, and convex polyhedral[2] subsets of Δ. Note that this space includes all singletons from Δ. Define an act h by:

$$h : S \to \mathscr{P}$$
$$h(s) = P_h^s \in \mathscr{P} \text{ for all } s = 1, \ldots, n.$$

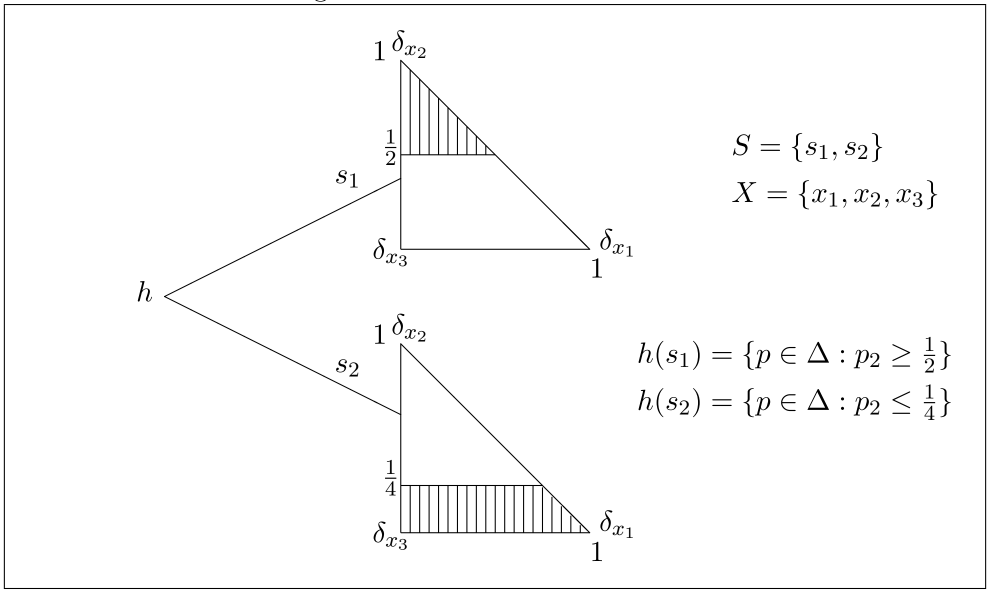
Hence, the acts are functions mapping states into $\mathscr{P}$ (or correspondences mapping states into sets of lotteries with the properties described above). Let $\mathscr{H}$ denote the set of all such acts. The decision maker has preferences over $\mathscr{H}$, represented by the binary relation $\succ$.

Figure 1 illustrates an act h in the case of 2 states and 3 outcomes. Each triangle $\delta_{x_1}\delta_{x_2}\delta_{x_3}$ is a probability simplex Δ. Here, δ_{x_i} is the lottery in the simplex that yields outcome x_i with probability 1. The sets of lotteries $h(s_1)$ and $h(s_2)$ are presented as hatched areas. Thus, all the decision maker knows is that if he chooses the act h the probability of getting x_2 is greater than $\frac{1}{2}$ if state 1 occurs and smaller than $\frac{1}{4}$ if state 2 occurs.

It is important to notice that the decision maker has to choose his action without knowing which state will occur. Each of the acts induces a set of conditional probabilities over the outcomes for each of the states. For some states and acts this set can be a singleton, but generally it is not, as was the case for the generic server in the introductory example. Allowing for the acts to return sets of lotteries in some states permits the decision maker

[2] I.e. with a finite number of vertices.

Figure 1: Illustration of an act h

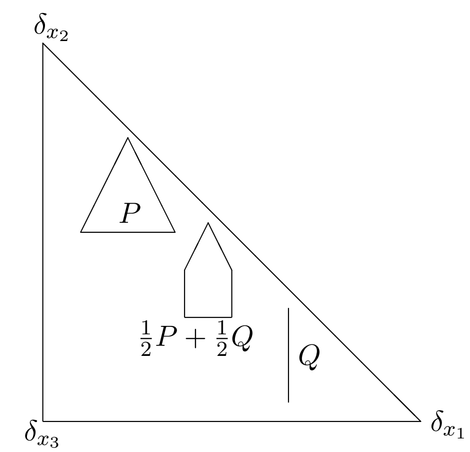


to have a vague idea about the likelihood of different outcomes. However, since the sets typically differ from act to act there is still a difference between acts in the vague states.

Two special cases are worth mentioning. First, if anything is possible in some state we have $h(s) = \Delta$ for that state, i.e. conditional on that state the set an act returns is equal to the entire probability simplex. Second, Anscombe and Aumann's model corresponds to for all acts having $h(s) \in \Delta$ for all states, i.e to having the set be a singleton for all states. Hence their environment is nested within the present environment.

Assuming that the sets are compact and convex with a finite number of vertices is the same as assuming that the decision maker knows the extreme points of the set and that everything in between is possible. It implies that the decision maker has an idea about the best and worst probability that could occur. It also implies that if vagueness of a state results in the agent not knowing the precise probability distribution implied by an act, then his knowledge is only precise enough to give a range of possible probabilities. We will not have a situation where, for example, the agent knows that he will face one of two known probability distributions in a state, or where, as another example, the probability of a particular outcome conditional on some state is either very high or very low, but not around $\frac{1}{2}$.

Figure 2: Convex combination of a triangular set P and a line segment Q



3 Optimism-Weighted Subjective Expected Utility

For $P, Q \in \mathscr{P}$ define the convex combination $\alpha P+(1-\alpha)Q = \{\alpha p+(1-\alpha)q \,|\, p \in P,\ q \in Q\}$. Note that the parameter α is fixed, while we run over all the elements of the sets P and Q. For $h, g \in \mathscr{H}$ the convex combination $\alpha h + (1-\alpha)g$ is taken this way pointwise, i.e. state by state. A convex combination of a triangular set P and a line segment Q in a state s is illustrated in Figure 2, where the triangle $\delta_{x_1}\delta_{x_2}\delta_{x_3}$ is a probability simplex.

I now turn to the axioms. The first three are standard Anscombe and Aumann axioms extended to the more general class of acts considered in the present paper.

Axiom 1 (Preference Relation) *$\succ$ on $\mathscr{H}$ is a preference relation, that is, it is asymmetric (if $h \succ g$ then $g \not\succ h$) and negatively transitive (if $h \not\succ g$ and $g \not\succ f$ then $h \not\succ f$).*

Axiom 2 (Set Independence) *For all acts $h, h', g \in \mathscr{H}$, and for all scalars $\alpha \in (0,1]$, $h \succ h' \Rightarrow \alpha h + (1-\alpha)g \succ \alpha h' + (1-\alpha)g$.*

Axiom 3 (Continuity) *For all acts $h, h', h'' \in \mathscr{H}$, if $h \succ h' \succ h''$ then there exist scalars $\alpha, \beta \in (0,1)$ such that $\alpha h + (1-\alpha)h'' \succ h' \succ \beta h + (1-\beta)h''$.*

Axiom 1 is subject to the standard critique, but the critique is no more severe for this type of acts than for standard acts. Axiom 3 is a standard Archimedean axiom. To see why the independence axiom reasonably extends to sets, note that the axiom implies that indifference curves between singleton lotteries are linear and that the convex combination of two singleton lotteries is worse than the better lottery and better than the worse lottery. The convex combination of two sets consists of lotteries that are all convex combinations of individual lotteries in the two sets. The set independence axiom implies that this convex combination set of lotteries that lie in between lotteries from the two sets in terms of preference will lie between the other two sets in terms of preference as well. Generally, we can think of independence in the present context the usual way, that a decision maker will focus his attention on the differences between acts, and hence making the same substitution for two acts will not alter the preference between them.

The set independence axiom is not rejected by the usual Ellsberg argument. To see this, note that the 2-urn Ellsberg experiment can be reinterpreted as a one-state version of my environment. Here we interpret the information Ellsberg gives about the urns as the decision maker's objective information. We have 1 state, 2 outcomes, and 4 acts, which in the one-state version of the environment are sets of lotteries. Then bets on the urn with unknown proportions of the balls are vague acts, while bets on the urn with known proportions are precise acts. Hence, the Ellsberg experiment fits naturally in the vague framework, and the axioms presented in this paper can generate behavior consistent with that observed in the Ellsberg experiment: for an optimism parameter less than $\frac{1}{2}$, the representation below is consistent with the majority of subjects' behavior of preferring bets on the urn with known proportions. Also, the representation is consistent with the behavior reported by Ellsberg of the two minority groups, who either prefer bets on the urn with unknown proportions or are indifferent between betting on either urn. In my representation these decision makers' optimism parameters are greater than $\frac{1}{2}$ or equal to $\frac{1}{2}$, respectively.

A few definitions are needed before I can present the next two axioms, which are non-standard. Define weak preference $\succsim$ by $h \succsim g$ if $g \not\succ h$, and indifference by $h \sim g$ if $h \not\succ g$ and $g \not\succ h$. Let $P_s h$ denote the act that returns the set $P \in \mathscr{P}$ in state s and agrees with act h in all other states $s' \neq s$:

$$\begin{aligned} &P_s h(s) = P \\ &P_s h(s') = h(s') \text{ for all } s' \neq s. \end{aligned}$$

As a special case $p_s h$ denotes the act that returns the singleton lottery $p \in \Delta$ in state s and agrees with act h in all other states. I.e. a lower case letter means that state s returns a singleton lottery.

For sets $P^1, \ldots, P^k \in \mathscr{P}$ let $Co(P^1, \ldots, P^k)$ denote the convex hull of the sets. For acts $h_1, \ldots, h_k$ the convex hull $Co(h_1, \ldots, h_k)$ is taken pointwise, i.e. state by state.

Axiom 4 (Set Convexity) *For all acts $h \in \mathscr{H}$, and for all sets of lotteries $P, Q \in \mathscr{P}$, if $P_s h \succsim Q_s h$ then $P_s h \succsim Co(P_s h, Q_s h) \succsim Q_s h$.*

Set convexity says that if two acts return the same sets in all other states, and if having the set P in state s is at least as good as having the set Q in state s, then having all mixes between P and Q as possibilities in state s is no better than having only the possibilities in P and no worse than having only the possibilities in Q in state s. Hence, the decision maker will not be made worse off by including weakly better possibilities, and will not be made better off by including weakly worse possibilities. Note that the axiom merely requires the decision maker to feel this way about acts that agree in all but a single state s. He is allowed to feel differently when making more complex comparisons of acts that differ in multiple states. As such the axiom is not very restrictive, and it is very natural in the presence of Axiom 2. The set independence axiom implies that any convex combination of two sets is no better than the best set and no worse than the worst set. When we take the convex hull we include all such convex combinations. Since we are only including sets that lie in between the original sets in terms of preference, it is reasonable that doing so will not make the decision maker better off than he is with only the best set and not worse off than he is with only the worst set. Intuitively, Axiom 4 says that the decision maker does not mind more vague acts if the additional vagueness is caused by including (weakly) better lotteries in the possible set. On the other hand, he would never like more vagueness if it is caused by including (weakly) worse lotteries.

Axiom 5 (Dominance) *For all acts $h \in \mathscr{H}$, for all lotteries $p \in \Delta$, and for all sets of lotteries $Q \in \mathscr{P}$, if $p_s h \succ q_s h$ for all $q \in Q$ then $p_s h \succ Co(p_s h, Q_s h) \succ Q_s h$.*

Dominance says that if two acts return the same set in all other states, and if having the lottery p in state s is preferred to having any of the lotteries in the set Q in state s then having all mixes of p and Q as possibilities in state s is worse than having the lottery p for sure, but better than having only the lotteries in Q. Thus, adding strictly better lotteries to a set makes the decision maker better off, while adding only worse possibilities makes him worse off. Intuitively, a vague act is strictly worse than a precise act if the vagueness is caused by including only strictly worse lotteries. On the other hand, the decision maker likes more vagueness if it results from including strictly better lotteries. Note that this axiom also only concerns acts that differ in just a single state.

Axioms 4 and 5 may at first glance seem quite similar, but the important differences are that Axiom 5 applies to strict preference and only concerns disjoint sets, while Axiom 4 applies to weak preference and also concerns how the decision maker feels about sets that might intersect. Axiom 4 implies that no set is better than its best element or worse than its worst element. Therefore the utility of the set can be expressed in terms of the utilities of its best and worst lotteries. Axiom 5 then ensures that the decision maker's weighting of the best and worst lotteries is the same for all sets.

Theorem 1 *Axioms 1 through 5 are necessary and sufficient for the existence of state-dependent Bernoulli utility functions $u_s(\cdot)$ over the outcomes*[3] *and unique state-dependent parameters $\alpha_s \in (0,1)$ that capture the decision maker's level of optimism in state s, such that*

$$\text{for all } h, g \in \mathscr{H},\ h \succsim g$$

$$\text{if and only if}$$

$$\sum_{s=1}^{n}\Big[\alpha_s \sum_{i=1}^{m} \overline{h}^s(x_i)u_s(x_i) + (1-\alpha_s)\sum_{i=1}^{m} \underline{h}^s(x_i)u_s(x_i)\Big] \geq$$

$$\sum_{s=1}^{n}\Big[\alpha_s \sum_{i=1}^{m} \overline{g}^s(x_i)u_s(x_i) + (1-\alpha_s)\sum_{i=1}^{m} \underline{g}^s(x_i)u_s(x_i)\Big],$$

where $\underline{h}^s$ and $\overline{h}^s$ are, respectively, the worst and best lotteries in P_h^s, while $\underline{g}^s$ and $\overline{g}^s$ are, respectively, the worst and best lotteries in P_g^s.

Outline of proof: I use a sequence of lemmas to prove that the axioms are sufficient for the representation. In Lemma 2 the mixture space theorem is used to show that Axioms 1 through 3 are equivalent to existence and linearity of the representation. Lemma 3 then shows that the representation is separable over states. Lemma 4 shows that for mixtures between two acts, mixtures with a higher weight on the better act are preferred to mixtures with a lower weight on the better act, while Lemma 5 shows that an act will be indifferent to a convex combination of a weakly better and a weakly worse act, and, as long as these two acts are not indifferent to each other, this convex combination is unique.

[3] $u_s(\cdot)$ is non-constant and unique up to a positive affine transformation if state s is non-null (defined below).

So far everything has followed from only Axioms 1, 2, and 3. Lemma 6 is the first of the lemmas that invokes Axiom 4. Lemmas 6 through 10 consider acts that differ in only a single state s. Lemma 6 shows that if $h(s)$ is a line, and if the acts that give the endpoints of this line in state s are not indifferent to each other, then there exists a unique convex combination of the latter two acts to which $h(s)$ is indifferent. Lemma 7 shows that if we consider an act h, where $h(s)$ is a subset of a line, then the weighting of the better and worse endpoints of this subset is the same as the weighting of the better and worse endpoints for the larger set, while Lemma 8 shows that this weighting has to be the same for two acts where the sets in state s are both lines and these lines are parallel. Lemmas 7 and 8 build on Lemma 6 but do not further invoke Axiom 4. Lemma 9 extends the result in Lemma 8 to acts where the lines are not parallel. Lemma 9 uses Axiom 4 and it is the only part of the proof of Theorem 1 that invokes Axiom 5. Lemma 10 then shows that if $h(s)$ is any set $P \in \mathscr{P}$ then the act h will be indifferent to an act which in state s returns a line between the best and worst lotteries in P. Lemma 10 directly invokes Axiom 4.

It is now straightforward to show that the weight α_s is the same for all sets in state s. I then show that the representation holds for singleton sets by induction on the size of the support of the lotteries and Lemma 3. Finally, Lemmas 3, 6, and 10 are used to show that the representation holds for general $P \in \mathscr{P}$.

It is fairly easy to show necessity of the axioms. The detailed proof of Theorem 1 is in the appendix. ■

If we were to only accept Axioms 1 through 3, we would get existence and linearity of the representation, and that it is additively separable over states. If we furthermore accept Axiom 4, we get that each set will be evaluated in terms of its best and worst lotteries. However, the weight on the best and worst could depend on the set. By also accepting Axiom 5 we get that the weighting is the same for all sets.

The interpretation of α_s as capturing the decision maker's optimism in state s can be given the following behavioral justification: Suppose two decision makers with preferences $\succsim_1$ and $\succsim_2$ have the same ranking of singleton sets in state s, i.e. $p_s h \succsim_1 r_s h \Leftrightarrow p_s h \succsim_2 r_s h \ \forall p, r \in \Delta,\ h \in \mathscr{H}$. Then we can think of $\succsim_2$ as more optimistic than than $\succsim_1$ if $P_s h \succsim_1 q_s h \Rightarrow P_s h \succsim_2 q_s h$ and $P_s h \succ_1 q_s h \Rightarrow P_s h \succ_2 q_s h \ \forall P \in \mathscr{P},\ q \in \Delta,\ h \in \mathscr{H}$. That is, if $\succsim_1$ prefers a vague act over a precise act in state s, then $\succsim_2$ always prefers the vague act as well. Therefore, $\succsim_2$ can be viewed as having a more optimistic view of what the outcome of the vagueness will be. A more optimistic decision maker will have a higher

α_s. If α_s approaches zero, we approach the case where the decision maker is extremely pessimistic and takes only the worst possibility into account. If, on the other hand, α_s approaches one, we approach the case where the decision maker is extremely optimistic and takes only the best possibility into account. The limits $\alpha = 0$ and $\alpha = 1$ require that we relax dominance. We get a representation with $\alpha = 0$ if we do not impose Axiom 5 and use the special instance of Axiom 4 where if $P_s h \succsim Q_s h$ then $Co(P_s h, Q_s h) \sim Q_s h$. We get a representation with $\alpha = 1$ if we do not impose Axiom 5 and use the special instance of Axiom 4 where if $P_s h \succsim Q_s h$ then $Co(P_s h, Q_s h) \sim P_s h$.

In each state, indifference curves between singleton lotteries are linear. For general sets there is either a unique best lottery or, if the highest von Neumann-Morgenstern utility is achieved along one of the edges of the set, a continuum thereof. The same applies for the worst lottery or lotteries. For each state, the von Neumann-Morgenstern utility of the best lottery is unique given the representation, as is the von Neumann-Morgenstern utility of the worst lottery.

If we restrict the model to a single state, the representation in Theorem 1 above is the same as that obtained in Olszewski (2007). Olszewski's Theorem 3 applies to a domain which is a one-state version of a vague environment. For this domain he provides axioms that are sufficient for the representation in Theorem 1 when we restrict it to a single state. In essence, by establishing that my axioms are necessary and sufficient for the representation, I show that by strengthening Olszewski's Set S-Independence and Set S-Solvability axioms to my Axioms 2 and 3, his Axiom 6 (Two-Set Union) becomes redundant.[4]

The representation in Theorem 1 has a lot of structure, but for applications even more structure, namely state-independent Bernoulli utility and uniquely determined subjective probabilities over states, is often desirable. The extra structure can be obtained by adding another two standard axioms: a non-triviality axiom and a state independence axiom, also extended to my more general class of acts. Whether the optimism-parameter is state-independent as well, depends on which set of acts the state-independence axiom is imposed

[4]Olszewski's Set S-Solvability follows from my Lemma 5, hence it is implied by my Axioms 1 through 3, and my Axiom 2 is stronger than, and implies, his Set S-Independence. My Axiom 4 and Olszewski's weak GDSB are quite similar, he imposes $A_1 \precsim A_2 \Rightarrow A_1 \precsim PA_1 + (1-P)A_2 \precsim A_2$ for all $P \subset [0,1]$ with the uniqueness property, while I just impose it for $P = [0,1]$ (the convex hull). Weak GDSB with $P = p$ (a singleton) is implied by my Axiom 2. My Axiom 5 is weaker than Olszewski's strict GDSB, since it allows one set to be singleton, and I only impose the axiom for P=[0,1], while he imposes it for all P different from $\{0\}$ and $\{1\}$. Strong GDSB with $P = p$ is implied by my Axioms 2 and 4.

on.

Axiom 6 (Non-triviality) *There exist acts $h, g \in \mathscr{H}$ such that $h \succ g$.*

Define a state s' to be null if for all acts $h, g \in \mathscr{H}$ for which $h(s) = g(s)$ in every $s \neq s'$ we have that $h \sim g$.

Axiom 7 (Set State Independence) *For all acts $h \in \mathscr{H}$, and for all sets of lotteries $P, Q \in \mathscr{P}$, if there exists some state s such that $P_s h \succ Q_s h$, then $P_{s'} h \succ Q_{s'} h$ for all non-null s'.*

Axiom 7' (State Independence) *For all acts $h \in \mathscr{H}$, and for all singleton lotteries $p, q \in \Delta$, if there exists some state s such that $p_s h \succ q_s h$, then $p_{s'} h \succ q_{s'} h$ for all non-null s'.*

Axiom 7 says that the decision maker's preference over sets of lotteries is state-independent, which implies that both the decision maker's preference over singleton lotteries and how he averages lotteries are state-independent. Axiom 7' only requires that the decision maker's preference over singleton lotteries is state-independent. Hence, Axiom 7' is weaker than Axiom 7.

Adding Axioms 6 and 7 to the first five axioms results in Theorem 2, which is the Optimism-Weighted Subjective Expected Utility (OWSEU) representation and the main result in this paper. If we instead impose Axioms 1 through 6 and 7', i.e. relax Set State Independence to State Independence, we get the Asymmetric Optimism-Weighted Subjective Expected Utility (AOWSEU) representation in Theorem 3.[5] Discussion of both these results follows after Theorem 3.

Theorem 2 (Optimism-Weighted Subjective Expected Utility) *Axioms 1 through 7 are necessary and sufficient for the existence of a non-constant, state-independent Bernoulli utility function $u(\cdot)$ over outcomes, a unique probability measure μ over states, and a unique state-independent parameter $\alpha \in (0, 1)$ that captures the decision maker's level of optimism, such that*

$$\text{for all } h, g \in \mathscr{H},\ h \succsim g$$

if and only if

[5] I am grateful to the anonymous referee for suggesting this.

$$\sum_{s=1}^{n} \mu(s) \Big[\alpha \sum_{i=1}^{m} \overline{h}^s(x_i) u(x_i) + (1-\alpha) \sum_{i=1}^{m} \underline{h}^s(x_i) u(x_i) \Big] \geq$$

$$\sum_{s=1}^{n} \mu(s) \Big[\alpha \sum_{i=1}^{m} \overline{g}^s(x_i) u(x_i) + (1-\alpha) \sum_{i=1}^{m} \underline{g}^s(x_i) u(x_i) \Big],$$

where $\underline{h}^s$ *and* $\overline{h}^s$ *are, respectively, the worst and best lotteries in* P_h^s*, while* $\underline{g}^s$ *and* $\overline{g}^s$ *are, respectively, the worst and best lotteries in* P_g^s*. The Bernoulli utility function* $u(\cdot)$ *is unique up to a positive affine transformation.*

Outline of proof: Axiom 6 guarantees that there exists at least one non-null state. That Axiom 7 has to hold for acts where the sets in state s are singletons is used to show that there exists a state-independent Bernoulli utility function. That Axiom 7 has to hold for any sets P and Q is used to show that the optimism-parameter must be the same for all states. Finally I get the subjective probabilities from the scaling of the Bernoulli utility function in the different states (remember that Bernoulli utility functions are unique up to a positive affine transformation). Necessity of the axioms is again fairly easy to show. The detailed proof of Theorem 2 is in the appendix.[6] ■

Theorem 3 (Asymmetric Optimism-Weighted Subjective Expected Utility) *Axioms 1 through 6 and 7' are necessary and sufficient for the existence of a non-constant, state-independent Bernoulli utility function* $u(\cdot)$ *over outcomes, a unique probability measure* μ *over states, and unique state-dependent parameters* $\alpha_s \in (0,1)$ *that capture the decision maker's level of optimism in state* s*, such that*

$$\text{for all } h, g \in \mathscr{H},\ h \succsim g$$

$$\text{if and only if}$$

$$\sum_{s=1}^{n} \mu(s) \Big[\alpha_s \sum_{i=1}^{m} \overline{h}^s(x_i) u(x_i) + (1-\alpha_s) \sum_{i=1}^{m} \underline{h}^s(x_i) u(x_i) \Big] \geq$$

$$\sum_{s=1}^{n} \mu(s) \Big[\alpha_s \sum_{i=1}^{m} \overline{g}^s(x_i) u(x_i) + (1-\alpha_s) \sum_{i=1}^{m} \underline{g}^s(x_i) u(x_i) \Big],$$

[6] There also exists a representation with state-dependent Bernoulli-utility where subjective probabilities over states are not identified. Since I am interested in the the representation most operational for applications, the focus here is on the state-independent representation.

where $\underline{h}^s$ *and* $\overline{h}^s$ *are, respectively, the worst and best lotteries in* P_h^s*, while* $\underline{g}^s$ *and* $\overline{g}^s$ *are, respectively, the worst and best lotteries in* P_g^s*. The Bernoulli utility function* $u(\cdot)$ *is unique up to a positive affine transformation.*

Proof: Please see the appendix. ■

The OWSEU representation in Theorem 2 shows that we can model the decision maker as if he evaluates an act by computing for each state the usual von Neumann-Morgenstern utility of the best lottery and of the worst lottery in that state's set and weighting them together, where the weight on the best lottery can be interpreted as the decision maker's level of optimism. The decision maker assigns unique subjective probabilities to the states and computes his overall utility using these and the weighted utility for each state.

Decision-making in vague environments where acts map states into sets of lotteries is potentially a very complex affair. Pushing the interpretation of the OWSEU representation a bit, we could view the decision maker as if he simplifies his decision making process by taking into consideration only the very best and the very worst among all the lotteries he could possibly come across in each state. Under this interpretation, his simplification of the problem even goes beyond this, since how much emphasis he puts on the best respectively worst lottery is independent of exactly how the possible set of lotteries looks. Furthermore, in Theorem 2 the emphasis on the best and worst is independent of which state will be realized, which is reflected by the state-independence of the optimism-parameter α: the decision maker puts weight α on the best possibility and weight $(1 - \alpha)$ on the worst possibility in the set, regardless of the state. Different acts are then compared using this weighting between the best and worst in constructing a subjective expected utility, assigning unique subjective probabilities to the states.

From the above, the appeal of Theorem 2 for applications is clear. All we need to characterize a decision maker is a Bernoulli utility function, an optimism-parameter, and a unique probability distribution over states characterizing his beliefs. Once we know the Bernoulli utility function, we get the best and worst lotteries in each set for free.

Theorem 3 lies between Theorems 1 and 2. The AOWSEU representation in Theorem 3 allows for asymmetric optimism across states, while at the same time the subjective probability measure is uniquely determined, as is the Bernoulli-utility function (up to a positive affine transformation). In any given state the decision maker's weighting of the best and worst lotteries is independent of the set of lotteries under evaluation, just as in

Theorem 2. However, with asymmetric optimism the weighting of the best and worst will generally differ between states.

Axioms 4, 5, and either Axiom 7 or 7', all play important parts in achieving the simplifying structure. Without the strong form of state independence, the decision maker's optimism will in general depend on the state. Without the weak version of state independence we cannot disentangle the Bernoulli utility functions from the subjective probabilities, and a representation with state-independent utilities will generally not exist. Without the dominance axiom the weighting between the best and the worst lotteries in each state could depend on the shape of the particular set the act returns in that state. If we were to further drop the set convexity axiom, the decision maker could potentially take all lotteries in the possible set into account, and modeling his decision-making process would be highly complex.

If the decision maker were boundedly rational it is very unlikely that he would be able to undertake such highly complex comparisons of acts. It is much more reasonable to expect that he simplifies the problem. Thus, Theorems 2 and 3 could be interpreted as an axiomatization and representation of decision making under some form of bounded rationality.

Returning to the server example from the introduction, we see that which server the decision maker will choose depends on his level of optimism. The more optimistic he is, the more inclined he will be to buy the no-brand-name server. This is the case because, for the brand-name server, the over-all emphasis the decision maker puts on server-failure is independent of his level of optimism since his information about this server is precise. For the generic server, on the other hand, the vagueness results in his over-all emphasis on server-failure being lower the more optimistic he is.

When α differs across states, my model captures behavior that is clearly different from behavior under the multiple priors approach. Again, I use the server example as an illustration. In my environment we have the two states {High load, Low load} $\equiv \{H, L\}$, the relevant outcomes are {Failure, No failure} $\equiv \{F, N\}$, and buying the generic server is a vague act g with $g(H) = [a, b]$ and $g(L) = [c, d]$, where the intervals are for the probability of No failure. Assume that $u(N) > u(F)$, that the states H and L reflect, respectively, a boom and a recession, and that buyers of the server become extremely pessimistic in a recession so that $\alpha_L = 0$, while $\alpha_H > 0$. Then

$$\begin{aligned} AOWSEU(g) &= \mu(H)[\alpha_H b + (1-\alpha_H)a]u(N) + \mu(H)[\alpha_H(1-b) + (1-\alpha_H)(1-a)]u(F) \\ &\quad + (1-\mu(H))cu(N) + (1-\mu(H))(1-c)u(F). \end{aligned}$$

Suppose the producer of the generic server can invest in the following 4 projects:
1) research or advertising raising a,
2) research or advertising raising b,
3) research or advertising raising c, and
4) research or advertising raising d.
Assume also, for simplicity, that he can perfectly price discriminate. When faced with consumers with the asymmetric optimism just described, the producer would pay positive amounts for projects 1), 2), and 3), but nothing for 4). If we were to redefine the problem to take the multiple priors approach, with $S = \{HN, HF, LN, LF\}$ and a set of priors $A = \{\Pi = (\mu_H p_H, \mu_H(1-p_H), (1-\mu_H)p_L, (1-\mu_H)(1-p_L) \mid p_H \in [a,b], p_L \in [c,d]\}$, we would get that the producer of the generic server would pay for 1) if and only if he would pay for 3), and that he would pay for 2) if and only if he would pay for 4), since the multiple priors approach does not allow for asymmetric ambiguity attitude. The example is simple, but the described behavior seems reasonable, and the example illustrates that my model captures more types of behavior than the multiple priors models do. Importantly, my model accomplishes this while maintaining a structure with Bernoulli utility functions, beliefs, and optimism parameters.

Comparison of Theorem 2 with the main representation result in Ghirardato (2001) shows important differences in the representations obtained in the two papers. In Ghirardato's main representation result (see his Theorem 2 and Corollary 1), an act is evaluated by weighting an optimistic and a pessimistic component, but his weights depend on the act under evaluation, his beliefs are non-additive and depend on the weight used in the representation, and the belief function for the pessimistic component is generally different from the belief function for the optimistic component. On the contrary, in my OWSEU-representation in Theorem 2, the weight α on the best lottery is the same across all sets, states, and acts, and beliefs are additive and independent of the act. Ghirardato also considers a representation in which the weight does not depend on the act (see his Corollary 2 and Theorem 3), but beliefs are still non-additive and indexed by the act under evaluation.

Finally, there are many applied problems that make use of probabilities being objective, and where my model is therefore much more appealing than a model with subjective multiple priors. This is, for example, the case in Vierø (2007) where I, among other things, revisit a problem considered in Holmström (1979) of when it will be valuable for a principal to condition a contract on an outside signal. I show that it can be optimal for the principal to condition the contract on an outside vague signal, even if this signal is orthogonal to the directly payoff relevant variables of interest, which provides a nice explanation for the granting of stock options to rank-and-file employees who individually have negligible influence on company performance. This problem is meaningful when probabilities are objective, since it will then be clear whether or not the signal is orthogonal. With subjective beliefs, the problem does not make as much sense.

4 How Vagueness Matters: A Contracting Problem

To illustrate the economic implications of the more general decision-making environments introduced above, this section considers the consequences of vagueness for a simple contracting problem. The introduction of vagueness substantially changes the problem and thus yields different predictions than the standard approach. Even more interesting is the resulting fundamental change in the mechanism behind the optimal contract. Vagueness gives room for the principal to affect which final scenario the agent puts most emphasis on through the design of the contract.

The canonical textbook principal-agent problem with hidden information (see e.g. Mas-Colell et al. (1995)) considers a risk neutral principal and a risk averse agent. The principal wants to hire the agent to complete a task. It is assumed that the agent's utility depends on a variable, here interpreted as his efficiency level, the value of which is realized after the contract is signed. Suppose that the agent's effort can be measured by a one-dimensional variable $e \in [0, \infty)$. The principal's gross profit is a function of effort, $\pi(e)$, with $\pi(0) = 0$, $\pi'(e) > 0$, and $\pi''(e) < 0 \ \forall e$.

The agent's Bernoulli utility function depends on his wage w, how much effort he chooses to exert, and his efficiency x, which affects how much disutility he experiences from effort. Assume for simplicity that there are only two possible values of x: the agent is either of high-efficiency type x_H or of low-efficiency type x_L. Assume that his Bernoulli utility function is of the form

$$u(w, e, x) = v(w - g(e, x)), \ v'(\cdot) > 0, \ v''(\cdot) < 0.$$

Assume also that $g(0, x_H) = g(0, x_L) = g_e(0, x_H) = g_e(0, x_L) = 0$, such that he suffers no disutility if he does not exert any effort, that $g_e(e, x) > 0 \ \forall e > 0$ and $g_{ee}(e, x) > 0 \ \forall e$, such that his disutility from effort is increasing at an increasing rate, and that $g(e, x_L) > g(e, x_H)$ and $g_e(e, x_L) > g_e(e, x_H)$, such that his disutility and marginal disutility from effort are higher if he is of low-efficiency type. Finally, let $\overline{u}$ denote the agent's reservation utility.

One could, for example, think of the principal as a large food processing company and the agent as a potato farmer entering an arrangement where the processing company provides the potato seed and financing for the crop and the farmer puts in his land and labor. The company has a standardized process all of its farmers must follow, a process which is new to the farmer. Therefore, he is not sure how much disutility he will suffer.

The contracting environment is as follows: The senate is currently debating whether to change environmental policies, which would affect farming. There are two possible states of the world. In state 1 the legislation remains unchanged, while in state 2 it is changed. With the current legislation, both parties know that the probability of a farmer being efficient with the production process is p_1. If, however, legislation is changed, both parties are less sure about the probability of the agent being efficient, since the two sides of the senate strongly disagree and the new legislation would likely be some compromise. Therefore, in state 2 the parties only know that the probability of the agent being of high-efficiency type is $p_2 \in Q \subseteq [0, 1]$.

Assume that the principal and the agent both maximize Optimism-Weighted Subjective Expected Utility (OWSEU), that their subjective probabilities of state 1 are μ_1^P and μ_1^A respectively, and that their optimism-parameters are α_P and α_A respectively.

Consider the first-best situation where the value of the parameter x is observable by both contracting parties and would also be verifiable by a court. In this situation the principal maximizes his OWSEU subject to a participation constraint for the agent:

$$\begin{aligned}\max_{w_L, e_L \geq 0, w_H, e_H \geq 0} \mu_1^P \Big[p_1(\pi(e_H) - w_H) + (1 - p_1)(\pi(e_L) - w_L) \Big] \\ +(1 - \mu_1^P) \Big[\alpha_P \Big\{ \overline{p}_2^P(\pi(e_H) - w_H) + (1 - \overline{p}_2^P)(\pi(e_L) - w_L) \Big\} \\ +(1 - \alpha_P) \Big\{ \underline{p}_2^P(\pi(e_H) - w_H) + (1 - \underline{p}_2^P)(\pi(e_L - w_L) \Big\} \Big]\end{aligned}$$

subject to

$$\begin{aligned}\mu_1^A \Big[p_1 v(w_H - g(e_H, x_H)) + (1 - p_1) v(w_L - g(e_L, x_L)) \Big] \\ +(1 - \mu_1^A) \Big[\alpha_A \Big\{ \overline{p}_2^A v(w_H - g(e_H, x_H)) + (1 - \overline{p}_2^A) v(w_L - g(e_L, x_L)) \Big\} \\ +(1 - \alpha_A) \Big\{ \underline{p}_2^A v(w_H - g(e_H, x_H)) + (1 - \underline{p}_2^A) v(w_L - g(e_L, x_L)) \Big\} \Big] \geq \overline{u}\end{aligned}$$

where $\overline{p}_2^P$ and $\underline{p}_2^P$ are the lotteries in Q that are best and worst, respectively, from the principal's point of view, and $\overline{p}_2^A$ and $\underline{p}_2^A$ are the lotteries in Q that are best and worst, respectively, from the agent's point of view. In the terminology of sections 2 and 3, the acts the principal chooses between are all the feasible contracts he could offer, while the acts the agent chooses between are accepting the offered contract or taking the outside option.

Using the first-order conditions of the problem, it is easy to show that, as in the standard model with no vagueness, the participation constraint binds such that the agent gets exactly his reservation utility, and the optimal contract will specify effort levels e_H^* and e_L^* that are both strictly positive with $\pi'(e_H^*) = g_e(e_H^*, x_H)$ and $\pi'(e_L^*) = g_e(e_L^*, x_L)$.

If $\mu_P = \mu_A = 1$ we have the standard problem with no vagueness. The optimal contract $(w_H^*, e_H^*, w_L^*, e_L^*)$ will fully insure the agent against all risk and exactly give him his reservation utility. That is, $w_H^* - g(e_H^*, x_H) = w_L^* - g(e_L^*, x_L)$ and $v(w_H^* - g(e_H^*, x_H)) = v(w_L^* - g(e_L^*, x_L)) = \overline{u}$.

If any of the contracting parties assigns positive probability to the vague state, the result changes substantially. The reason is a fundamental change in the mechanism behind the contracts. Suppose the principal thinks the legislation will remain unchanged and hence has beliefs $\mu_P = 1$ that the precise state will occur, while the agent assigns positive probability to the vague state and hence has beliefs $\mu_A < 1$. Thus, there is vagueness for the agent, but no vagueness for the principal.[7]

To illustrate the intuition and the mechanism that arises from vagueness in an easy and tractable way, consider the following example:

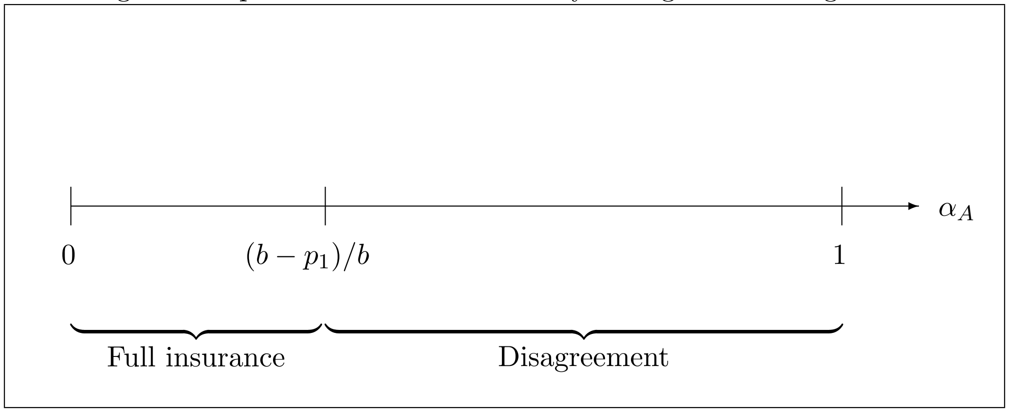
$$\pi(e) = e^{1/2}, \;\; v(\cdot) = log(\cdot), \;\; g(e,x) = \frac{e^2}{x}, \;\; x_L = 1, \;\; x_H = 8, \;\text{ and }\; \overline{u} = 0.$$

Let $\mu_A = p_1 = 1/2$ and $p_2 \in [0, b]$ with $b > \frac{1}{2}$.

Figure 3 illustrates the optimal contract as a function of the agent's optimism. If he is sufficiently optimistic, that is, if $\alpha_A > \frac{b-p_1}{b}$, the optimal contract does not fully insure the agent. Instead, the principal will offer a contract where the parties disagree on which final scenario is the best. The contract makes the agent better off if he turns out to be of type x_L than if he turns out to be of type x_H, but makes the principal better off if the agent turns out to be of type x_H. The agent gets a relatively low compensation if he turns out to be of type x_H, but a relatively high compensation if he turns out to be of type x_L. Since the optimistic agent puts most emphasis on the best scenario for him, in which he is of type x_L, he does not mind the relatively low compensation if he is of type x_H.

[7]The mechanism is even stronger when both parties face vagueness. The assumption that only the agent faces vagueness is made for simplicity. The consequences of vagueness for contracting problems are investigated more comprehensively in Vierø (2007).

Figure 3: Optimal contracts when only the agent faces vagueness



What drives this result is that which lotteries are best and worst for the agent depends on the contract offered. Hence, the presence of vagueness gives room for the principal to affect which final scenario the agent puts most emphasis on through the design of the contract. The principal can exploit the presence of optimism to offer contracts that are better from his point of view. This is a crucial difference from the standard model with no vagueness where, even if the parties have heterogeneous beliefs, the beliefs do not depend on the contract and thus the over-all weights on the different final scenarios are also independent of the contract.

5 Concluding Remarks

In this paper I have derived a representation of preferences for a choice theory with vague environments; vague in the sense that the agent does not know the precise lotteries over outcomes conditional on states. Instead, he knows only a possible set of these lotteries. The result is the Optimism-Weighted Subjective Expected Utility (OWSEU) and Asymmetric Optimism-Weighted Subjective Expected Utility (AOWSEU) representations, where the decision maker evaluates acts by computing for each state the von Neumann-Morgenstern utility of the best and worst lotteries within the set and weighting them together according to his optimism, and then computes his overall utility using these weighted utilities and unique subjective probabilities over the states. The model is consistent with the behavior observed in the Ellsberg experiment. It can capture the same type of behavior as the multiple priors models, but can also result in behavior that is different from both the behavior implied by standard subjective expected utility models and the behavior implied by the multiple priors models.

Given the important implications for contracting problems[8] we can expect vagueness to generate very interesting predictions in other economic applications. For example, vagueness could have interesting effects for no trade results and lemons problems. We can expect trade to occur more often in a vague world than in a world with precise information. With vagueness, willingness to sell an item does not mean that the seller necessarily has information that the item is not worth much, such willingness can now arise as a result of the seller being pessimistic. Consequently, prices will not reveal as much information about asset values in a vague world as they do when there is no vagueness.

Vagueness could also have effects on asset pricing in the following way: the presence of vagueness changes individuals' investment problems and hence their demand for different assets. Specifically, some agents, depending on their level of optimism/pessimism, could want to not participate in the markets for risky assets. The existence of such agents will affect prices in general equilibrium.

These further investigations of the implications of vagueness for economic problems are left for future research.

Appendix

Lemma 1 shows that the set of acts $\mathscr{H}$ is a mixture space.

Lemma 1 *The set of acts $\mathscr{H}$ with the family of functions $\phi_\alpha : \mathscr{H} \times \mathscr{H} \to \mathscr{H}$ for $\alpha \in [0,1]$ defined by $\phi_\alpha(h,g) = \alpha h + (1-\alpha)g$ is a mixture space.*

Proof: I have to show that the three defining properties of a mixture space are satisfied:
(i) $\phi_1(h,g) = 1 * h + 0 * g = h$.
(ii) $\phi_\alpha(h,g) = \alpha h + (1-\alpha)g = \phi_{1-\alpha}(g,h)$.
(iii) I have to show that $\phi_\alpha(\phi_\beta(h,g),g) = \phi_{\alpha\beta}(h,g)$. In order to see this, note that $\phi_\alpha(\phi_\beta(h,g),g) = \{\alpha\beta x + \alpha(1-\beta)y + (1-\alpha)y' \,|\, x \in h, y \in g, y' \in g\}$. To see that $\phi_{\alpha\beta}(h,g) \subseteq \phi_\alpha(\phi_\beta(h,g),g)$ note that $y = y' \Rightarrow \alpha(1-\beta)y + (1-\alpha)y' = (1-\alpha\beta)y$. To see that $\phi_\alpha(\phi_\beta(h,g),g) \subseteq \phi_{\alpha\beta}(h,g)$ we will show that $\{\alpha(1-\beta)y + (1-\alpha)y' \,|\, y \in g, y' \in g\} \subseteq \{(1-\alpha\beta)y \,|\, y \in g\}$. Suppose $t \in \{\alpha(1-\beta)y + (1-\alpha)y' \,|\, y \in g, y' \in g\}$. Then there exists $y, y' \in g$ such that $t = \alpha(1-\beta)y + (1-\alpha)y' = (\alpha(1-\beta)+1-\alpha)\left[\frac{\alpha(1-\beta)}{\alpha(1-\beta)+1-\alpha}y + \frac{1-\alpha}{\alpha(1-\beta)+1-\alpha}y'\right]$,

[8] These are investigated further in Vierø (2007).

where the expression in square brackets is an element of g, since g is convex. Thus there exists $y'' \in g$ such that $t = (\alpha(1-\beta)+1-\alpha)y'' = (1-\alpha\beta)y'' \Rightarrow t \in \{(1-\alpha\beta)y \,|\, y \in g\}$. ■

Theorem 1 is proved in a sequence of lemmas. Lemma 2 shows existence and linearity of the representation.

Lemma 2 *$\succ$ on $\mathscr{H}$ satisfies Axioms 1-3 if and only if there exists $F : \mathscr{H} \to \Re$ such that*
(i) $h \succ g \Leftrightarrow F(h) > F(g)$,
(ii) $F(\alpha h + (1-\alpha)g) = \alpha F(h) + (1-\alpha)F(g)$.

Proof: This follows directly from the mixture space theorem, see Kreps (1988). ■

Lemmas 3, 4, and 5 are pretty standard (see for example Kreps (1988)) as are their proofs. Lemma 3 shows that the representation is separable over states.

Lemma 3 *The function F, defined in Lemma 2, satisfies $F(h) = \sum_{s=1}^{n} F_s(h_s)$.*

Proof: Fix some $h^* = (h^{*1}, \ldots, h^{*n}) \in \mathscr{H}$. For any $h \in \mathscr{H}$, let $h_1 = (h^1, h^{*2}, \ldots, h^{*n})$, $h_2 = (h^{*1}, h^2, \ldots, h^{*n})$, ..., $h_n = (h^{*1}, \ldots, h^{*n-1}, h^n)$. For any $h \in \mathscr{H}$,

$$\frac{1}{n}h + \frac{n-1}{n}h^* = \sum_{s=1}^{n} \frac{1}{n}h^s. \tag{1}$$

To see this recall that convex combinations are done pointwise. Consider state 1 (the proof is the same for the other states). For the coordinate corresponding to state 1 in the acts we have $\sum_{s=1}^{n} \frac{1}{n}h_s^1 = \frac{1}{n}h_1^1 + \frac{1}{n}h_2^1 + \cdots + \frac{1}{n}h_n^1 = \frac{1}{n}h^1 + \frac{1}{n}h^{*1} + \cdots + \frac{1}{n}h^{*1} = \frac{1}{n}h^1 + \frac{n-1}{n}h^{*1}$. Using (1) and Lemma 2,

$$\begin{aligned} \frac{1}{n}\sum_{s=1}^{n} F(h_s) &= F(\sum_{s=1}^{n} \frac{1}{n}h_s) = F(\frac{1}{n}h + \frac{n-1}{n}h^*) \\ &= \frac{1}{n}F(h) + \frac{n-1}{n}F(h^*). \end{aligned} \tag{2}$$

For $s = 1, \ldots, n$, define $F_s : \mathscr{P} \to \Re$ by $F_s(P) = F(h^{*1}, \ldots, h^{*s-1}, P, h^{*s+1}, \ldots, h^{*n}) - \frac{n-1}{n}F(h^*)$. For $h \in \mathscr{H}$, this definition gives $F_s(h^s) = F(h_s) - \frac{n-1}{n}F(h^*)$, which implies that $\frac{1}{n}\sum_{s=1}^{n} F_s(h^s) = \frac{1}{n}\sum_{s=1}^{n} F(h_s) - \frac{n-1}{n}F(h^*)$. Combining this with (2) we get that $\frac{1}{n}\sum_{s=1}^{n} F_s(h^s) = \frac{1}{n}F(h) + \frac{n-1}{n}F(h^*) - \frac{n-1}{n}F(h^*) \Rightarrow F(h) = \sum_{s=1}^{n} F_s(h^s)$. ■

Lemma 4 shows that for mixtures between two acts, mixtures with a higher weight on the better act are preferred to mixtures with a lower weight on the better act.

Lemma 4 *For all* $h, g \in \mathscr{H}$*, if* $h \succ g$ *and* $0 \leq \alpha < \beta \leq 1$ *then* $\beta h + (1-\beta)g \succ \alpha h + (1-\alpha)g$.

Proof: First consider $\alpha = 0$. By Axiom 2, $h \succ g$ and $\beta \in (0,1] \Rightarrow \beta h + (1-\beta)g \succ \beta g + (1-\beta)g = g = \alpha h + (1-\alpha)g$.
Now consider $\alpha > 0$. Since $0 < \frac{\alpha}{\beta} < 1$ and $\beta h + (1-\beta)g \succ g$ by Axiom 2, using Axiom 2 again implies that $\beta h + (1-\beta)g = (1-\frac{\alpha}{\beta})(\beta h + (1-\beta)g + \frac{\alpha}{\beta}(\beta h + (1-\beta)g) \succ (1-\frac{\alpha}{\beta})g + \frac{\alpha}{\beta}(\beta h + (1-\beta)g) = \alpha h + (1-\alpha)g$. ■

Lemma 5 shows that an act will be indifferent to a convex combination of a weakly better and a weakly worse act, and as long as these two acts are not indifferent to each other this convex combination is unique.

Lemma 5 *For all* $h, h_1, h_2 \in \mathscr{H}$*, if* $h_1 \succsim h \succsim h_2$ *and* $h_1 \succ h_2$*, then there exists a unique* $\alpha^* \in [0,1]$ *such that* $h \sim \alpha^* h_1 + (1-\alpha^*)h_2$.

Proof: By Lemma 4, if α^* exists, it is unique. Thus it suffices to show existence.
If $h_1 \sim h$ then $\alpha^* = 1$, and if $h \sim h_2$ then $\alpha^* = 0$. So consider $h_1 \succ h \succ h_2$.
Define $\alpha^* = sup\Big\{\alpha \in [0,1] \,:\, h \succsim \alpha h_1 + (1-\alpha)h_2\Big\}$. $\alpha = 0$ is in the set, guaranteeing we are not taking the sup over an empty set.

Suppose, in order to reach a contradiction, that $\alpha^* h_1 + (1-\alpha^*)h_2 \succ h$. By Axiom 3, since $h \succ h_2$, there exists $\beta \in (0,1)$ such that $\beta(\alpha^* h_1 + (1-\alpha^*)h_2) + (1-\beta)h_2 \succ h \Leftrightarrow \beta\alpha^* h_1 + (1-\beta\alpha^*)h_2 \succ h$. By Lemma 4, if $\alpha^* > \alpha \geq 0$ then $h \succ \alpha h_1 + (1-\alpha)h_2$, so since $\beta\alpha^* < \alpha^*$, $h \succ \beta\alpha^* h_1 + (1-\beta\alpha^*)h_2$. Hence we have a contradiction with Axiom 1.

Suppose now, again to reach a contradiction, that $h \succ \alpha^* h_1 + (1-\alpha^*)h_2$. By Axiom 3, since $h_1 \succ h$, there exists $\beta \in (0,1)$ such that $h \succ \beta(\alpha^* h_1 + (1-\alpha^*)h_2) + (1-\beta)h_1 \Leftrightarrow h \succ (1-\beta(1-\alpha^*))h_1 + \beta(1-\alpha^*)h_2$. By definition of α^*, if $\alpha^* < \alpha \leq 1$ then $\alpha h_1 + (1-\alpha)h_2 \succ h$, so since $\beta < 1 \Rightarrow \beta(1-\alpha^*) < 1-\alpha^* \Rightarrow \alpha^* < 1-\beta(1-\alpha^*)$ we get that $(1-\beta(1-\alpha^*))h_1 + \beta(1-\alpha^*)h_2 \succ h$, and we have a contradiction with Axiom 1.

Since Axiom 1 implies completeness of $\succsim$ we are left with $h \sim \alpha^* h_1 + (1-\alpha^*)h_2$.■

Lemmas 6 through 10 consider acts that differ in only a single state s. Lemma 6 shows that if the set such an act gives in state s is a line, and if the acts that give the endpoints of the line in state s are not indifferent to each other, then the first act is indifferent to a unique convex combination of the latter two. Lemma 6 is the first of the lemmas that invokes Axiom 4. Lemmas 1 through 5 have followed from only Axioms 1, 2 and 3.

Lemma 6 *Let $p_s h, q_s h \in \mathscr{H}$ be acts that differ only in state s. If $p_s h \succ q_s h$ then there exists a unique $\alpha^* \in [0,1]$ such that $Co(p_s h, q_s h) \sim \alpha^* p_s h + (1-\alpha^*) q_s h$.*

Proof: By Axiom 4, $p_s h \succ q_s h \Rightarrow p_s h \succsim Co(p_s h, q_s h) \succsim q_s h$. Then applying Lemma 5 gives the result. ■

Also, note that if $p_s h \sim q_s h$ then $Co(p_s h, q_s h) \sim p_s h$ since $p_s h \succsim Co(p_s h, q_s h) \succsim q_s h \sim p_s h$.

Before reading the following lemmas note that for all $p^1, \ldots, p^k \in \Delta$, $Co(p^1_s h, \ldots, p^k_s h) = Co(p^1, \ldots, p^k)_s h$ since the convex hull is taken pointwise.

Lemma 7 shows that if the set an act gives in state s is a subset of a line, then the weight on the act that gives the better endpoint of this subset in state s is the same as the weight for the act which gives the larger set. Lemma 7 builds on Lemma 6 but does not further invoke Axiom 4.

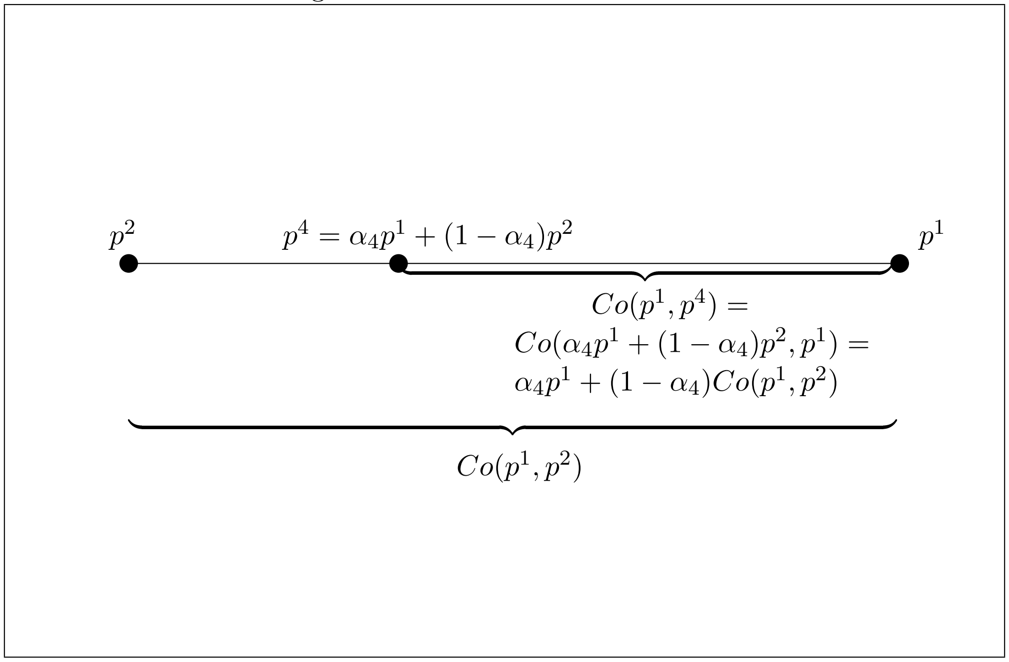
Lemma 7 *Consider $p^1_s h, p^2_s h \in \mathscr{H}$ with $p^1_s h \succ p^2_s h$. There exists a unique $\beta^* \in [0,1]$ such that for all $p^3, p^4 \in Co(p^1, p^2)$ with $p^3_s h \succ p^4_s h$, $Co(p^3_s h, p^4_s h) \sim \beta^* p^3_s h + (1-\beta^*) p^4_s h$.*

Proof: Figure 4 should be helpful when reading the proof. By Lemma 6 there exists a unique $\beta^* \in [0,1]$ such that $Co(p^1_s h, p^2_s h) \sim \beta^* p^1_s h + (1-\beta^*) p^2_s h$. Since $p^3 \in Co(p^1, p^2)$ there exists a unique $\alpha_3 \in [0,1]$ such that $p^3_s h = \alpha_3 p^1_s h + (1-\alpha_3) p^2_s h$, and since $p^4 \in Co(p^1, p^2)$ there exists a unique $\alpha_4 \in [0,1]$ such that $p^4_s h = \alpha_4 p^1_s h + (1-\alpha_4) p^2_s h$.

Note that $\alpha_4 p^1 + (1-\alpha_4) Co(p^1, p^2) = Co(p^1, \alpha_4 p^1 + (1-\alpha_4) p^2) = Co(p^1, p^4)$. Hence $\alpha_4 p^1_s h + (1-\alpha_4) Co(p^1, p^2)_s h = Co(p^1, p^4)_s h$. Thus, by Lemma 2,

$$\begin{aligned} F(Co(p^1, p^4)_s h) &= F(\alpha_4 p^1_s h + (1-\alpha_4) Co(p^1, p^2)_s h) \\ &= \alpha_4 F(p^1_s h) + (1-\alpha_4) F(Co(p^1, p^2)_s h) \\ &= \alpha_4 F(p^1_s h) + (1-\alpha_4)[\beta^* F(p^1_s h + (1-\beta^*) F(p^2_s h)] \\ &= \beta^* F(p^1_s h) + (1-\beta^*) F(p^4_s h). \end{aligned}$$

Figure 4: Illustration for Lemma 7



Next, if $\lambda = \frac{\alpha_3 - \alpha_4}{1-\alpha_4}$ then $\lambda p^1 + (1-\lambda)p^4 = \alpha_3 p^1 + (1-\alpha_3)p^2 = p^3$. Thus $\lambda Co(p^1, p^4) + (1-\lambda)p^4 = Co(\lambda p^1 + (1-\lambda)p^4, p^4) = Co(p^3, p^4)$ and hence $\lambda Co(p^1, p^4)_s h + (1-\lambda)p^4_s h = Co(p^3, p^4)_s h$

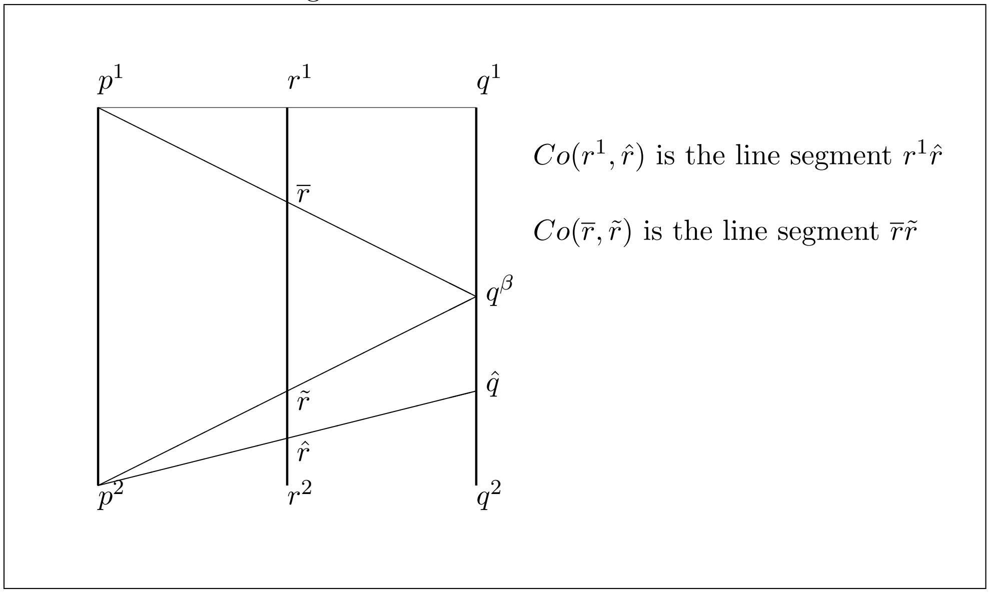
Now, by Lemma 2,

$$\begin{aligned}
F(Co(p^3, p^4)_s h) &= \lambda F(Co(p^1, p^4)_s h) + (1-\lambda)F(p^4_s h) \\
&= \lambda[\beta^* F(p^1_s h) + (1-\beta^*)F(p^4_s h)] + (1-\lambda)F(p^4_s h) \\
&= \lambda\beta^* F(p^1_s h) + (1-\lambda\beta^*)F(p^4_s h) \\
&= \frac{\alpha_3 - \alpha_4}{1-\alpha_4}\beta^* F(p^1) + (1 - \frac{\alpha_3 - \alpha_4}{1-\alpha_4}\beta^*)[\alpha_4 F(p^1_s h) + (1-\alpha_4)F(p^2_s h)] \\
&= \beta^*[\alpha_3 F(p^1_s h) + (1-\alpha_3)F(p^2_s h)] + (1-\beta^*)[\alpha_4 F(p^1_s h) + (1-\alpha_4)F(p^2_s h)] \\
&= \beta^* F(p^3_s h) + (1-\beta^*)F(p^4_s h). \quad \blacksquare
\end{aligned}$$

Note that if $p^1_s h \succ p^2_s h$ and $p^3_s h \sim p^4_s h$ then we have $Co(p^3_s h, p^4_s h) \sim \alpha p^3_s h + (1-\alpha)p^4_s h$ for all $\alpha \in [0,1]$, and thus we still have $Co(p^3_s h, p^4_s h) \sim \beta^* p^3_s h + (1-\beta^*)p^4_s h$, but we have lost uniqueness. If $p^1_s h \sim p^2_s h$ we lose uniqueness as well.

Lemma 8 shows that if we are considering two acts which both give a set that is a line in state s, the better endpoint acts of the two lines are indifferent, the worse endpoint acts

Figure 5: Illustration for Lemma 8



of the two lines are indifferent, and the lines are parallel, then the weight on the act that gives the better endpoint in state s must be the same for the two acts. Again Lemma 8 builds on the previous lemmas but does not directly use Axiom 4.

Lemma 8 *Consider* $p_s^1h, p_s^2h \in \mathscr{H}$ *with* $p_s^1h \succ p_s^2h$ *and* $Co(p_s^1h, p_s^2h) \sim \beta^* p_s^1h + (1-\beta^*)p_s^2h$. *If for* $q_s^1h, q_s^2h \in \mathscr{H}$ *we have* $q_s^1h \sim p_s^1h$, $q_s^2h \sim p_s^2h$ *and* $Co(q^1, q^2)$ *parallel to* $Co(p^1, p^2)$ *then* $Co(q_s^1h, q_s^2h) \sim \beta^* q_s^1h + (1-\beta^*)q_s^2h$.

Proof: By Lemma 6, β^* is unique and there exists a unique $\alpha^* \in [0,1]$ such that $Co(q_s^1h, q_s^2h) \sim \alpha^* q_s^1h + (1-\alpha^*)q_s^2h \equiv q_s^\alpha h$. Suppose, in order to reach a contradiction, that $\beta^* > \alpha^*$.

Suppose first that $\beta^* < 1$. Let $q_s^\beta h \equiv \beta^* q_s^1h + (1-\beta^*)q_s^2h$ and $p_s^\beta h \equiv \beta^* p_s^1h + (1-\beta^*)p_s^2h$. By Lemmas 2 and 4, $p_s^\beta h \sim q_s^\beta h$ and $q_s^\beta h \succ q_s^\alpha h$, so by Axiom 1, $p_s^\beta h \succ q_s^\alpha h$. Furthermore, if we define $\hat{q}_s h \equiv \hat{\alpha} q_s^1 h + (1-\hat{\alpha})q_s^2h$ with $\hat{\alpha} \equiv \frac{\beta^* - \alpha^*}{1-\alpha^*} \in (0, \beta^*]$ then $q_s^1h \succ \hat{q}_s h \succ q_s^2h$ by Lemma 4 and $\hat{q} = \hat{\alpha} q^1 + (1-\hat{\alpha})q^2 \in Co(q^1, q^2)$, and thus $Co(q_s^1h, \hat{q}_s h) \sim \alpha^* q_s^1h + (1-\alpha^*)\hat{q}_s h$ by Lemma 7. Note that $\alpha^* q_s^1h + (1-\alpha^*)\hat{q}_s h = \alpha^* q_s^1h + (1-\alpha^*)[\hat{\alpha} q_s^1h + (1-\hat{\alpha})q_s^2h] = \beta^* q_s^1h + (1-\beta^*)q_s^2h$. Thus $Co(q_s^1h, \hat{q}_s h) \sim Co(p_s^1h, p_s^2h)$.
By Lemma 2, $q_s^\beta h \sim Co(p_s^1h, p_s^2h) \Leftrightarrow F(q_s^\beta h) = F(Co(p_s^1h, p_s^2h))$ and $Co(q_s^1h, \hat{q}_s h) \sim$

$Co(p^1_s h, p^2_s h) \Leftrightarrow F(Co(q^1_s h, \hat{q}_s h)) = F(Co(p^1_s h, p^2_s h)$. Using this and Lemma 2 again,

$$F(\lambda q^\beta_s h + (1-\lambda)Co(p^1_s h, p^2_s h)) = \lambda F(q^\beta_s h) + (1-\lambda)F(Co(p^1_s h, p^2_s h)) = F(Co(p^1_s h, p^2_s h))$$

for all $\lambda \in [0,1]$ and

$$F(\mu Co(q^1_s h, \hat{q}_s h) + (1-\mu)Co(p^1_s h, p^2_s h)) = \mu F(Co(q^1_s h, \hat{q}_s h)) + (1-\mu)F(Co(p^1_s h, p^2_s h)) = F(Co(p^1_s h, p^2_s h))$$

for all $\mu \in [0,1]$. Hence,

$$F(\lambda q^\beta_s h + (1-\lambda)Co(p^1_s h, p^2_s h)) = F(\mu Co(q^1_s h, \hat{q}_s h) + (1-\mu)Co(p^1_s h, p^2_s h)) \tag{3}$$

for all $\lambda, \mu \in [0,1]$.

Figure 5 should be helpful for the following. Consider $\lambda = \mu = \frac{1}{2}$ and let $r^1 \equiv \frac{1}{2}p^1 + \frac{1}{2}q^1$ and $r^2 \equiv \frac{1}{2}p^2 + \frac{1}{2}q^2$. $\frac{1}{2}q_\beta + \frac{1}{2}Co(p^1, p^2) = Co(\overline{r}, \tilde{r})$ where $\overline{r} \equiv \frac{1}{2}p^1 + \frac{1}{2}q^\beta$ and $\tilde{r} \equiv \frac{1}{2}p^2 + \frac{1}{2}q^\beta$. Also, $\frac{1}{2}Co(q^1, \hat{q}) + \frac{1}{2}Co(p^1, p^2) = Co(r^1, \hat{r})$ where $\hat{r} \equiv \frac{1}{2}p^2 + \frac{1}{2}\hat{q}$.

Since $Co(p^1, p^2)$ and $Co(q^1, q^2)$ are parallel, $Co(\overline{r}^\beta, \widetilde{r^\beta})$ and $Co(r^1, \hat{r})$ are both subsets of $Co(r^1, r^2)$. Thus, by Lemma 7 there exists a unique $\gamma^* \in [0,1]$ such that $Co(\overline{r}, \tilde{r})_s h \sim \gamma^* \overline{r}_s h + (1-\gamma^*)\tilde{r}_s h$ and $Co(r^1, \hat{r})_s h \sim \gamma^* r^1_s h + (1-\gamma^*)\hat{r}_s h$.

Since by (3) $Co(\overline{r}_s h, \tilde{r}_s h) \sim Co(r^1_s h, \hat{r}_s h)$, we have $\gamma^* F(\overline{r}_s h) + (1-\gamma^*)F(\tilde{r}_s h) = \gamma^* F(r^1_s h) + (1-\gamma^*)F(\hat{r}_s h)$. Using Lemma 2, we can express $F(\overline{r}_s h)$, $F(\tilde{r}_s h)$, $F(\hat{r}_s h)$, and $F(r^1_s h)$ in terms of $F(p^1_s h)$ and $F(p^2_s h)$:

$$\begin{aligned}
F(\overline{r}_s h) &= \frac{1}{2}F(p^1_s h) + \frac{1}{2}F(q^\beta_s h) = \frac{1}{2}F(p^1_s h) + \frac{1}{2}[\beta^* F(p^1_s h) + (1-\beta^*)F(p^2_s h)] \\
&= \frac{1}{2}(1+\beta^*)F(p^1_s h) + \frac{1}{2}(1-\beta^*)F(p^2_s h), \\
F(\tilde{r}_s h) &= \frac{1}{2}F(p^2_s h) + \frac{1}{2}F(q^\beta_s h) = \frac{1}{2}F(p^2_s h) + \frac{1}{2}[\beta^* F(p^1_s h) + (1-\beta^*)F(p^2_s h)] \\
&= \frac{\beta^*}{2}F(p^1_s h) + (1-\frac{\beta^*}{2})F(p^2_s h), \\
F(\hat{r}_s h) &= \frac{1}{2}F(p^2_s h) + \frac{1}{2}F(\hat{q}_s h) = \frac{1}{2}\Big(\frac{\beta^*-\alpha^*}{1-\alpha^*}F(p^1_s h) + (1-\frac{\beta^*-\alpha^*}{1-\alpha^*})F(p^2_s h)\Big) + \frac{1}{2}F(p^2_s h) \\
&= \frac{1}{2}\frac{\beta^*-\alpha^*}{1-\alpha^*}F(p^1_s h) + (1-\frac{1}{2}\frac{\beta^*-\alpha^*}{1-\alpha^*})F(p^2_s h), \text{ and} \\
F(r^1_s h) &= \frac{1}{2}F(p^1_s h) + \frac{1}{2}F(q^1_s h) = F(p^1_s h).
\end{aligned}$$

Therefore, $\gamma^* F(\overline{r}_s h) + (1-\gamma^*)F(\tilde{r}_s h) = \gamma^*\Big[\frac{1}{2}(1+\beta^*)F(p^1_s h) + \frac{1}{2}(1-\beta^*)F(p^2_s h)\Big] + (1-\gamma^*)\Big[\frac{\beta^*}{2}F(p^1_s h) + (1-\frac{\beta^*}{2})F(p^2_s h)\Big] = \frac{\gamma^*+\beta^*}{2}F(p^1_s h) + (1-\frac{\gamma^*+\beta^*}{2})F(p^2_s h)$. Also, $\gamma^* F(r^1_s h) +$

$(1-\gamma^*)F(\hat{r}_sh) = \gamma^*F(p^1_sh) + (1-\gamma^*)\Big[\frac{1}{2}\frac{\beta^*-\alpha^*}{1-\alpha^*}F(p^1_sh) + (1-\frac{1}{2}\frac{\beta^*-\alpha^*}{1-\alpha^*})F(p^2_sh)\Big]$. Thus, in order to have $\gamma^*F(\overline{r}_sh)+(1-\gamma^*)F(\tilde{r}_sh) = \gamma^*F(r^1_sh)+(1-\gamma^*)F(\hat{r}_sh)$, we must have

$$\frac{\gamma^*+\beta^*}{2} = \gamma^* + (1-\gamma^*)\frac{1}{2}\frac{\beta^*-\alpha^*}{1-\alpha^*} \Leftrightarrow \gamma^* = \alpha^*. \tag{4}$$

Now I will do the same exercise the other way around. If $\overline{p}_sh \equiv \overline{\alpha}p^1_sh + (1-\overline{\alpha})p^2_sh$ with $\overline{\alpha} \equiv \frac{\alpha^*}{\beta^*} \in (0,1)$ then $p^1_sh \succ \overline{p}_sh \succ p^2_sh$ by Lemma 4 and $\overline{p} = \overline{\alpha}p^1 + (1-\overline{\alpha})p^2 \in Co(p^1,p^2)$ and thus $Co(\overline{p}_sh, p^2_sh) \sim \beta^*\overline{p}_sh + (1-\beta^*)p^2_sh$ by Lemma 7. Note that $\beta^*\overline{p}_sh + (1-\beta^*)p^2_sh = \alpha^*p^1_sh + (1-\alpha^*p^2_sh$, and therefore $Co(\overline{p}_sh, p^2_sh) \sim Co(q^1_sh, q^2_sh)$. We have that $p^\alpha_sh \sim Co(q^1_sh,q^2_sh) \Leftrightarrow F(p^\alpha_sh) = F(Co(q^1_sh,q^2_sh))$ and $Co(\overline{p}_sh,p^2_sh) \sim Co(q^1_sh,q^2_sh) \Leftrightarrow F(Co(\overline{p}_sh,p^2_sh)) = F(Co(q^1_sh,q^2_sh))$.

Using this and Lemma 2, we get

$$F(\lambda p^\alpha_sh + (1-\lambda)Co(q^1,q^2)) = \lambda F(p^\alpha_sh) + (1-\lambda)F(Co(q^1_sh,q^2_sh)) = F(Co(q^1_sh,q^2_sh))$$

for all $\lambda \in [0,1]$ and

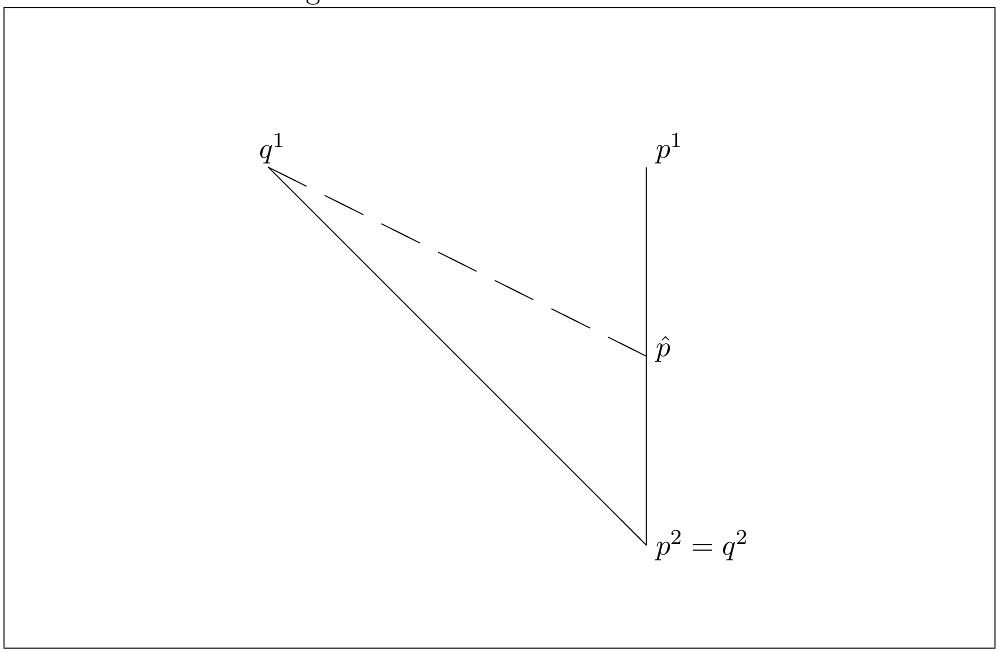
$$\begin{aligned}&F(\mu Co(\overline{p}_sh,p^2_sh) + (1-\mu)Co(q^1_sh,q^2_sh))\\&= \mu F(Co(\overline{p}_sh,p^2_sh)) + (1-\mu)F(Co(q^1_sh,q^2_sh)) = F(q^1_sh,q^2_sh).\end{aligned}$$

Hence

$$F(\lambda p^\alpha_sh + (1-\lambda)Co(q^1,q^2)) = F(\mu Co(\overline{p}_sh,p^2_sh) + (1-\mu)Co(q^1_sh,q^2_sh)). \tag{5}$$

Consider $\lambda = \mu = \frac{1}{2}$. Define $\frac{1}{2}p^\alpha + \frac{1}{2}Co(q^1,q^2) = Co(\tilde{t},\hat{t})$ where $\tilde{t} = \frac{1}{2}p^\alpha + \frac{1}{2}q^1$ and $\hat{t} = \frac{1}{2}p^\alpha + \frac{1}{2}q^2$. Also, $\frac{1}{2}Co(\overline{p},p^2) + \frac{1}{2}Co(q^1,q^2) = Co(\overline{t},t^2)$ where $\overline{t} = \frac{1}{2}\overline{p} + \frac{1}{2}q^1$ and $t^2 = \frac{1}{2}p^2 + \frac{1}{2}q^2$. Since $Co(\tilde{t},\hat{t})$ and $Co(\overline{t},t^2)$ are subsets of $Co(r^1,r^2)$, $Co(\tilde{t},\hat{t})_sh \sim \gamma^*\tilde{t}_sh + (1-\gamma^*)\hat{t}_sh$ and $Co(\overline{t},t^2)_sh \sim \gamma^*\overline{t}_sh + (1-\gamma^*)t^2_sh$. By (5) we have $\gamma^*F(\tilde{t}_sh) + (1-\gamma^*)F(\hat{t}_sh) = \gamma^*F(\overline{t}_sh) + (1-\gamma^*)F(t^2_sh)$. Using Lemma 2, we have
$F(\tilde{t}_sh) = \frac{1}{2}[\alpha^*F(p^1_sh) + (1-\alpha^*)F(p^2_sh)] + \frac{1}{2}F(p^1_sh) = \frac{1}{2}(1+\alpha^*)F(p^1_sh) + \frac{1}{2}(1-\alpha^*)F(p^2_sh)$,
$F(\hat{t}_sh) = \frac{1}{2}[\alpha^*F(p^1_sh) + (1-\alpha^*)F(p^2_sh)] + \frac{1}{2}F(q^2_sh) = \frac{\alpha^*}{2}F(p^1_sh) + (1-\frac{\alpha^*}{2})F(p^2_sh)$,
$F(\overline{t}_sh) = \frac{1}{2}[\overline{\alpha}F(p^1_sh) + (1-\overline{\alpha})F(p^2_sh)] + \frac{1}{2}F(p^1_sh) = \frac{1}{2}(1+\frac{\alpha}{\beta})F(p^1_sh) + \frac{1}{2}(1-\frac{\alpha}{\beta})F(p^2_sh)$, and
$F(t^2_sh) = F(p^2_sh)$.
Thus, $\gamma^*F(\tilde{t}_sh) + (1-\gamma^*)F(\hat{t}_sh) = \gamma^*[\frac{1}{2}(1+\alpha^*)F(p^1_sh) + \frac{1}{2}(1-\alpha^*)F(p^2_sh] + (1-\gamma^*)[\frac{\alpha^*}{2}F(p^1_sh) + (1-\frac{\alpha^*}{2})F(p^2_sh)]$ and $\gamma^*F(\overline{t}_sh) + (1-\gamma^*)F(t^2_sh) = \gamma^*[\frac{1}{2}(1+\frac{\alpha}{\beta})F(p^1_sh) + \frac{1}{2}(1-\frac{\alpha}{\beta})F(p^2_sh)] + (1-\gamma^*)F(p^2_sh)$. Hence $\gamma^*F(\tilde{t}_sh) + (1-\gamma^*)F(\hat{t}_sh) = \gamma^*F(\overline{t}_sh) + (1-\gamma^*)F(t^2_sh) \Leftrightarrow \frac{\gamma^*}{2}(1+\alpha^*) + (1-\gamma^*)\frac{\alpha^*}{2} \Leftrightarrow \gamma^* = \beta^*$. But since by (4) we have $\gamma^* = \alpha^*$ and γ^* is unique this contradicts that $\beta^* > \alpha^*$.

Figure 6: Illustration for Lemma 9



Now suppose that $\beta^* = 1$. Let $p^\alpha_s h = \alpha^* p^1_s h + (1-\alpha^*) p^2_s h$ and note that $Co(q^1, q^2)_s h \sim p^\alpha_s h$. Also, $Co(q^1, q^2)_s h \sim Co(p^\alpha_s h, p^2_s h)$. Therefore, if we define $\hat{u} = \frac{1}{2}p^\alpha + \frac{1}{2}q^1$ and $\tilde{u} = \frac{1}{2}p^\alpha + \frac{1}{2}q^2$, and $u^2 = \frac{1}{2}p_1 + \frac{1}{2}q^2$, we have $Co(\hat{u}, \tilde{u})_s h \sim Co(\hat{u}, u^2)_s h \Leftrightarrow \gamma^* \hat{u} + (1-\gamma^*)\tilde{u} = \gamma^* \hat{u} + (1-\gamma^*) u^2 \Leftrightarrow \gamma^* = 1$.

Let $q_3 \equiv \frac{1+\alpha^*}{2} q^1 + \frac{1-\alpha^*}{2} q^2$ and $q^4 \equiv \frac{\alpha^*}{2} q^1 + (1 - \frac{\alpha^*}{2}) q^2$ and note that $Co(q^3, q^4)_s h \sim p^\alpha_s h$. Therefore $Co(\frac{1}{2}p^\alpha + \frac{1}{2}q^3, \frac{1}{2}p^\alpha + \frac{1}{2}q^4)_s h \sim p^\alpha_s h$. Since $\gamma^* = 1$, $Co(\frac{1}{2}p^\alpha + \frac{1}{2}q^3, \frac{1}{2}p^\alpha + \frac{1}{2}q^4)_s h \sim \frac{1}{2}p^\alpha + \frac{1}{2}q^3 \sim \frac{1}{2}(\alpha^* p^1 + (1-\alpha^*)p^2) + \frac{1}{2}(\frac{1+\alpha^*}{2}p^1 + \frac{1-\alpha^*}{2}p^2)$. Thus $Co(\frac{1}{2}p^\alpha + \frac{1}{2}q^3, \frac{1}{2}p^\alpha + \frac{1}{2}q^4)_s h \sim p^\alpha_s h \Leftrightarrow \frac{\alpha^*}{2} + \frac{1+\alpha^*}{4} = \alpha^* \Leftrightarrow \alpha^* = 1$.

A similar argument gives a contradiction if we assume that $\beta^* < \alpha^*$. Hence we must have $\alpha^* = \beta^*$. ■

Lemma 9 extends the result in Lemma 8 to acts where the lines are not parallel. Lemma 9 uses Axiom 4 and it is the only part of the proof of Theorem 1 that invokes Axiom 5.

Lemma 9 *Consider* $p^1_s h, p^2_s h, q^1_s h, q^2_s h \in \mathscr{H}$. *If* $p^1_s h \sim q^1_s h$ *and* $p^2_s h \sim q^2_s h$ *then* $Co(p^1, p^2)_s h \sim Co(q^1, q^2)_s h$.

Proof: Suppose first that $p^1_s h \sim p^2_s h$. Then the result follows from Axiom 4 and Lemma 2.

So suppose now without loss of generality that $p^1_s h \succ p^2_s h$. (The case $p^2_s h \succ p^1_s h$ is analogous.) By Lemma 6, there exists a unique $\beta^* \in [0,1]$ such that $Co(p^1, p^2)_s h \sim \beta^* p^1_s h + (1-\beta^*) p^2_s h$, and there exists a unique $\alpha^* \in [0,1]$ such that $Co(q^1_s h, q^2_s h) \sim \alpha^* q^1_s h + (1-\alpha^*) q^2_s h$.

Suppose, in order to reach a contradiction, that $\beta^* > \alpha^*$. Then $Co(p^1_s h, p^2_s h) \succ Co(q^1_s h, q^2_s h)$ by Lemmas 2 and 4. Let $\hat{p}_s h \equiv \hat{\alpha} p^1_s h + (1-\hat{\alpha}) p^2_s h$ with $\hat{\alpha} \equiv \frac{\alpha^*}{\beta^*} \in (0,1)$. Since $\hat{p} \in Co(p^1, p^2)$ we have $Co(\hat{p}_s h, p^2_s h) \sim \beta^* \hat{p}_s h + (1-\beta^*) p^2_s h$ by Lemma 7. See figure 6 for an illustration.

Thus $Co(\hat{p}_s h, p^2_s h) \sim \beta^* \hat{p}_s h + (1-\beta^*) p^2_s h = \beta^* \left[\frac{\alpha^*}{\beta^*} p^1_s h + (1 - \frac{\alpha^*}{\beta^*}) p^2_s h\right] + (1-\beta^*) p^2_s h = \alpha^* p^1_s h + (1-\alpha^*) p^2_s h \sim \alpha^* q^1_s h + (1-\alpha^*) q^2_s h \sim Co(q^1_s h, q^2_s h)$. By Lemma 8, without loss of generality we can let $q^2 = p^2$. (If $q^2 \neq p^2$ we can replace it with an act with a parallel segment that has $q^2 = p^2$ by Lemma 8.)

Note that $Co\Big(Co(q^1_s h, q^2_s h), Co(\hat{p}_s h, p^2_s h)\Big) = Co\Big(q^1_s h, Co(\hat{p}_s h, p^2_s h)\Big)$, i.e. the set for state s consists of the triangle $(q^1, \hat{p}, q^2)$.

By Axiom 4, since $Co(q^1_s h, q^2_s h) \sim Co(\hat{p}_s h, p^2_s h)$ we have

$$Co\Big(Co(q^1_s h, q^2_s h), Co(\hat{p}_s h, p^2_s h)\Big) \sim Co(\hat{p}_s h, p^2_s h).$$

By Axiom 5, since $q^1_s h \succ p_s h$ for all $p_s h \in Co(\hat{p}_s h, p^2_s h)$ we have

$$Co\Big(q^1_s h, Co(\hat{p}_s h, p^2_s h)\Big) \succ Co(\hat{p}_s h, p^2_s h).$$

But since $Co\Big(q^1_s h, Co(\hat{p}_s h, p^2_s h)\Big) = Co\Big(Co(q^1_s h, q^2_s h), Co(\hat{p}_s h, p^2_s h)\Big)$ we then have

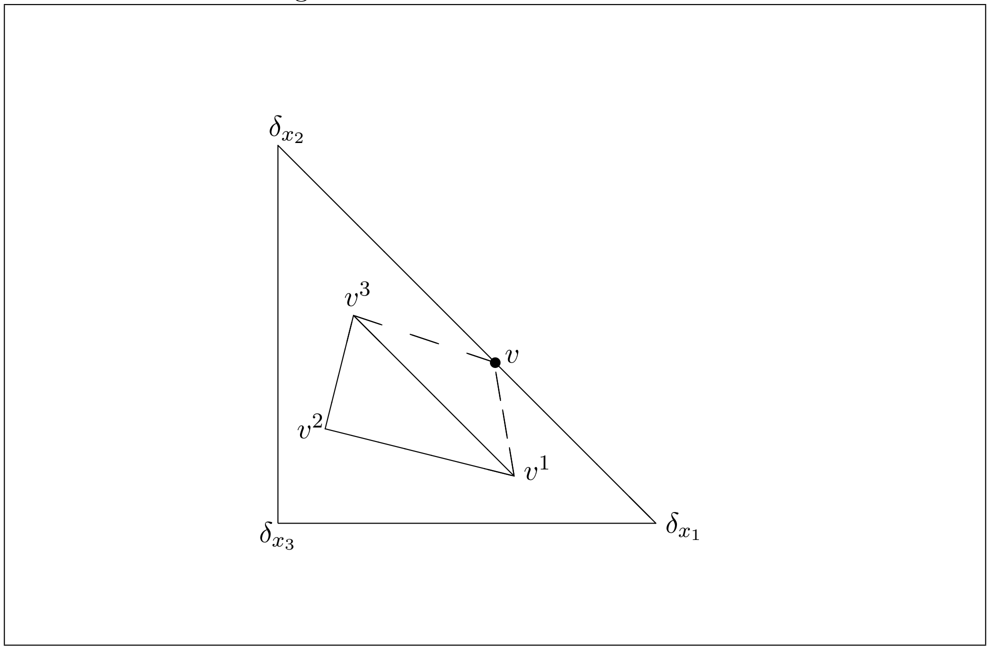
$$Co\Big(q^1_s h, Co(\hat{p}_s h, p^2_s h)\Big) \succ Co\Big(q^1_s h, Co(\hat{p}_s h, p^2_s h)\Big),$$

which is a contradiction with Axiom 1. ■

Lemma 10 shows that if the set the act gives in state s is any set $P \in \mathscr{P}$ then the act will be indifferent to an act that gives a line in state s where the endpoints of the line are the singleton lotteries from P that are considered respectively best and worst in state s. Lemma 10 uses Axiom 4.

Lemma 10 *Consider $P_s h \in \mathscr{H}$. Let $v^k_s h$ denote the act in $P_s h$ for which $v^k_s h \succsim p_s h$ for all $p \in P$, and let $v^1_s h$ denote the act in $P_s h$ for which $p_s h \succsim v^1_s h$ for all $p \in P$. Then $P_s h \sim Co(v^1, v^k)_s h$.*

Figure 7: Illustration for Lemma 10



Proof: Since $P \in \mathscr{P}$, there exist constant singleton acts $v^1, \dots, v^k$ such that $P = Co(v^1, \dots, v^k)$. Without loss of generality, let $v_s^k h \succsim v_s^{k-1} h \succsim \dots \succsim v_s^1 h$. By Lemma 2 $v_s^k h \succsim p_s h$ for all $p \in P$ and $p_s h \succsim v_s^1 h$ for all $p \in P$.

I will proceed by induction on the number of vertices of P (that is, on k). We already have the result for $k = 2$.

For the inductive step, suppose the result is true for acts where P has k vertices. Then $Co(v_s^1 h, \dots, v_s^k h) \sim Co(v_s^1 h, v_s^k h)$. Consider $v_s h$ such that $v_s^k h \succsim v_s h \succsim v_s^1 h$. We want to show that $Co(v_s^1 h, \dots, v_s^k h, v_s h) \sim Co(v_s^1 h, v_s^k h)$. See Figure 7 for an illustration.

If $v_s^k h \sim v_s^1 h$ then $Co(v_s^1 h, v_s^k h) \sim v_s^k h \sim v_s h$, and since $Co(v_s^1 h, \dots, v_s^k h, v_s h) = Co\Big(Co(v_s^1 h, \dots, v_s^k h), v_s h\Big)$ we have $Co(v_s^1 h, \dots, v_s^k h, v_s h) \sim v_s h \sim Co(v_s^1 h, v_s^k h)$ by Axiom 4.

So suppose $v_s^k h \succ v_s^1 h$. Also suppose $Co(v_s^1 h, v_s^k h) \succsim v_s h$. (If $v_s h \succsim Co(v_s^1 h, v_s^k h)$ a similar proof applies.) Note that $Co(v_s^1 h, \dots, v_s^k h, v_s h) = Co\Big(Co(v_s^1 h, \dots, v_s^k h), v_s h\Big)$, so

$$Co(v_s^1 h, v_s^k h) \succsim v_s h \Rightarrow Co(v_s^1 h, v_s^k h) \succsim Co(v_s^1 h, \dots, v_s^k h, v_s h) \tag{6}$$

by Axiom 4.

Now we have to show that $Co(v_s^1 h, \dots, v_s^k h, v_s h) \succsim Co(v_s^1 h, v_s^k h)$. Since $v_s^k h \succsim v_s h \succsim v_s^1 h$ there exists a unique $\hat{v} \in Co(v^1, v^k)$ such that $v_s h \sim \hat{v}_s h \equiv \beta v_s^k h + (1 - \beta) v_s^1 h$. Note that $\beta \leq \alpha^*$ since $Co(v_s^1 h, v_s^k h) \succsim v$.

Case 1: $v_s^k h \succ v_s h \succ v_s^1 h$. Then $v_s^k h \succ \hat{v}_s h \succ v_s^1 h$, which implies that there exists $\gamma \in (0,1)$ such that $\gamma v_s^k h + (1-\gamma) v_s^1 h \equiv v_s^\gamma h \succ \hat{v}_s h$ by Axiom 3. If $\gamma \geq 1 - (1-\alpha^*)\frac{\beta}{\alpha^*}$ then, since $\hat{v}, v_\gamma \in Co(v^1, v^k)$, by Lemma 7 we have $Co(\hat{v}_s h, v_s^\gamma h) \sim \alpha^* v_s^\gamma h + (1-\alpha^*)\hat{v}_s h = \alpha^*[\gamma v_s^k h + (1-\gamma) v_s^1 h] \succsim \alpha^* v_s^k h + (1-\alpha^*) v_s^1 h \sim Co(v_s^1 h, v_s^k h)$. Therefore, because $\beta \leq \alpha^* \Rightarrow 1 - (1-\alpha^*)\frac{\beta}{\alpha^*} \in (0,1)$, it is possible to pick γ such that $Co(\hat{v}_s h, v_s^\gamma h) \succsim Co(v_s^1 h, v_s^k h)$.

By Lemma 9, since $v_s h \sim \hat{v}_s h$, we have $Co(v_s h, v_s^\gamma h) \sim Co(\hat{v}_s h, v_s^\gamma h)$, and since $Co(\hat{v}_s h, v_s^\gamma h) \succsim Co(v_s^1 h, \ldots, v_s^k h)$ we then have $Co(v_s h, v_s^\gamma h) \succsim Co(v_s^1 h, \ldots, v_s^k h)$. Now, since $v_s^\gamma h \in Co(v_s^1 h, v_s^k h)$, it follows that $v_s^\gamma h \in Co(v_s^1 h, \ldots, v_s^k h)$ and therefore

$$Co\Big(Co(v_s^1 h, v_s^\gamma h), Co(v_s^1 h, \ldots, v_s^k h)\Big) = Co(v_s^1 h, \ldots, v_s^k h, v_s h).$$

By Axiom 4,

$$Co(v_s h, v_s^\gamma h) \succsim Co(v_s^1 h, \ldots, v_s^k h) \Rightarrow$$
$$Co\Big(Co(v_s h, v_s^k h), Co(v_s^1 h, \ldots, v_s^k h)\Big) \succsim Co\Big(v_s^1 h, \ldots, v_s^k h\Big).$$

By inductive assumption, $Co\Big(v_s^1 h, \ldots, v_s^k h\Big) \sim Co(v_s^1 h, v_s^k h)$ and thus

$$Co(v_s^1 h, \ldots, v_s^k h, v_s h) \succsim Co(v_s^1 h, v_s^k h).$$

Combining with (6) we have that $Co(v_s^1 h, \ldots, v_s^k h, v_s h) \sim Co(v_s^1 h, v_s^k h)$.

Case 2: $v_s^k h \succ v_s h \sim v_s^1 h$. Then $Co(v_s h, v_s^k h) \sim Co(v_s^1 h, v_s^k h)$ by Lemma 9. Thus, $Co(v_s h, v_s^k h) \succsim Co(v_s^1 h, \ldots v_s^k h)$, which implies

$$Co\Big(Co(v_s h, v_s^k h), Co(v_s^1 h, \ldots, v_s^k h)\Big) \succsim Co(v_s^1 h, \ldots, v_s^k h)$$

by Axiom 4 and thus $Co(v_s^1 h, \ldots, v_s^k h, v_s h) \succsim Co(v_s^1 h, v_s^k h)$. Also, $Co(v_s^1 h, \ldots, v_s^k h) \succsim Co(v_s^1 h, v_s^k h) \Rightarrow Co(v_s^1 h, \ldots, v_s^k h) \succsim Co\Big(Co(v_s^1 h, \ldots, v_s^k h), Co(v_s h, v_s^k h)\Big)$ by Axiom 4, so $Co(v_s^1 h, v_s^k h) \succsim Co(v_s^1 h, \ldots, v_s^k h, v_s h)$.
Combining we get $Co(v_s^1 h, \ldots, v_s^k h, v_s h) \sim Co(v_s^1 h, v_s^k h)$

Case 3: $v_s^k h \sim v_s h \succ v_s^1 h$. Then $Co(v_s^1 h, v_s h) \sim Co(v_s^1 h, v_s^k h)$ by Lemma 9, and the same procedure as in case 2 works. ■

Proof of Theorem 1: **Sufficiency of axioms:** Fix some $h^* \in \mathscr{H}$. Let $\overline{\delta} \in \Delta$ denote the lottery for which $\overline{\delta}_s h^* \succsim p_s h^*$ for all $p \in \Delta$, and let $\underline{\delta} \in \Delta$ denote the singleton lottery for which $p_s h \succsim \underline{\delta}_s h$ for all $p \in \Delta$. If $\overline{\delta}_s h^* \sim \underline{\delta}_s h^*$ then Lemmas 1 through 10 imply that the decision maker is indifferent between anything that could happen in state s.

So suppose $\overline{\delta}_s h^* \sim \underline{\delta}_s h^*$. By Lemma 6, there exists a unique $\alpha_s \in [0,1]$ such that $Co(\overline{\delta}_s h^*, \underline{\delta}_s h^*) \sim \alpha_s \overline{\delta}_s h^* + (1-\alpha_s)\underline{\delta}_s h^*$. By Lemma 7 this implies that for all $p^1, p^2 \in Co(\overline{\delta}, \underline{\delta})$ with $p^1_s h^* \succsim p^2_s h^*$, $Co(p^1_s h^*, p^2_s h^*) \sim \alpha_s p^1_s h^* + (1-\alpha_s) p^2_s h^*$. Since for all $q^1, q^2 \in \Delta$ with $q^1_s h^* \succsim q^2_s h^*$ there exists $p^1, p^2 \in Co(\overline{\delta}, \underline{\delta})$ such that $p^1_s h^* \sim q^1_s h^*$ and $p^2_s h^* \sim q^2_s h^*$, Lemma 9 implies that $Co(q^1_s h^*, q^2_s h^*) \sim \alpha_s q^1_s h^* + (1-\alpha_s) q^2_s h^*$. Then Lemma 10 implies that for all $P \in \mathscr{P}$, $P_s h^* \sim \alpha_s v^k_s h^* + (1-\alpha_s) v^1_s h^*$, where the notation is the same as in Lemma 10. Hence, the weight α_s is the same for all sets in state s.

Now, let $\delta^{x_i} \in \Delta$ denote the lottery that gives the outcome x_i for sure. Define

$$u_s(x_i) \equiv F_s(\delta^{x_i}) \tag{7}$$

Recall that $F_s(\delta^{x_i}) = F(h^{*1}, \ldots, h^{*s-1}, \delta^{x_i}, h^{*s+1}, \ldots, h^{*n}) - \frac{n-1}{n} F(h^*)$. Note that $F_s(\delta^{x_i}) = u_s(x_i) = \alpha_s \sum_{i=1}^m \delta^{x_i}(x_i) u_s(x_i) + (1-\alpha_s) \sum_{i=1}^m \delta^{x_i}(x_i) u_s(x_i)$, so it satisfies Theorem 1.

Now I will proceed by proving that it holds for all singleton lotteries p by induction on the size of the support of p. Suppose that for all $p \in \Delta$ with $|supp\{p\}| = k-1$, $F_s(p) = \sum_{i=1}^{k-1} p(x_i) u_s(x_i)$. Consider $q \in \Delta$ with $|supp\{q\}| = k$ and let $\hat{x} \in supp\{q\}$. Define $\hat{q}$ by

$$\hat{q}(\hat{x}) = 0,$$
$$\hat{q}(x) = \tfrac{q(x)}{1-q(\hat{x})} \text{ for all } x \neq \hat{x}.$$

Note that $|supp\{\hat{q}\}| = k-1$ and $q = q(\hat{x})\delta_{\hat{x}} + (1-q(\hat{x}))\hat{q}$. Hence,

$$\begin{aligned}
F_s(q) &= F(h^{*1}, \ldots, h^{*s-1}, q, h^{*s+1}, \ldots, h^{*n}) - \frac{n-1}{n} F(h^*) \\
&= F(h^{*1}, \ldots, h^{*s-1}, q(\hat{x})\delta_{\hat{x}} + (1-q(\hat{x}))\hat{q}, h^{*s+1}, \ldots, h^{*n}) - \frac{n-1}{n} F(h^*) \\
&= q(\hat{x}) F(h^{*1}, \ldots, h^{*s-1}, \delta^{\hat{x}}, h^{*s+1}, \ldots, h^{*n}) \\
&\quad + (1-q(\hat{x})) F(h^{*1}, \ldots, h^{*s-1}, \hat{q}, h^{*s+1}, \ldots, h^{*n}) - \frac{n-1}{n} F(h^*) \\
&= q(\hat{x}) F_s(\delta^{\hat{x}}) + (1-q(\hat{x})) F_s(\hat{q}) = q(\hat{x}) u_s(\hat{x}) + (1-q(\hat{x})) \sum_{x_i \neq \hat{x}} \frac{q_{x_i}}{1-q(x_i)} u_s(x_i)
\end{aligned}$$

by (7) and the inductive assumption. Thus, we have $F_s(q) = \sum_{i=1}^k q(x_i) u_s(x_i)$. Note that $F_s(q) = \sum_{i=1}^k q(x_i) u_s(x_i) = \alpha_s \sum_{i=1}^m q(x_i) u_s(x_i) + (1-\alpha_s) \sum_{i=1}^m q(x_i) u_s(x_i)$, so again it satisfies Theorem 1.

Finally, for general $P \in \mathscr{P}$, let v^1 denote the worst vertex of P and let v^k denote the best vertex of P. By Lemmas 6 and 10

$$F_s(P) = F(h^{*1}, \ldots, h^{*s-1}, P, h^{*s+1}, \ldots, h^{*n}) - \frac{n-1}{n} F(h^*)$$

$$
\begin{aligned}
&= F(h^{*1},\dots,h^{*s-1},\alpha_s v^k+(1-\alpha_s)v^1,h^{*s+1},\dots,h^{*n})-\frac{n-1}{n}F(h^*)\\
&= \alpha_s F(h^{*1},\dots,h^{*s-1},v^k,h^{*s+1},\dots,h^{*n})\\
&\quad +(1-\alpha_s)F(h^{*1},\dots,h^{*s-1},v^1,h^{*s+1},\dots,h^{*n})-\frac{n-1}{n}F(h^*)\\
&= \alpha_s F_s(v^k)+(1-\alpha_s)F_s(v^1)=\alpha_s\sum_{i=1}^{k}v^k(x_i)u_s(x_i)+(1-\alpha_s)\sum_{i=1}^{k}v^1(x_i)u_s(x_i).
\end{aligned}
$$

Combining this with Lemma 3 we have that $F(h)=\sum_{s=1}^{n}\Big[\alpha_s\sum_{i=1}^{m}\overline{h}^s(x_i)u_s(x_i)+(1-\alpha_s)\sum_{i=1}^{m}\underline{h}^s(x_i)u_s(x_i)\Big]$.

Necessity of axioms: Necessity of Axioms 1 through 3 follows from Lemma 2. To see that Axiom 4 is necessary, suppose the representation holds and that $P_{s'}h\succsim Q_{s'}h$. Denote by $\overline{f}^s$ and $\underline{f}^s$ the best respectively worst lottery in $P_{s'}h(s)$ and by $\overline{g}^s$ and $\underline{g}^s$ the best respectively worst lottery in $Q_{s'}h(s)$. Then, since the representation holds,

$$
\begin{aligned}
P_{s'}h\succsim Q_{s'}h \;\Leftrightarrow\;& \sum_{s=1}^{n}\Big[\alpha_s\sum_{i=1}^{m}\overline{f}^s(x_i)u_s(x_i)+(1-\alpha_s)\sum_{i=1}^{m}\underline{f}^s(x_i)u_s(x_i)\Big]\geq\\
& \sum_{s=1}^{n}\Big[\alpha_s\sum_{i=1}^{m}\overline{g}^s(x_i)u_s(x_i)+(1-\alpha_s)\sum_{i=1}^{m}\underline{g}^s(x_i)u_s(x_i)\Big]\\
\Leftrightarrow\;& \alpha_{s'}\sum_{i=1}^{m}\overline{p}(x_i)u_{s'}(x_i)+(1-\alpha_{s'})\sum_{i=1}^{m}\underline{p}(x_i)u_{s'}(x_i)\geq\\
& \alpha_{s'}\sum_{i=1}^{m}\overline{q}(x_i)u_{s'}(x_i)+(1-\alpha_{s'})\sum_{i=1}^{m}\underline{q}(x_i)u_s(x_i),
\end{aligned}
$$

where

$$
\overline{p}=argmax_{p\in P}\sum_{i=1}^{m}p(x_i)u_{s'}(x_i)\Leftrightarrow\sum_{i=1}^{m}\overline{p}(x_i)u_{s'}(x_i)\geq\sum_{i=1}^{m}p(x_i)u_{s'}(x_i)\text{ for all }p\in P, \quad (8)
$$

$$
\overline{q}=argmax_{q\in Q}\sum_{i=1}^{m}q(x_i)u_{s'}(x_i)\Leftrightarrow\sum_{i=1}^{m}\overline{q}(x_i)u_{s'}(x_i)\geq\sum_{i=1}^{m}q(x_i)u_{s'}(x_i)\text{ for all }q\in Q, \quad (9)
$$

$$
\underline{p}=argmin_{p\in P}\textstyle\sum_{i=1}^{m}p(x_i)u_{s'}(x_i)\Leftrightarrow\sum_{i=1}^{m}p(x_i)u_{s'}(x_i)\geq\sum_{i=1}^{m}\underline{p}(x_i)u_{s'}(x_i)\text{ for all }p\in P, \quad (10)
$$

and

$$
\underline{q}=argmin_{q\in Q}\textstyle\sum_{i=1}^{m}q(x_i)u_{s'}(x_i)\Leftrightarrow\sum_{i=1}^{m}q(x_i)u_{s'}(x_i)\geq\sum_{i=1}^{m}\underline{q}(x_i)u_{s'}(x_i)\text{ for all }q\in Q. \quad (11)
$$

Define $\overline{t} = argmax_{t \in Co(P,Q)} \sum_{i=1}^{m} t(x_i)u_{s'}(x_i)$ and $\underline{t} = argmin_{t \in Co(P,Q)} \sum_{i=1}^{m} t(x_i)u_{s'}(x_i)$. We want to show that

$$\begin{array}{c} \alpha_{s'} \sum_{i=1}^{m} \overline{p}(x_i)u_{s'}(x_i) + (1-\alpha_{s'}) \sum_{i=1}^{m} \underline{p}(x_i)u_{s'}(x_i) \geq \\ \alpha_{s'} \sum_{i=1}^{m} \overline{t}(x_i)u_{s'}(x_i) + (1-\alpha_{s'}) \sum_{i=1}^{m} \underline{t}(x_i)u_{s}(x_i) \geq \\ \alpha_{s'} \sum_{i=1}^{m} \overline{q}(x_i)u_{s'}(x_i) + (1-\alpha_{s'}) \sum_{i=1}^{m} \underline{q}(x_i)u_{s}(x_i). \end{array}$$

Note that for all $t \in Co(P,Q)$ there exists $p \in P, q \in Q, \lambda \in [0,1]$ such that $t = \lambda p + (1-\lambda)q$. Thus, there exists $p^* \in P, q^* \in Q, \lambda^* \in [0,1]$ such that $\overline{t} = \lambda^* p^* + (1-\lambda^*)q^*$. Hence,

$$\begin{aligned} \sum_{i=1}^{m} \overline{t}(x_i)u_{s'}(x_i) &= \sum_{i=1}^{m} (\lambda^* p^*(x_i) + (1-\lambda^*)q^*(x_i))u_{s'}(x_i) \\ &= \lambda^* \sum_{i=1}^{m} p^*(x_i)u_{s'}(x_i) + (1-\lambda^*) \sum_{i=1}^{m} q^*(x_i))u_{s'}(x_i). \end{aligned}$$

From this, (8), and (9), we see that we must have $\sum_{i=1}^{m} p^*(x_i)u_{s'}(x_i) = \sum_{i=1}^{m} \overline{p}(x_i)u_{s'}(x_i)$ and $\sum_{i=1}^{m} q^*(x_i)u_{s'}(x_i) = \sum_{i=1}^{m} \overline{q}(x_i)u_{s'}(x_i)$ (otherwise the sum would not be maximized).

Also, there exists $\hat{p} \in P, \hat{q} \in Q, \mu^* \in [0,1]$ such that $\underline{t} = \lambda^* \hat{p} + (1-\mu^*)\hat{q}$. Hence $\sum_{i=1}^{m} \underline{t}(x_i)u_{s'}(x_i) = \sum_{i=1}^{m} (\mu^* \hat{p}(x_i) + (1-\mu^*)\hat{q}(x_i))u_{s'}(x_i) = \mu^* \sum_{i=1}^{m} \hat{p}(x_i)u_{s'}(x_i) + (1-\mu^*) \sum_{i=1}^{m} \hat{q}(x_i))u_{s'}(x_i)$.

From this, (10), and (11), we see that we must have $\sum_{i=1}^{m} \hat{p}(x_i)u_{s'}(x_i) = \sum_{i=1}^{m} \underline{p}(x_i)u_{s'}(x_i)$ and $\sum_{i=1}^{m} \hat{q}(x_i)u_{s'}(x_i) = \sum_{i=1}^{m} \underline{q}(x_i)u_{s'}(x_i)$ (otherwise the sum would not be minimized). Thus,

$$\begin{array}{l} \alpha_{s'} \sum_{i=1}^{m} \overline{t}(x_i)u_{s'}(x_i) + (1-\alpha_{s'}) \sum_{i=1}^{m} \underline{t}(x_i)u_{s}(x_i) = \\ \alpha_{s'} \sum_{i=1}^{m} [\lambda^* \overline{p}(x_i) + (1-\lambda^*)\overline{q}(x_i)]u_{s'}(x_i) + (1-\alpha_{s'}) \sum_{i=1}^{m} [\mu^* \underline{p}(x_i) + (1-\mu^*)q^*(x_i)]u_{s}(x_i), \end{array}$$

which is equal to

$$\begin{array}{l} \lambda^* \alpha_{s'} \sum_{i=1}^{m} \overline{p}(x_i)u_{s'}(x_i) + \mu^*(1-\alpha_{s'}) \sum_{i=1}^{m} \underline{p}(x_i)u_{s'}(x_i) \\ +(1-\lambda^*)\alpha_{s'} \sum_{i=1}^{m} \overline{q}(x_i)u_{s'}(x_i) + (1-\mu^*)(1-\alpha_{s'}) \sum_{i=1}^{m} \underline{q}(x_i)u_{s'}(x_i). \end{array} \quad (12)$$

Since $\alpha_{s'} \sum_{i=1}^{m} \overline{p}(x_i)u_{s'}(x_i) + (1-\alpha_{s'}) \sum_{i=1}^{m} \underline{p}(x_i)u_{s'}(x_i) \geq \alpha_{s'} \sum_{i=1}^{m} \overline{q}(x_i)u_{s'}(x_i) + (1-\alpha_{s'}) \sum_{i=1}^{m} \underline{q}(x_i)u_{s}(x_i)$ we must have that $\sum_{i=1}^{m} \overline{p}(x_i)u_{s'}(x_i) \geq \sum_{i=1}^{m} \overline{q}(x_i)u_{s'}(x_i)$ or $\sum_{i=1}^{m} \underline{p}(x_i)u_{s'}(x_i) \geq \sum_{i=1}^{m} \underline{q}(x_i)u_{s}(x_i)$.

Thus, if $\lambda^* = \mu^* = 1$, (12) $= \alpha_{s'} \sum_{i=1}^{m} \overline{p}(x_i)u_{s'}(x_i) + (1-\alpha_{s'}) \sum_{i=1}^{m} \underline{p}(x_i)u_{s'}(x_i)$, if $\lambda^* \leq 1$ or $\mu^* \leq 1$, $\alpha_{s'} \sum_{i=1}^{m} \overline{p}(x_i)u_{s'}(x_i) + (1-\alpha_{s'}) \sum_{i=1}^{m} \underline{p}(x_i)u_{s'}(x_i) \geq$ (12) $\geq \alpha_{s'} \sum_{i=1}^{m} \overline{q}(x_i)u_{s'}(x_i) + (1-\alpha_{s'}) \sum_{i=1}^{m} \underline{q}(x_i)u_{s'}(x_i)$, and if $\lambda^* = \mu^* = 0$,

$$(12) = \alpha_{s'} \sum_{i=1}^{m} \overline{q}(x_i)u_{s'}(x_i) + (1-\alpha_{s'}) \sum_{i=1}^{m} \underline{q}(x_i)u_{s'}(x_i).$$

Hence, we have that

$$\begin{aligned}
&\alpha_{s'} \textstyle\sum_{i=1}^m \overline{p}(x_i)u_{s'}(x_i) + (1-\alpha_{s'}) \sum_{i=1}^m \underline{p}(x_i)u_{s'}(x_i) \geq \\
&\alpha_{s'} \textstyle\sum_{i=1}^m \overline{t}(x_i)u_{s'}(x_i) + (1-\alpha_{s'}) \sum_{i=1}^m \underline{t}(x_i)u_s(x_i) \geq \\
&\alpha_{s'} \textstyle\sum_{i=1}^m \overline{q}(x_i)u_{s'}(x_i) + (1-\alpha_{s'}) \sum_{i=1}^m \underline{q}(x_i)u_s(x_i) \\
&\Leftrightarrow P^{s'}h \succsim Co(P,Q)_s h \succsim Q_{s'}h.
\end{aligned}$$

To see that Axiom 5 is necessary, suppose that we have the representation and that $p_{s'}h \succ q_{s'}h$ for all $q \in Q$. Then $p_{s'}h \succ q_{s'}h$ for all $q \in Q$ is equivalent to

$$\sum_{i=1}^m p(x_i)u_{s'}(x_i) \geq \sum_{i=1}^m q(x_i)u_{s'}(x_i) \text{ for all } q \in Q. \tag{13}$$

First, I will show that $p_{s'}h \succ Co(p,Q)_{s'}h$ is necessary if $\alpha^{s'} < 1$. So suppose $\alpha^{s'} < 1$. Since for all $t \in Co(p,Q)$ there exists $q \in Q$ and $\lambda \in [0,1]$ such that $t = \lambda p + (1-\lambda)q$, we have that $\sum_{i=1}^m t(x_i)u_{s'}(x_i) \leq \sum_{i=1}^m p(x_i)u_{s'}(x_i)$ for all $t \in Co(p,Q)$. Also, since $Q \subseteq Co(p,Q)$, (13) $\Rightarrow$ there exists $t \in Co(p,Q)$ such that $\sum_{i=1}^m p(x_i)u_{s'}(x_i) > \sum_{i=1}^m t(x_i)u_{s'}(x_i)$ $\Rightarrow \sum_{i=1}^m p(x_i)u_{s'}(x_i) > \alpha_{s'} \sum_{i=1}^m p(x_i)u_{s'}(x_i) + (1-\alpha_{s'}) \sum_{i=1}^m \underline{t}(x_i)u_{s'}(x_i)$.

I will now show that $Co(p,Q)_{s'}h \succ Q_{s'}h$ is necessary. Note that

$$\overline{t} = argmax_{t \in Co(p,Q)} \sum_{i=1}^m t(x_i)u_{s'}(x_i) = argmax_{\lambda \in [0,1], q \in Q} \lambda \sum_{i=1}^m p(x_i)u_{s'}(x_i) + (1-\lambda) \sum_{i=1}^m q(x_i)u_{s'}(x_i).$$

Since

$$max_{\lambda \in [0,1], q \in Q} \lambda \sum_{i=1}^m p(x_i)u_{s'}(x_i) + (1-\lambda) \sum_{i=1}^m q(x_i)u_{s'}(x_i) > max_{q \in Q} \sum_{i=1}^m q(x_i)u_{s'}(x_i)$$

we now have that

$$\sum_{i=1}^m \overline{t}(x_i)u_{s'}(x_i) > \sum_{i=1}^m \overline{q}(x_i)u_{s'}(x_i). \tag{14}$$

Also, for all $t \in Co(p,Q)$, $\sum_{i=1}^m t(x_i)u_{s'}(x_i) = \sum_{i=1}^m (\lambda p(x_i) + (1-\lambda)q(x_i))u_{s'}(x_i) \geq \sum_{i=1}^m (\lambda p(x_i) + (1-\lambda)\underline{q}(x_i))u_{s'}(x_i) \geq \sum_{i=1}^m \underline{q}(x_i))u_{s'}(x_i)$ and thus $\sum_{i=1}^m \underline{t}(x_i)u_{s'}(x_i) \geq \sum_{i=1}^m \underline{q}(x_i)u_{s'}(x_i)$. Combining this with (14) we have that $\alpha_{s'} \sum_{i=1}^m \overline{t}(x_i)u_{s'}(x_i) + (1-\alpha_{s'}) \sum_{i=1}^m \underline{t}(x_i)u_{s'}(x_i) > \alpha_{s'} \sum_{i=1}^m \overline{q}(x_i)u_{s'}(x_i) + (1-\alpha_{s'}) \sum_{i=1}^m \underline{q}(x_i)u_{s'}(x_i) \Leftrightarrow Co(p,Q)^{s'}h \succ Q^{s'}h$. ■

Proof of Theorem 2: **Sufficiency of axioms:** By Axioms 1 through 5,

$$\sum_{s=1}^n \left[\alpha_s \sum_{i=1}^m \overline{h}^s(x_i)u_s(x_i) + (1-\alpha_s) \sum_{i=1}^m \underline{h}^s(x_i)u_s(x_i)\right] \geq$$

$$\sum_{s=1}^{n}\Big[\alpha_s\sum_{i=1}^{m}\overline{g}^s(x_i)u_s(x_i)+(1-\alpha_s)\sum_{i=1}^{m}\underline{g}^s(x_i)u_s(x_i)\Big]$$

By Axiom 6 there exists at least one non-null state. Let $\hat{s}$ be a non-null state. Consider $P, Q \in \mathscr{P}$. Let s' be any non-null state. By Axiom 7

$$\begin{aligned}
& \alpha_{s'}\sum_{i=1}^{m}\overline{p}(x_i)u_{s'}(x_i)+(1-\alpha_{s'})\sum_{i=1}^{m}\underline{p}(x_i)u_{s'}(x_i)\\
& >\alpha_{s'}\sum_{i=1}^{m}\overline{q}(x_i)u_{s'}(x_i)+(1-\alpha_{s'})\sum_{i=1}^{m}\underline{q}(x_i)u_{s'}(x_i)\\
\Leftrightarrow\quad & \sum_{s\neq s'}\Big[\alpha_s\sum_{i=1}^{m}\overline{h}^s(x_i)u_s(x_i)+(1-\alpha_s)\sum_{i=1}^{m}\underline{h}^s(x_i)u_s(x_i)\Big]\\
& +\alpha_{s'}\sum_{i=1}^{m}\overline{p}(x_i)u_{s'}(x_i)+(1-\alpha_{s'})\sum_{i=1}^{m}\underline{p}(x_i)u_{s'}(x_i)\\
& >\sum_{s\neq s'}\Big[\alpha_s\sum_{i=1}^{m}\overline{h}^s(x_i)u_s(x_i)+(1-\alpha_s)\sum_{i=1}^{m}\underline{h}^s(x_i)u_s(x_i)\Big]\\
& +\alpha_{s'}\sum_{i=1}^{m}\overline{q}(x_i)u_{s'}(x_i)+(1-\alpha_{s'})\sum_{i=1}^{m}\underline{q}(x_i)u_{s'}(x_i)\\
\Leftrightarrow\quad & P_{s'}h\succ Q_{s'}h\Leftrightarrow(h^1,\dots,h^{s'-1},P,h^{s'+1},\dots,h^n)\succ(h^1,\dots,h^{s'-1},Q,h^{s'+1},\dots,h^n)\\
\Leftrightarrow\quad & P_{\hat{s}}h\succ Q_{\hat{s}}h\Leftrightarrow(h^1,\dots,h^{\hat{s}-1},P,h^{\hat{s}+1},\dots,h^n)\succ(h^1,\dots,h^{\hat{s}-1},Q,h^{\hat{s}+1},\dots,h^n)\\
\Leftrightarrow\quad & \sum_{s\neq \hat{s}}\Big[\alpha_s\sum_{i=1}^{m}\overline{h}^s(x_i)u_s(x_i)+(1-\alpha_s)\sum_{i=1}^{m}\underline{h}^s(x_i)u_s(x_i)\Big]\\
& +\alpha_{\hat{s}}\sum_{i=1}^{m}\overline{p}(x_i)u_{\hat{s}}(x_i)+(1-\alpha_{\hat{s}})\sum_{i=1}^{m}\underline{p}(x_i)u_{\hat{s}}(x_i)\\
& >\sum_{s\neq \hat{s}}\Big[\alpha_s\sum_{i=1}^{m}\overline{h}^s(x_i)u_s(x_i)+(1-\alpha_s)\sum_{i=1}^{m}\underline{h}^s(x_i)u_s(x_i)\Big]\\
& +\alpha_{\hat{s}}\sum_{i=1}^{m}\overline{q}(x_i)u_{\hat{s}}(x_i)+(1-\alpha_{\hat{s}})\sum_{i=1}^{m}\underline{q}(x_i)u_{\hat{s}}(x_i)\\
\Leftrightarrow\quad & \alpha_{\hat{s}}\sum_{i=1}^{m}\overline{p}(x_i)u_{\hat{s}}(x_i)+(1-\alpha_{\hat{s}})\sum_{i=1}^{m}\underline{p}(x_i)u_{\hat{s}}(x_i)\\
& >\alpha_{\hat{s}}\sum_{i=1}^{m}\overline{q}(x_i)u_{\hat{s}}(x_i)+(1-\alpha_{\hat{s}})\sum_{i=1}^{m}\underline{q}(x_i)u_{\hat{s}}(x_i).
\end{aligned}$$

For singleton sets p and q this implies that

$$\sum_{i=1}^{m} p(x_i)u_{s'}(x_i) > \sum_{i=1}^{m} q(x_i)u_{s'}(x_i) \Leftrightarrow \sum_{i=1}^{m} p(x_i)u_{\hat{s}}(x_i) > \sum_{i=1}^{m} q(x_i)u_{\hat{s}}(x_i).$$

Since p and q are simple lotteries this means that they are evaluated according to the same von Neumann-Morgenstern expected utility function in states $\hat{s}$ and s'. Since von Neumann-Morgenstern expected utility functions are unique up to a positive linear transformation there exist scalars $a_{s'} > 0$ and $b_{s'}$ such that $u_{s'}(\cdot) = a_{s'}u_{\hat{s}}(\cdot) + b_{s'}$.

Define $u(x_i) = u_{\hat{s}}(x_i)$. Since $\hat{s}$ is non-null, we have that $u_{\hat{s}}(\cdot)$ is non-constant, since a state s is null if and only if $u_s(\cdot)$ is constant. Then for general $h, g \in \mathscr{H}$,

$$\begin{aligned} h \succ g \quad \Leftrightarrow \quad & \sum_{s=1}^{n} \Big[\alpha_s \sum_{i=1}^{m} \overline{h}^s(x_i)[a_s u(x_i) + b_s] + (1-\alpha_s) \sum_{i=1}^{m} \underline{h}^s(x_i)[a_s u(x_i) + b_s]\Big] > \\ & \sum_{s=1}^{n} \Big[\alpha_s \sum_{i=1}^{m} \overline{g}^s(x_i)[a_s u(x_i) + b_s] + (1-\alpha_s) \sum_{i=1}^{m} \underline{g}^s(x_i)[a_s u(x_i) + b_s]\Big] \\ \Leftrightarrow \quad & \sum_{s=1}^{n} \Big[a_s\alpha_s \sum_{i=1}^{m} \overline{h}^s(x_i)u(x_i) + a_s(1-\alpha_s) \sum_{i=1}^{m} \underline{h}^s(x_i)u(x_i)\Big] > \\ & \sum_{s=1}^{n} \Big[a_s\alpha_s \sum_{i=1}^{m} \overline{g}^s(x_i)u(x_i) + a_s(1-\alpha_s) \sum_{i=1}^{m} \underline{g}^s(x_i)u(x_i)\Big]. \end{aligned}$$

Suppose now, in order to reach a contradiction, that $\alpha_{s'} > \alpha_{\hat{s}}$. For $P, Q \in \mathscr{P}$ with $\sum_{i=1}^m \underline{q}(x_i)u(x_i) > \sum_{i=1}^m \underline{p}(x_i)u(x_i)$ and $\frac{1-\alpha_{\hat{s}}}{\alpha_{\hat{s}}}\Big[\sum_{i=1}^m \underline{q}(x_i)u(x_i) - \sum_{i=1}^m \underline{p}(x_i)u(x_i)\Big] \geq \sum_{i=1}^m \overline{p}(x_i)u(x_i) - \sum_{i=1}^m \overline{q}(x_i)u(x_i) > \frac{1-\alpha_{s'}}{\alpha_{s'}}\Big[\sum_{i=1}^m \underline{q}(x_i)u(x_i) - \sum_{i=1}^m \underline{p}(x_i)u(x_i)\Big]$ we have a violation of Axiom 7, because this is equivalent to

$$\alpha_{\hat{s}} \sum_{i=1}^{m} \overline{p}(x_i)u(x_i) + (1-\alpha_{\hat{s}}) \sum_{i=1}^{m} \underline{p}(x_i)u(x_i) \leq \alpha_{\hat{s}} \sum_{i=1}^{m} \overline{q}(x_i)u(x_i) + (1-\alpha_{\hat{s}}) \sum_{i=1}^{m} \underline{q}(x_i)u(x_i)$$

and $\alpha_{s'} \sum_{i=1}^m \overline{p}(x_i)u(x_i) + (1-\alpha_{s'}) \sum_{i=1}^m \underline{p}(x_i)u(x_i) > \alpha_{s'} \sum_{i=1}^m \overline{q}(x_i)u(x_i) + (1-\alpha_{s'}) \sum_{i=1}^m \underline{q}(x_i)u(x_i)$, which is again equivalent to $P^{s'}h \succ Q^{s'}h$ but $P^{\hat{s}}h \not\succ Q^{\hat{s}}h$. Thus, since Axiom 7 has to hold for all $P, Q \in \mathscr{P}$ we cannot have $\alpha_{s'} > \alpha_{\hat{s}}$.

Suppose similarly, again in order to reach a contradiction, that $\alpha_{\hat{s}} > \alpha_{s'}$. For $P, Q \in \mathscr{P}$ with $\sum_{i=1}^m \underline{q}(x_i)u(x_i) > \sum_{i=1}^m \underline{p}(x_i)u(x_i)$ and $\frac{1-\alpha_{s'}}{\alpha_{s'}}\Big[\sum_{i=1}^m \underline{q}(x_i)u(x_i) - \sum_{i=1}^m \underline{p}(x_i)u(x_i)\Big] \geq \sum_{i=1}^m \overline{p}(x_i)u(x_i) - \sum_{i=1}^m \overline{q}(x_i)u(x_i) > \frac{1-\alpha_{\hat{s}}}{\alpha_{\hat{s}}}\Big[\sum_{i=1}^m \underline{q}(x_i)u(x_i) - \sum_{i=1}^m \underline{p}(x_i)u(x_i)\Big]$ we have another violation of Axiom 7. Therefore, since Axiom 7 has to hold for all $P, Q \in \mathscr{P}$ we cannot have $\alpha_{\hat{s}} > \alpha_{s'}$.

Hence, the only possibility is $\alpha_{s'} = \alpha_{\hat{s}} \equiv \alpha$. Thus,

$$h \succ g \quad \Leftrightarrow \quad \sum_{s=1}^{n} a_s \Big[\alpha \sum_{i=1}^{m} \overline{h}^s(x_i)u(x_i) + (1-\alpha)\sum_{i=1}^{m} \underline{h}^s(x_i)u(x_i)\Big] > \sum_{s=1}^{n} a_s \Big[\alpha \sum_{i=1}^{m} \overline{g}^s(x_i)u(x_i) + (1-\alpha)\sum_{i=1}^{m} \underline{g}^s(x_i)u(x_i)\Big].$$

Define $\mu(s) = \frac{a_s}{\sum_s a_s}$. We have that $\sum_s \mu(s) = 1$, and since $a_s \geq 0$ for all s and $a_s > 0$ for some s, we also have $0 \leq \mu(s) \leq 1$. Thus μ is a probability measure over the states.

We therefore have that

$$\sum_{s=1}^{n} \mu(s) \Big[\alpha \sum_{i=1}^{m} \overline{h}^s(x_i)u(x_i) + (1-\alpha)\sum_{i=1}^{m} \underline{h}^s(x_i)u(x_i)\Big] \geq$$

$$\sum_{s=1}^{n} \mu(s) \Big[\alpha \sum_{i=1}^{m} \overline{g}^s(x_i)u(x_i) + (1-\alpha)\sum_{i=1}^{m} \underline{g}^s(x_i)u(x_i)\Big].$$

Necessity of axioms: The necessity of Axioms 1 through 5 follows from Theorem 1. To prove necessity of Axiom 6, suppose we have the representation in Theorem 2 and that $h \sim g$ for all $h, g \in \mathscr{H}$. However, then $h \sim g$ for all $h, g \in \mathscr{H}$ if and only if for all $h, g \in \mathscr{H}$,

$$\sum_{s=1}^{n} \mu(s) \Big[\alpha \sum_{i=1}^{m} \overline{h}^s(x_i)u(x_i) + (1-\alpha)\sum_{i=1}^{m} \underline{h}^s(x_i)u(x_i)\Big]$$

$$= \sum_{s=1}^{n} \mu(s) \Big[\alpha \sum_{i=1}^{m} \overline{g}^s(x_i)u(x_i) + (1-\alpha)\sum_{i=1}^{m} \underline{g}^s(x_i)u(x_i)\Big],$$

which implies that $u(\cdot)$ must be constant. Hence we must have that there exists $h, g \in \mathscr{H}$ such that $h \succ g$.

To prove the necessity of Axiom 7, suppose there exists $\hat{s}$ such that $P_{\hat{s}}h \succ Q_{\hat{s}}h$. Then

$$\mu(\hat{s}) \Big[\alpha \sum_{i=1}^{m} \overline{p}(x_i)u(x_i) + (1-\alpha)\sum_{i=1}^{m} \underline{p}(x_i)u(x_i)\Big]$$

$$+ \sum_{s \neq \hat{s}} \mu(s) \Big[\alpha \sum_{i=1}^{m} \overline{h}^s(x_i)u(x_i) + (1-\alpha)\sum_{i=1}^{m} \underline{h}^s(x_i)u(x_i)\Big]$$

$$> \mu(\hat{s}) \Big[\alpha \sum_{i=1}^{m} \overline{q}(x_i)u(x_i) + (1-\alpha)\sum_{i=1}^{m} \underline{q}(x_i)u(x_i)\Big]$$

$$+ \sum_{s \neq \hat{s}} \mu(s) \Big[\alpha \sum_{i=1}^{m} \overline{h}^s(x_i)u(x_i) + (1-\alpha)\sum_{i=1}^{m} \underline{h}^s(x_i)u(x_i)\Big]$$

$$\Leftrightarrow \quad \alpha \sum_{i=1}^{m} \overline{p}(x_i)u(x_i) + (1-\alpha)\sum_{i=1}^{m} \underline{p}(x_i)u(x_i)$$
$$> \alpha \sum_{i=1}^{m} \overline{q}(x_i)u(x_i) + (1-\alpha)\sum_{i=1}^{m} \underline{q}(x_i)u(x_i)$$
$$\Rightarrow \quad \mu(s')\Big[\alpha \sum_{i=1}^{m} \overline{p}(x_i)u(x_i) + (1-\alpha)\sum_{i=1}^{m} \underline{p}(x_i)u(x_i)\Big]$$
$$+ \sum_{s\neq s'} \mu(s)\Big[\alpha \sum_{i=1}^{m} \overline{h}^s(x_i)u(x_i) + (1-\alpha)\sum_{i=1}^{m} \underline{h}^s(x_i)u(x_i)\Big]$$
$$> \mu(s')\Big[\alpha \sum_{i=1}^{m} \overline{q}(x_i)u(x_i) + (1-\alpha)\sum_{i=1}^{m} \underline{q}(x_i)u(x_i)\Big]$$
$$+ \sum_{s\neq s'} \mu(s)\Big[\alpha \sum_{i=1}^{m} \overline{h}^s(x_i)u(x_i) + (1-\alpha)\sum_{i=1}^{m} \underline{h}^s(x_i)u(x_i)\Big].$$

This holds for all non-null s', and thus $P_{s'}h \succ Q_{s'}h$ for all non-null s'.■

Proof of Theorem 3: <u>**Sufficiency of axioms:**</u> The proof follows that of Theorem 2 until the point where I establish that for general $h, g \in \mathscr{H}$,

$$h \succ g \quad \Leftrightarrow \quad \sum_{s=1}^{n} \Big[a_s\alpha_s \sum_{i=1}^{m} \overline{h}^s(x_i)u(x_i) + a_s(1-\alpha_s)\sum_{i=1}^{m} \underline{h}^s(x_i)u(x_i)\Big] >$$
$$\sum_{s=1}^{n} \Big[a_s\alpha_s \sum_{i=1}^{m} \overline{g}^s(x_i)u(x_i) + a_s(1-\alpha_s)\sum_{i=1}^{m} \underline{g}^s(x_i)u(x_i)\Big].$$

In the proof of Theorem 2 we get a violation of Axiom 7 when we assume $\alpha_{s'} > \alpha_{\hat{s}}$ for $P, Q \in \mathscr{P}$ with $\sum_{i=1}^{m} \underline{q}(x_i)u(x_i) > \sum_{i=1}^{m} \underline{p}(x_i)u(x_i)$ and $\frac{1-\alpha_{\hat{s}}}{\alpha_{\hat{s}}}\Big[\sum_{i=1}^{m} \underline{q}(x_i)u(x_i) - \sum_{i=1}^{m} \underline{p}(x_i)u(x_i)\Big] \geq \sum_{i=1}^{m} \overline{p}(x_i)u(x_i) - \sum_{i=1}^{m} \overline{q}(x_i)u(x_i) >$ $\frac{1-\alpha_{s'}}{\alpha_{s'}}\Big[\sum_{i=1}^{m} \underline{q}(x_i)u(x_i) - \sum_{i=1}^{m} \underline{p}(x_i)u(x_i)\Big]$. Since Axiom 7' only applies to singleton lotteries p and q, we do not get this violation here. Similarly, we do not get a contradiction by assuming $\alpha_{s'} < \alpha_{\hat{s}}$.

Define $\mu(s) = \frac{a_s}{\sum_s a_s}$. We have that $\sum_s \mu(s) = 1$, and since $a_s \geq 0$ for all s and $a_s > 0$ for some s, we also have $0 \leq \mu(s) \leq 1$. Thus μ is a probability measure over the states.

We therefore have that

$$\sum_{s=1}^{n} \mu(s)\Big[\alpha_s \sum_{i=1}^{m} \overline{h}^s(x_i)u(x_i) + (1-\alpha_s)\sum_{i=1}^{m} \underline{h}^s(x_i)u(x_i)\Big] \geq$$
$$\sum_{s=1}^{n} \mu(s)\Big[\alpha_s \sum_{i=1}^{m} \overline{g}^s(x_i)u(x_i) + (1-\alpha_s)\sum_{i=1}^{m} \underline{g}^s(x_i)u(x_i)\Big].$$

Necessity of axioms: The proof of necessity of the axioms follows that of Theorem 2 with $\overline{p}(x_i) = \underline{p}(x_i) = p(x_i)$ and $\overline{q}(x_i) = \underline{q}(x_i) = q(x_i)$.

References

[1] Ahn, D. S. (2008): "Ambiguity without a State Space," *Review of Economic Studies*, 75, 3-28.

[2] Anscombe, F.J. and R. J. Aumann (1963): "A definition of subjective probability," *Annals of Mathematical Statistics*, 34, 199-205.

[3] Bewley, T.F. (1986): "Knightian decision theory: Part I," Cowles Foundation Discussion Paper no. 807, Yale University.

[4] Dekel, E., B. L. Lipman, and A. Rustichini (2001): "Representing preferences with a unique subjective state space," *Econometrica* 69, 891-934.

[5] Ellsberg (1961): "Risk, ambiguity, and the Savage axioms," *Quarterly Journal of Economics* 75, 643-669.

[6] Epstein, L. G. , M. Marinacci, and K. Seo (2007): "Coarse contingencies and ambiguity," *Theoretical Economics*, 2, 355-394.

[7] Epstein, L. G. and M. Schneider (2003): "Recursive multiple-priors," *Journal of Economic Theory* 113, 1-31.

[8] Gajdos, T., T. Hayashi, J.-M. Tallon, and J.-C. Vergnaud (forthcoming): "Attitude toward imprecise information," forthcoming, *Journal of Economic Theory.*

[9] Ghirardato, P. (2001): "Coping with ignorance: unforeseen contingencies and non-additive uncertainty," *Economic Theory* 17, 247-276.

[10] Ghirardato, P., Maccheroni, F., and Marinacci, M. (2004): "Differentiating ambiguity and ambiguity attitude," *Journal of Economic Theory* 118, 133-173.

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

[12] Gul, F. and W. Pesendorfer (2001): "Temptation and self-control," *Econometrica* 69, 1403-1435.

[13] Holmström, Bengt (1979): "Moral hazard and observability," *Bell Journal of Economics* 10, 74-91.

[14] Klibanoff, P., M. Marinacci, and S. Mukerji (2005): "A smooth model of decision making under ambiguity," *Econometrica* 73, 1849-1892.

[15] Kreps, D. M. (1979): "A representation theorem for preference for flexibility," *Econometrica* 47, 565-577.

[16] Kreps, D. M. (1988): "Notes on the theory of choice," *Westview Press*, Boulder.

[17] Kreps, D. M. (1992): "Static choice in the presence of unforeseen contingenies," in *Economic Analysis of Markets and Games*, ed. by P. Dasgupta et al., MIT Press, 258-281.

[18] Manski, C. F. and E. Tamer (2002): "Inference on regressions with interval data on a regressor or outcome," *Econometrica* 70, 519-546.

[19] Mas-Colell, A., M. D. Whinston, and J. R. Green (1995): "Microeconomic theory," *Oxford University Press*, New York.

[20] Mukerji, S. (1997): "Understanding the nonadditive probability decision model," *Economic Theory* 9, 23-46.

[21] Olszewski, W. (2007): "Preferences over sets of lotteries," *Review of Economic Studies* 74, 567-595.

[22] Ozdenoren, E. (2002): "Completing the state space with subjective states," *Journal of Economic Theory* 105, 531-539.

[23] Nehring, K. (1999): "Preference for flexibility in a Savage framework," *Econometrica* 67, 101-119.

[24] Savage, L. J. (1954): "The foundations of statistics," *Dover Publications*, New York.

[25] Vierø, M.-L. (2007): "Contracting in vague environments," Working Paper, Queen's University.