

## CHOICE WITH AMBIGUITY AS SETS OF PROBABILITIES

MAXWELL B. STINCHCOMBE

ABSTRACT. An outcome is ambiguous if it is an incomplete description of the probability distribution over consequences. An incomplete description is identified with the set of probabilities that satisfy the incomplete description. A choice problem is ambiguous if choices lead to sets of probabilities. This paper develops the theory of ambiguous choice problems as a continuous, linear extension of expected utility preferences from probabilities to sets of probabilities. The axiomatic foundation and representation theorem lead to a demand theory for ambiguity reduction, to an analysis of the value of information in the presence of ambiguity, and to a theory of efficient allocations of risk and ambiguity.

Roughly, risk refers to situations where the likelihood of relevant events can be represented by a probability measure, while ambiguity refers to situations where there is insufficient information available for the decision maker to assign probabilities to events. (Epstein and Zhang [6])

### 1. INTRODUCTION

This paper takes this rough distinction as the defining difference between risky choice problems and ambiguous choice problems. “Uncertainty” is an umbrella term for situations in which actions do not lead deterministically to outcomes. Risky situations are uncertain situations in which the decision maker knows the probability distributions associated with their various choices. Ambiguous situations are uncertain situations in which the decision maker has only partial information about the probability distributions associated with their various choices. Under study are choice problems and games with ambiguous outcomes.

Without ambiguity, expected utility preferences can be represented by continuous linear functions on probabilities. Such functions have continuous, linear extensions from probabilities to sets of probabilities. Except for the difference

---

*Date:* First version: February 12, 2003. This version: November 20, 2003

*Key words and phrases.* Ambiguity, decision theory, valuing information.

in domain, the axiomatic foundations of these continuous affine preferences are indistinguishable from those of expected utility theory.

There is a direct sum decomposition of the domain. This decomposition leads to a complete separation of attitudes toward risk and attitudes toward ambiguity. Such a separation is exceedingly difficult to obtain when working with a Savage model of choice under uncertainty. The essential distinction is that the present approach puts ambiguity into the description of the problem. The Savage modeling approach tries to extract subjective ambiguity from properties of the preferences.

The dual space representation of the preferences lead to a demand theory for ambiguity reduction, to a general theory of ambiguous equilibria in finite games, and to an analysis of the value of information in the presence of ambiguity. These applications are only sketched. There is a fuller treatment of Knight's [10] treatment of extra-economic, or "true," profits arising from ambiguity.

**1.1. Knight's extra-economic profits.** There is at least a century's worth of social science commentary on the critical role of new goods and production techniques in capitalist economies.<sup>1</sup> Definitionally, a **new** good or production process is one for which the distribution of the rewards associated is not and can not be known. This work is an attempt to trace through Knight's arguments about how ambiguity leads to extra-economic profits, profits above and beyond those that show up in competitive equilibria with risk. Knight presents a preference based argument and a role based argument.

1.1.1. *Preferences of innovators and efficient contracts.* Schumpeter (1942, p. 132) briefly, and Knight (1921 (1971) Part III) much more extensively discussed the psychological makeup of people who introduce new products and new methods of production, people who make decisions not knowing the distribution of the consequences.

---

<sup>1</sup>See Marx on the role of positive feedback cycles in expansion (1864 (1906) Vol. I, Ch. XXIV, esp. p. 663 *et seq.*), Weber's (1930) discussion of the congruence between Puritan ideals and the accumulation and expansion of capital, Schumpeter (1942, Ch. 12) on the entrepreneurial function in expansion, and especially, Knight (1921 (1971) Ch. 11) on the role of uncertainty in social progress.

My model-based version of this argument:  $X$  is a random reward to be efficiently split between two agents, one less ambiguity averse than the other. The distribution of  $X$  belongs to some set. This is an ambiguous problem if the set contains more than one point. An allocation is a mapping  $x \mapsto (f_1(x), f_2(x))$ . In principle,  $f_i(x)$  can be a distribution over rewards. The question is “What are the efficient allocations?”

In general, the efficient allocations are extreme points of the set of allocations. They reflect tradeoffs between attitudes towards risk and ambiguity. In special cases, arguably the ones that Knight had in mind, one achieves Knight’s result that if  $i$  is more ambiguity averse than  $j$ , then  $f_i(x) \equiv c$  is efficient. In other words, those most temperamentally suited to dealing with ambiguity bear all of it. They become the residual claimants, the entrepreneurs.

There are two comments. First, the extra-economic profit is the premium, above and beyond the risk premium, that the residual claimant receives for absorbing the ambiguity. Second, this is a preference based argument for a particular pattern. Self-selection into roles is implicit.

1.1.2. *Roles and efficient contracts.* Knight also gave a role-based argument for who will be the residual claimant. If there is more ambiguity attached to the choices of one agent, then one expects that agent to be the residual claimant. In modern terms, I believe that Knight’s argument is about incentives to exercise the effort to control the ambiguity, and about the other party to the contract to be unwilling to be subject to ambiguity as well as risk due to moral hazard problems.

My model-based version of this argument: Two players much choose their actions/effort levels simultaneously. The choices of one agent have larger effects on the ambiguity of the outcome. This structure defines the roles of the two players.

In general, tradeoffs between incentive effects and ambiguity effects are more complicated than the tradeoffs between incentive effects, and those are already very complicated. One can, however, see the patterns that lead to (something like) the residual claimant being the one whose actions have a large effect on the ambiguity of the outcome.

**1.2. Equilibrium with ambiguity.** A Nash equilibrium without ambiguity is a vector of beliefs which, when held by all players, belongs, componentwise, to the set of best responses to the beliefs. An equilibrium with ambiguity is a vector of sets of beliefs which, when held by all players, is a subset of, componentwise, the set of best responses to the set of beliefs.

One idea of equilibrium is as a point where “things have settled down.” The present view of ambiguity is as a description of incomplete knowledge about the probabilities over consequences associated with different actions. There is a tension between “incomplete knowledge” and having “settled down.” This tension appears in the interpretations of the ambiguous equilibria of several of the two person, simultaneous move games examined here.

**1.3. A thumbnail sketch of some intellectual history.** Thumbnail sketches of intellectual history are much like theory models, organizing a huge amount of material in a way that highlights the major themes and influences. Both ignore complexities. With that in mind, there are three organizing paradoxes in the intellectual history of choice under risk and uncertainty, the St. Petersburg paradox, the Allais paradox, and the Ellsberg paradox.

**1.3.1. *St. Petersburg.*** The St. Petersburg paradox is an example of bounded willingness to pay for a random variable,  $X \geq 0$  with unbounded expectation,  $E X = \infty$ . This implies that expectations are, in many cases, not the right tool for choosing between random outcomes. Bernoulli’s expected utility resolution is to choose between random variables on the basis of  $U(X) := E u(X)$  for a non-linear  $u(\cdot)$ . This class of preferences are well adapted to needs in the foundations of game theory and Bayesian statistics, von Neumann and Morgenstern [15] and Savage [16].

**1.3.2. *Allais.*** Expected utility theory preferences are continuous linear functions on the set of probabilities. Allais’s paradox [1] is an example of indifference curves not being parallel in the set of probabilities. This implies that linearity is, in many cases, not the right tool for choosing between distributions. Identify random variables,  $X$ , with their distributions,  $P = P_X$ . Allais [1, p. 513] suggests studying preferences that can be represented by non-linear functions of  $P$ . Machina [12] was the first to show that this non-expected utility resolution of the paradox

delivers a workable theory of choice under uncertainty. Crawford [3] showed that it can be adapted to the needs of game theory.

1.3.3. *Ellsberg*. The Ellsberg paradox is an example of people strictly preferring to know the distribution of random variables. This implies that preferences, linear or non-linear, over probability distributions is not the right tool for choosing in all uncertain situations.

1.4. **Map**. The next two sections develop the theory of preferences over compact convex sets of probability distributions over two consequences and then over an arbitrary finite number of consequences. The subsequent section applies this theory to Knight's theory of extra-economic profits that accrue to entrepreneurs. Following this is the theory of Nash equilibria with ambiguity. A number of two person games illustrate the tension between equilibrium and ambiguity. Following this are sketches of: "irrational" reactions to some kinds of small probability events; portfolio choice theory; the demand for ambiguity reduction.

## 2. AMBIGUOUS CHOICE WITH SIMPLE URNS

A decision problem is a triple  $(A, C, K)$  where  $A$  is a (usually finite) set of actions,  $C$  is a finite set of consequences, and  $K$  is a mapping from  $A$  to  $\mathbb{K}(\Delta(C))$ , the non-empty, compact subsets of  $\Delta(C)$ . The decision rule under study is

$$a^*(K, U) = \operatorname{argmax}\{U(K(a)) : a \in A\},$$

where  $U$  is a continuous, affine function. There is an informative decomposition of  $U(\cdot)$ . It is most easily seen in the case of just two outcomes.

2.1. **Ellsberg urns as intervals of probabilities**. The best-known ambiguous choice problem is Ellsberg's paradox. It is an example of people strictly preferring to know the distribution of random variables. If we believe that this is a real phenomenon, then preferences, linear or non-linear, over probability distributions, are not the right tool for choosing in all uncertain situations.

An urn is known to contain 90 balls, 30 of which are Red, each of the remaining 60 can be either Green or Blue. The decision maker is faced with the urn, the description just given, and two pairs of choice situations.

1. Single ticket choices:

(a) The choice between the Red and the Green ticket.

- (b) The choice between the Red and the Blue ticket.
- 2. Pairs of ticket choices:
  - (a) The choice of the R&B or the G&B pair.
  - (b) The choice of the R&G or the B&G pair.

In each situation, after the DM makes her choice, one of the 90 balls will be picked at random. If the ball's color matches the color of (one of) the chosen ticket(s), the decision maker gets \$1,000, otherwise they get nothing.

Typical preferences are

$$R \succ G \text{ and } R \succ B,$$

$$R\&B \prec G\&B \text{ and } R\&G \prec B\&G.$$

There is no possible Bayesian explanation for these preferences. If there was, we'd have both

$$P(R) > P(G) \text{ and } P(R) > P(B),$$

as well as

$$P(R) + P(B) < P(G) + P(B), \quad P(R) + P(G) < P(B) + P(G).$$

Note that this argument encompasses *any* story about beliefs about the distribution of the numbers of Blues and Greens.<sup>2</sup>

The probability that the Red ticket wins is  $\frac{1}{3}$ . That is, the action "choose Red" is risky, with the associated probability  $\frac{1}{3}$ . The actions "choose Blue" and "choose Green" are ambiguous, leading to the interval of probabilities  $[0, \frac{2}{3}]$ . Choosing the Blue&Green pair is risky,  $\frac{2}{3}$ , choosing the other two pairs is ambiguous,  $[\frac{1}{3}, 1]$ . As noted, the typical preferences are

$$\{\frac{1}{3}\} \succ [0, \frac{2}{3}] \quad \text{and} \quad \{\frac{2}{3}\} \succ [\frac{1}{3}, 1].$$

People prefer knowing the center of the interval to the interval itself.

**2.2. Preferences on sets of probabilities.** Under study are continuous, linear preferences,  $\succeq$ , on,  $\mathbb{K} = \mathbb{K}([0, 1])$ , the closed subsets of  $[0, 1]$ . For the moment, restrict attention to convex subsets, that is, to intervals  $[a, b]$ . Every interval

---

<sup>2</sup>It does not encompass stories about people believing that experimentalists (or a vengeful god) cheating by changing the number of balls in the urn after the ticket is chosen.

$[a, b]$  is of the form  $[c - r, c + r]$  for  $c = (a + b)/2$ ,  $r \geq 0$ . Adding intervals and multiplying them by non-negative constants involves adding the centers and the radii and multiplying them by constants.

The singleton subsets of  $[0, 1]$  will be denoted either  $[p, p]$  or  $\{p\}$ . Non-trivial, continuous, affine preference can be represented by a function  $U : \mathbb{K} \rightarrow \mathbb{R}$  satisfying

1. (continuity)  $c^n \rightarrow c$  and  $r^n \rightarrow r$ , implies  $U([c^n - r^n, c^n + r^n]) \rightarrow U([c - r, c + r])$ ,
2. (affineness)  $U(\alpha[c - r, c + r] + (1 - \alpha)[c' - r', c' + r']) = \alpha U([c - r, c + r]) + (1 - \alpha)U([c' - r', c' + r'])$ , and
3. (non-triviality, normalized)  $U(\{0\}) = 0$  and  $U(\{1\}) = 1$ .

**Lemma 2.1.** *The unique normalized representation of continuous, affine preferences is  $U([c - r, c + r]) = c - vr$ .*

Given the normalization  $u(\$0) = 0$  and  $u(\$1,000) = 1$ , the parameter  $v$  measures the tradeoff between ambiguity and risk, and characterizes the preferences. Ambiguity aversion corresponds  $v > 0$ , a liking for ambiguity corresponds to  $v < 0$ , and ambiguity neutrality corresponds to  $v = 0$ . The typical Ellsberg choices correspond to  $v > 0$ , but give no information about the size of the tradeoff between ambiguity and risk.

**2.3. Estimating ambiguity aversion.** Eliciting the  $v$ 's that characterize different decision makers can be done by offering choices between ambiguous choices and risky choices. Introduce a second urn with  $p$  of the balls known to be Yellow, the rest being White. The decision maker is faced with the description of the first urn given above, the description of the second urn just given, and the following choice situations.

1. The choice between the Green (Blue) or the Yellow ticket.
2. The choice between the Red-Green (Red-Blue) pair of tickets, or the White ticket.

If the DM chooses a ticket, a ball is drawn at random from the corresponding urn, and if the color of the ball matches the color of (one of) the chosen ticket(s), the DM gets \$1,000, otherwise they get nothing.

The  $p$  that makes the Green (Blue) and the Yellow ticket indifferent satisfies  $p = \frac{1}{3} - v \cdot \frac{1}{3}$  so that  $v = (1 - 3p)$ . Positive  $v$ 's correspond to  $p < \frac{1}{3}$ . The  $p'$  that

makes the Red-Green (Red-Blue) pair and the White ticket indifferent satisfies  $(1 - p') = \frac{2}{3} - v \cdot \frac{1}{3}$  so that  $v = 3p' - 1$ . In this case, positive  $v$ 's correspond to  $p' > \frac{1}{3}$ , that is,  $1 - p' < \frac{2}{3}$ .

Insert Diagram here: for  $p \in [0, \frac{1}{3}(1+v))$ ,  $W \succ RG \sim RB$ , for  $p \in (\frac{1}{3}(1+v), 1]$ ,  $RG \sim RB \succ W$ , for  $p \in [0, \frac{1}{3}(1-v))$ ,  $Y \succ G \sim B$ , and for  $p \in (\frac{1}{3}(1-v), 1]$ ,  $G \sim B \succ Y$ . For  $v$  positive, this gives the left-to-pattern  $[G \sim B \succ Y] \cap [W \succ RG \sim RB]$ , then  $[Y \succ G \sim B] \cap [W \succ RG \sim RB]$ , then  $[Y \succ G \sim B] \cap [RG \sim RB \succ W]$ . The change point/indifference point should be symmetric about  $\frac{1}{3}$  in either case.

Inspection shows that  $-1 \leq v \leq +1$  is necessary and sufficient to keep the  $p$  delivering indifference in the interval  $[0, 1]$ . This is not the only reason for such a restriction on  $v$ .

**2.4. Bounds on the risk-ambiguity tradeoff.** An absolute pessimist evaluates a set  $A$  of probabilities by focusing on the worst element of  $A$ ,  $U_{pess}(A) = \underline{u}(A) := \min\{u(x) : x \in A\}$ . An absolute optimist evaluates a set  $A$  of probabilities by focusing on the best element of  $A$ ,  $U_{opt}(A) = \bar{u}(A) := \max\{u(x) : x \in A\}$ . In the normalized case,  $U_{pess}$  corresponds to  $v = 1$  and  $U_{opt}$  to  $v = -1$ . The risk-ambiguity tradeoff is **balanced** if for all  $A \in \mathbb{K}$ ,  $\underline{u}(A) \leq U(A) \leq \bar{u}(A)$ . Unbalanced preferences correspond to being willing to pay to avoid better outcomes or being willing to pay to add worse outcomes.

For  $A, B \in \mathbb{K}$ , define  $A \vee B$  as the convex hull of the union of  $A$  and  $B$ ,  $A \vee B = \mathbf{co}(A \cup B)$ . Preferences satisfy **betweenness** if for all  $A, B \in \mathbb{K}$ ,  $[U(A) > U(B)] \Rightarrow U(A) \geq U(A \vee B) \geq U(B)$ .

Let  $w$  be the worst outcome and  $b$  the best outcome. The **risk equivalent** of  $A \in \mathbb{K}$  is that number  $p_A \in \mathbb{R}$  that satisfies  $p_A U(\{b\}) + (1 - p_A)U(\{w\}) = U(A)$ . The risk equivalent is **allowable for**  $A$  if  $p_A \in [0, 1]$ . The risk-ambiguity tradeoff is **allowable** if  $p_A$  is allowable for all  $A \in \mathbb{K}$ . If the risk-ambiguity tradeoff is allowable, then in the normalized case,  $p_A = U(A)$ .

**Theorem 2.2.** *In the normalized two outcome case, the following are equivalent:*

1. *the risk-ambiguity tradeoff is allowable,*
2. *the risk-ambiguity tradeoff is balanced,*
3. *preferences satisfy betweenness, and*
4.  $-1 \leq v \leq 1$ .

**Proof:** The equivalence of allowable risk equivalents and  $-1 \leq v \leq 1$  was established above.

In the normalized two outcome case,  $U([a, b]) = \frac{a+b}{2} - v\frac{b-a}{2}$ , and a balanced risk ambiguity tradeoff corresponds to the two inequalities,

$$(1) \quad \underline{u}([a, b]) = a \leq U([a, b]) = \frac{a+b}{2} - v\frac{b-a}{2} \leq b = \bar{u}([a, b]).$$

Rearrangement shows that  $-1 \leq v \leq +1$  iff for all  $0 \leq a \leq b \leq 1$ , (1) holds.

Suppose that  $-1 \leq v \leq +1$  and  $U([a, b]) > U([c, d])$ . Betweenness is

$$(2) \quad U([a, b]) \geq U([a, b] \vee [c, d]) \geq U([c, d]).$$

If  $[a, b] \subset [c, d]$  or  $[c, d] \subset [a, b]$ , both inequalities in (2) are satisfied, one of them as an equality. The remaining two cases are  $a < c$  and  $b < d$ , which is ruled out by  $U([a, b]) > U([c, d])$  and  $|v| \leq 1$ , and  $a > c$  and  $b > d$ . In this case,  $[a, b] \vee [c, d] = [c, b]$ . The first inequality in (2) is  $(a+b) - v(b-a) \geq (c+b) - v(b-c)$ , which reduces to  $a(1+v) \geq c(1+v)$ , satisfied because  $a > c$  and  $|v| \leq 1$ . The second inequality in (2) is  $(c+b) - v(b-c) \geq (c+d) - v(d-c)$ , which reduces to  $b(1-v) \geq d(1-v)$ , satisfied because  $b > d$  and  $|v| \leq 1$ .

Finally, suppose that  $|v| > 1$ . To show that betweenness is violated, pick  $0 < a < b \leq 1$  and  $s > 0$  so that  $[c, d] := [a-2s, b-2s] \subset [0, 1]$  and  $[a, b] \vee [c, d] = [c, b]$ . Because the radius of  $[a, b]$  is the same as the radius of  $[c, d]$ ,  $U([a, b]) > U([c, d])$ . The center of  $[c, b]$  is  $s$  less than the center of  $[a, b]$  and  $s$  greater than the center of  $[c, d]$ , while the radius of  $[c, b]$  is  $s$  greater than the common radius of  $[a, b]$  and  $[c, d]$ . With  $|v| > 1$ , this increase in radius outweighs the change in center. ■

**2.5. Removing Convexity.** If  $A$  is a closed subset of  $[0, 1]$ , and  $\alpha_k$  is a convex set of weights, then  $\sum_k \alpha_k A$  is approximately convex if  $\max_k \alpha_k$  is small. Formally, the Starr-Shapley-Folkman theorem gives

**Lemma 2.3.** *If  $U$  is a continuous affine mapping on the compact subsets of  $[0, 1]$ , then for all compact  $A$ ,  $U(A) = U(\text{conv}(A))$ .*

There is no loss in restricting attention to convex sets.

### 3. FOUNDATIONS AND REPRESENTATIONS

The finite space of consequences is  $\mathfrak{C}$ , and  $\Delta = \Delta(\mathfrak{C})$  is the set of probabilities on  $\mathfrak{C}$ .  $\mathbb{K} = \mathbb{K}(\Delta)$  is the set of non-empty, compact, convex subsets of  $\Delta$ . The Hausdorff metric,  $d_H$ , makes  $\mathbb{K}$  into a compact metric space. Under study are continuous, affine preferences on  $\mathbb{K}$ .

An extension of compound lottery logic interprets  $\alpha A + (1 - \alpha)B$  for  $A, B \in \mathbb{K}$ . This and a direct re-write of the independence axiom of von Neumann-Morgenstern expected utility theory delivers continuous, affine preferences. The dual space of  $\mathbb{K}$  delivers a representation theorem for these preferences, and a direct sum decomposition of  $\mathbb{K}$  yields more information on the shape of these preferences.

**3.1. Notation.** For  $A, B \subset \mathbb{R}^X$ ,  $A + B := \{a + b : a \in A, b \in B\}$  and  $\lambda A := \{\lambda a : a \in A\}$  for  $\lambda \geq 0$ . For  $a \in \mathbb{R}^X$ ,  $\|a\| = \sqrt{a \cdot a}$ , and this norm induces the usual (Euclidean) metric on  $\mathbb{R}^X$ . The open unit ball is  $\mathbb{U} = \{a \in \mathbb{R}^X : \|a\| < 1\}$ . Its boundary is  $\partial\mathbb{U} = \{a \in \mathbb{R}^X : \|a\| = 1\}$ . For  $a, b \in \mathbb{R}^X$ ,  $[a, b]$  is the line segment joining  $a$  and  $b$ ,  $[a, b] = \{\alpha a + (1 - \alpha)b : \alpha \in [0, 1]\}$ .

For compact  $A \subset \mathbb{R}^X$  and  $\epsilon > 0$ ,  $A^\epsilon := A + \epsilon\mathbb{U}$ . The Hausdorff metric is  $d_H(A, B) := \inf\{\epsilon > 0 : A \subset B^\epsilon, B \subset A^\epsilon\}$ . It is well-known that  $(\mathbb{K}, d_H)$  is a compact metric space.

For  $A, B \in \mathbb{K}$  and  $\alpha \in [0, 1]$ ,  $A\alpha B := \alpha A + (1 - \alpha)B \in \mathbb{K}$ . A function  $U : \mathbb{K} \rightarrow \mathbb{R}$  is **affine** if  $U(A\alpha B) = \alpha U(A) + (1 - \alpha)U(B)$ .

A **rigid motion** is an isometry of  $\mathbb{R}^X$  onto itself, it is a **rotation** if it is an isometry fixing the origin. Every rigid motion is a composition of a translation and a rotation.

Convex sets  $U$  and  $V$  form a **direct sum decomposition of a convex  $K$** , written  $K \subset U \oplus V$ , if for every  $k \in K$ , there is a unique pair  $u \in U, v \in V$ , such that  $k = u + v$ . Any affine function on  $K \subset U \oplus V$  can be decomposed into the sum of an affine function on  $U$  and a linear function on  $V$ .

**3.2. Compound lotteries.** If  $A$  and  $B$  are singleton sets, then  $A\alpha B$  has a compound lottery interpretation as a initial lottery over lotteries. The initial lottery picks the lottery  $A$  with probability  $\alpha$  and picks the lottery  $B$  with probability  $(1 - \alpha)$ . There is a similar interpretation for convex combinations of ambiguous outcomes.

Consider a two outcome example: ambiguous Situation  $A$  corresponds to being told (only) that the probability of the bad outcome is at least twice as large as the good outcome; in ambiguous Situation  $B$ , the role of the good and bad outcome are reversed. In terms of intervals,  $A = [0, \frac{1}{3}]$  and  $B = [\frac{2}{3}, 1]$ .

If the DM is told that the situation that obtains will be determined by the toss of a fair coin, the corresponding ambiguous situation is  $A\frac{1}{2}B = [\frac{1}{3}, \frac{2}{3}]$ . The left (lower) end of the interval is the least possible probability assignable to the good outcome,  $\frac{1}{3} = \frac{1}{2}0 + \frac{1}{2}\frac{2}{3}$ . The right end of the interval is the highest possible probability assignable to the good outcome,  $\frac{2}{3} = \frac{1}{2}\frac{1}{3} + \frac{1}{2}1$ .

**3.3. The extension of von Neumann-Morgenstern preferences.** If a preference relation,  $\succeq$ , on  $\Delta$  satisfies continuity and independence, then it can be represented as a continuous linear function on  $\Delta$ . The same is true for preferences on  $\mathbb{K}$ . A preference relation on  $\mathbb{K}$  is **continuous** if for all  $A \in \mathbb{K}$ , the sets  $\{B \in \mathbb{K} : A \succeq B\}$  and  $\{B \in \mathbb{K} : B \succeq A\}$  are closed.

**Axiom 1** (Independence). *For all  $A, B, C \in \mathbb{K}$  and all  $\alpha \in (0, 1)$ ,  $A \succeq B$  if and only if  $A\alpha C \succeq B\alpha C$ .*

When  $A$ ,  $B$ , and  $C$  are singleton sets, this is the usual independence axiom, and represents an assumption about compound lotteries. When  $A$  and  $B$  are sets,  $A\alpha B$  is a combination of an objective lottery, represented by  $\alpha$ , and the ambiguous lotteries, represented by  $A$  and  $B$ . The assumption is that the order of resolution of the types of uncertainty does not matter.

**Theorem 3.1.** *Preferences on  $\mathbb{K}$  are continuous and satisfy independence if and only if they can be represented by a continuous affine function.*

**Proof:** Standard. ■

The function  $p \mapsto u(p) := U(\{p\})$ , being continuous and affine, is a von Neumann-Morgenstern utility function on  $\Delta$ . The function  $U$  is an extension of  $u$  from the singleton subsets in  $\mathbb{K}$  to all of  $\mathbb{K}$ .

### 3.4. A dual space representation. **WARNING: THIS SUBSECTION IS STILL BEING WORKED OUT**

The dual space representation of  $U$  can be given in three steps. (1) The use of support functions will show that  $\mathbb{K}$  is linearly homeomorphic to a compact, convex subset of the almost everywhere continuously differentiable functions on the compact subset of  $\partial\mathbb{U}$ . (2) The span of the set of support functions is dense in this space of continuous functions. (3) The Riesz representation theorem delivers a dense subset of the dual space as the set of signed measures.

Let  $H_\Delta$  be the smallest hyperplane in  $\mathbb{R}^X$  containing  $\Delta$ , and let  $H = H_\Delta - p$  be the translate of  $H_\Delta$  by some  $p$  in the relative interior of  $\Delta$ . The origin is in the relative interior of the translate of  $\Delta$ , itself a subset of  $H$ . Therefore, every compact convex subset of  $H$  is of the form  $\lambda A$  for some  $A \in \mathbb{K}$  and  $\lambda \geq 0$ .

Steps (1) and (2) use  $S^1(\partial\mathbb{U}_H)$ , the set of almost everywhere continuously differentiable functions on  $\partial\mathbb{U} \cap H$  with the Sobolev norm. The support function,  $s_A$ , of a non-empty, compact, convex  $A \subset H$  is defined by  $s_A(u) = \max\{u \cdot a : a \in A\}$  for  $u \in \partial\mathbb{U} \cap H$ . The mapping  $A \mapsto s_A$  is continuous, and satisfies  $(\alpha A + \beta B) \mapsto \alpha s_A + \beta s_B$  for  $\alpha, \beta \geq 0$ . Further, the span of  $\{s_A : A \in \mathbb{K}\}$  is dense in  $C(\partial\mathbb{U}_H)$  (e.g. from [17, Lemma 1.7.9, p. 45], or more directly from [7]). The Reisz representation theorem and the denseness directly imply

**Theorem 3.2.** *A dense set of the continuous preferences  $\succeq$  on  $\mathbb{K}$  satisfying independence can be represented by a function of the form*

$$U(A) = \int_{\partial\mathbb{U} \cap H} \max\{u \cdot a : a \in A\} dm_\succeq(u)$$

for some signed, countably additive, Borel measure  $m_\succeq$  on  $\partial\mathbb{U} \cap H$ .

In the two-outcome case, this theorem is particularly easy to visualize.  $H_\Delta$  is  $\{(r, 1-r) : r \in \mathbb{R}\}$ , translating by (say)  $p = (\frac{1}{2}, \frac{1}{2})$  means that  $\Delta$  is represented by the line segment  $[(-\frac{1}{2}, +\frac{1}{2}), (+\frac{1}{2}, -\frac{1}{2})]$ . A signed measure on  $H \cap \partial\mathbb{U}$  puts mass,  $v_a$  and  $v_b$  on the two points,  $x_a = (-\sqrt{\frac{1}{2}}, +\sqrt{\frac{1}{2}})$  and  $x_b = (+\sqrt{\frac{1}{2}}, -\sqrt{\frac{1}{2}})$ . Up to multiplication by a constant,  $\int_{\partial\mathbb{U} \cap H} \max\{u \cdot [a, b]\} dm_\succeq(u) = v_a a + v_b b$ , as above.

Unless specifically noted, preferences are henceforth assumed to be continuous and to satisfy independence.

**3.5. A direct sum decomposition.** In the two outcome case, every  $[a, b] \in \mathbb{K}$  can be represented as  $\{c\} + [-r, +r]$  where  $c = \frac{a+b}{2}$  is the center of the interval and  $r = \frac{b-a}{2}$  is the radius. Attitudes toward risk are captured by an affine function of  $c$ , attitudes toward ambiguity are captured by a linear function of  $r$ , and the risk-ambiguity tradeoff by the relation between the two slopes. The general form of this decomposition is a direct sum.

Still in the two outcome case, let  $\mathbb{K}_S$  be the class of singleton sets, and  $\mathbb{K}_0$  the class of intervals centered at 0. The class  $\mathbb{K}$  of interval subsets of  $[0, 1]$  is a subset of the direct sum  $\mathbb{K}_S \oplus \mathbb{K}_0$ . This means that any linear function on  $\mathbb{K}$  can

be decomposed into the sum of its action on  $\mathbb{K}_S$  and its action on  $\mathbb{K}_0$ . This is what we showed by direct algebraic manipulation above. Steiner points provide the definition of the center of a compact convex subset of  $\mathbb{R}^n$  that generalize this direct sum decomposition.

**Definition 3.3.** *The Steiner point of a compact convex  $A \subset \mathbb{R}^n$  is the vector-valued integral*

$$St_A = \int_{\partial U} u s_A(u) d\mu(u)$$

where  $\mu$  is normalized Lebesgue measure on  $\partial U$ .

There are other informative representations of Steiner points. Define  $\mathbf{U} = \{\mathbf{u} \in H : \mathbf{u} \cdot \mathbf{u} = 1\}$ .  $K \mapsto h_K(\mathbf{u}) := \max\{\mathbf{u} \cdot x : x \in K\}$ .  $h_K(\mathbf{u})$  is a continuous function on  $\mathbf{U}$ . For  $1 < p < \infty$ ,

$$d_p(K, K') = \left( \int_{\mathbf{U}} |h_K(\mathbf{u}) - h_{K'}(\mathbf{u})|^p d\mu(\mathbf{u}) \right)^{1/p},$$

$\mu$  being normalized Lebesgue measure on  $\mathbf{U}$ . Define  $C_p(K)$  as the necessarily unique singleton set minimizing  $d_p(\{c\}, K)$ . The Steiner point is  $C_2(K)$ . The other  $C_p$  are also nice, but violate the Axiomatics of Steiner points:

It is known [17, Theorem 3.4.2, p. 167] that the mapping  $A \mapsto St_A$  is the unique mapping from  $\mathbb{K}$  to  $\mathbb{R}^n$  that is linear, continuous, and equivariant under rigid motions (i.e. if  $R : \mathbb{R}^n \rightarrow \mathbb{R}^n$  is a rigid motion, then  $R(St_A) = St_{R(A)}$ ).

For present purposes, the important observations are:

1.  $St_A$  is in the relative interior of  $A$ ,
2.  $[A^n \rightarrow A] \Rightarrow [St_{A^n} \rightarrow St_A, St_{A^n - St_A^n} \rightarrow St_{A - St_A}]$ , and
3.  $St_{A - St_A} = 0$ .

For the sake of concreteness, we work in  $H = H_\Delta - St_\Delta$ , and identify  $\Delta$  with  $\Delta - St_\Delta$ , viewed as subset of the vector space  $H$ . Let  $\mathbb{K}_S$  be the class of singleton subsets of  $H$  and  $\mathbb{K}_0 = \{A \in \mathbb{K}(H) : St_A = 0\}$ . It is immediate that  $\mathbb{K}(H) = \mathbb{K}_S \oplus \mathbb{K}_0$ , giving the direct sum decomposition  $\mathbb{K}(\Delta) \subset \mathbb{K}_S \oplus \mathbb{K}_0$ .

Any affine function,  $U$ , on  $\mathbb{K}(\Delta)$  can be expressed as the sum of an affine function,  $U_S$ , on  $\mathbb{K}_S$ , and a linear function,  $U_0$ , on  $\mathbb{K}_0$ . This decomposes  $\succeq$  into a part containing the attitude toward risk and a part containing the attitude toward ambiguity.

**3.6. Comparative ambiguity aversion.** Sets of probabilities are part of the primitive description of a choice problem. Differing reactions to sets contain the information necessary to distinguish between degrees of ambiguity aversion. The starting point is

**Definition 3.4.** *Let  $\succeq$  be a continuous, independent preference relation on  $\mathbb{K}$ . If for all  $x \in \Delta$  and all  $A \in \mathbb{K}_0$  such that  $x + A \subset \Delta$ , and for all  $0 \leq \lambda < \lambda' \leq 1$ ,*

- $x + \lambda A \succeq x + \lambda' A$ , then  $\succeq$  is **ambiguity averse**,*
- $x + \lambda A \sim x + \lambda' A$ , then  $\succeq$  is **ambiguity neutral**, and*
- $x + \lambda A \preceq x + \lambda' A$ , then  $\succeq$  is **ambiguity loving**.*

The direct sum decomposition and linearity of  $U_0$  delivers

**Lemma 3.5.** *A continuous affine  $U : \mathbb{K} \rightarrow \mathbb{R}$  represents ambiguity averse (neutral, loving) preferences if and only if the associated  $U_0 : \mathbb{K}_0 \rightarrow \mathbb{R}$  satisfies  $U_0 \leq 0$  ( $U_0 \equiv 0$ ,  $U_0 \geq 0$ ).*

For  $\lambda \geq 0$ ,  $U_0(\lambda \cdot A) = \lambda U_0(A)$ . Geometrically, this means that for  $x \in \Delta$  and  $A \in \mathbb{K}_0$ , the utility of  $x + \lambda A$  is (weakly) decreasing in  $\lambda$  for ambiguity averse preferences. As the “size” of a set centered at  $p$  increases in the  $A$ -direction, it makes an ambiguity averse DM (weakly) worse off. The more sensitive to increases in the size, the more ambiguity averse the preferences.

Let  $w$  be the worst outcome in  $X$  and  $b$  the best outcome. As above, the **risk equivalent** of  $A \in \mathbb{K}$  is that number  $p_A \in \mathbb{R}$  that satisfies  $p_A U(\{b\}) + (1 - p_A) U(\{w\}) = U(A)$ .

**Definition 3.6.**  *$U$  is more ambiguity averse than  $U^*$  if for all  $x \in \Delta$  and all  $A \in \mathbb{K}_0$  such that  $x + A \subset \Delta$ ,  $\partial p_{x+\lambda A} / \partial \lambda \geq \partial p_{x+\lambda A}^* / \partial \lambda$ .*

The present paper considers only the case of a finite  $\mathfrak{C}$ . Looking ahead to the case of consequences being an interval  $[w, b] \subset \mathbb{R}$  delivers some insight into the definition of  $U$  being more ambiguity averse than  $U^*$ .

Comparisons of degrees of risk aversion can be done by the use of certainty equivalents when consequences are  $[w, b] \subset \mathbb{R}$  —  $u$  is more risk averse than  $u^*$  if for all random variables  $X$ ,  $u(c_X) = E u(X)$  and  $u^*(c_X^*) = E u^*(X)$  imply that  $c_X \leq c_X^*$ . One can fruitfully define the ambiguity-certainty equivalent of an  $A \in \mathbb{K}$  by  $u(c_A) (= U(\{\delta_{c_A}\}) = U(A)$ . However, using certainty equivalents to compare degrees of ambiguity aversion conflates attitudes toward risk and

attitudes toward ambiguity. The ambiguity-certainty equivalent  $c_A$  is equal to the usual certainty equivalent of the lottery  $p_A\delta_b + (1 - p_A)\delta_w$ . This depends on how  $U$  behaves on  $\mathbb{K}_S$ , that is, on the attitude toward risk rather than the attitude toward ambiguity.

Sets of distributions over consequences arise when the function is not constant over “ambiguous” parts of the state space. [5] and [8] have provided two related definitions of preferences being more and less ambiguity averse for such models of choice under uncertainty. The difficulties encountered in that work is here circumvented by using sets of probability distributions over consequences as a description of what the decision maker knows, and focusing on preferences over sets.

**3.7. Bounds on risk-ambiguity tradeoffs.** Recall, the risk-ambiguity tradeoff is **balanced** if for all  $A \in \mathbb{K}$ ,  $\underline{u}(A) \leq U(A) \leq \bar{u}(A)$ , it is **allowable** if for all  $A \in \mathbb{K}$ ,  $0 \leq p_A \leq 1$ , and it satisfies **betweenness** if for all  $A, B \in \mathbb{K}$ ,  $[U(A) > U(B)] \Rightarrow U(A) \geq U(A \vee B) \geq U(B)$  where  $A \vee B = \mathbf{co}(A \cup B)$ .

This directly parallels Theorem 2.2, the first two are almost certainly equivalent, the third probably is. There is probably a fourth equivalence, but as the reader will observe, this is no more than a partially informed guess.

**Theorem 3.7.** *The following are equivalent:*

1. *the risk-ambiguity tradeoff is allowable,*
2. *the risk-ambiguity tradeoff is balanced, and*
3. *preferences satisfy betweenness.*

There must be some fourth set of derivative conditions on the  $\partial p_{x+\lambda A}/\partial \lambda$ 's that is equivalent to these.

#### 4. TWO KNIGHTIAN APPLICATIONS

Expected utility models of risky choice specify preferences over distributions over consequences, that is, over  $\mathbb{K}_S$ , as well as a *function* from choices to distributions. For example, in modeling the demand for insurance, the distribution over monetary outcomes is a function of the choice to buy insurance, the deductible if insurance is purchased, and the degree of care exercised. The potential demander of insurance either knows, or acts as if s/he knows the distributions, and

maximizes expected utility. The applications here replace the function with a *correspondence*, and the decision maker either knows the set of probabilities, or acts as if s/he knows them.

There are two ways that the present models can reduce to expected utility models. Most simply, if the correspondence takes values only in  $\mathbb{K}_S$ , it is indistinguishable from a function, and the part of preferences that depends on  $\mathbb{K}_0$  does not matter. Also, when the correspondence takes values in  $\mathbb{K} \setminus \mathbb{K}_S$  but preferences are ambiguity neutral, one can replace any set  $A$  by the single distribution,  $St_A$ . In this sense, an ambiguity neutral decision maker is not distinguishable from a classical expected utility maximizer.

**4.1. A role-based efficiency theorem.** Knight's [10] theory of profits above and beyond economic profits is based on ambiguity. In his analysis, entrepreneurs have extra ability or willingness to deal with the choices that must be made in the course of endeavors in which experience provides little or no guide in the assignment of probabilities.

Knight argues that ambiguity is more prevalent in economies introducing new products [10]. A new product is one for which there is no actuarial experience with which to estimate the distribution of sales. For Knight, the ability to deal with such ambiguities is what defines entrepreneurs, leading them to be the residual claimants. This ability earns a premium above and beyond normal economic profits. We see Knight's pattern in the efficient contracts between an ambiguity averse and an ambiguity neutral agent.

$X \in M = \{1, \dots, M\}$  is a random reward to be efficiently split between two agents, one less ambiguity averse than the other.  $X \sim P, P \in A \in \mathbb{K}(\Delta(M))$ . This is an ambiguous problem if  $A$  contains more than one point. An allocation is a mapping  $x \mapsto f(x)$  where  $f(x)$  belongs to  $\Delta(x)$ , the set of probability distributions on  $\{(s_1, s_2) : s_1 + s_2 = x\}$ . The set of such  $f$ 's is the compact convex set  $F = \Pi_x \Delta(x)$ .

The two agents receive  $w_i + s_i \in C_i$  where  $w_i$  is their initial wealth. I assume that  $|s_i| \leq M < w_i$ , I need some bounds, these seem to cover many interesting possibilities.

Remember the results when  $A$  contains only point — efficient contracts have the least risk averse agent bearing all the risk. If  $i$  is the least risk averse, then

$f_j(x) \equiv s_1^\circ$  for some  $s_i^\circ$ . The less risk averse person typically receives a premium for bearing the risk, and the more risk averse person is happy to pay it. Adding ambiguity introduces another layer of potential complication.

The preferences of each  $i$  is represented by  $U_i$ , a continuous affine function on  $\mathbb{K}(\Delta(C_i))$ . For a distribution  $\mu \in \Delta(M)$ ,  $f_i(\mu)$  is the marginal distribution of  $f(\mu)$  on  $C_i$ .  $f_i(A) := \{f_i(\mu) : \mu \in A\}$ . The mapping  $f \mapsto f_i(A)$  is linear. Therefore the efficient  $f$ 's are the solutions to

$$\max_{f \in F} \lambda U_1(f_1(A)) + (1 - \lambda) U_2(f_2(A))$$

for some  $\lambda \in [0, 1]$ .

Given all the linearity, we have

**Lemma 4.1.** *The set of efficient contracts always contains an extreme point of  $F$ .*

Both agents are risk neutral but differ in their ambiguity aversion. The agents know, or act as if they know, that the distribution of  $X$  belongs to  $B$ , which has non-empty relative interior. Agent  $i$  is ambiguity averse with  $U_0^i(A) < 0$  for any  $A$  having non-zero diameter. Agent  $j$  is effectively ambiguity neutral, either having  $U_0^j(A) \equiv 0$  or acting as if they know that the distribution is  $St_B$ .

Efficient allocations are the solutions to

$$(3) \quad \max U^i(W_i + Y_i) \quad \text{subject to} \quad U^j(W_j + Y_j) \geq v^j.$$

Since  $j$  is risk and ambiguity neutral, the constraint can be re-written as  $W_j + E(X - f(X)) \geq v^j$ , equivalently,  $E f(X) \leq y_j$  where expectations are taken with respect to  $P := St_B$ . Let  $C = f(B)$ , that is,  $C = \{x : \exists x' \in B, x = f(x')\}$ . Since  $Y_i = f(X)$ , the efficient allocations solve

$$(4) \quad \max (W_i + E Y_i) + U_0^i(C - St_C) \quad \text{subject to} \quad E Y_i \leq y_i.$$

**Lemma 4.2.** *For any  $y_j \in \mathfrak{C}$ , the efficient allocation is  $f \equiv y_j$ .*

**Proof:** The function  $f \equiv y_j$  is efficient among the constant functions, and gives  $i$  a utility of  $W_i + y_j$ . Since  $B$  has non-empty relative interior,  $C = \{St_C\}$  if and only if  $f$  is constant. If  $C = \{St_C\}$ , then  $U_0^i(C - St_C) = 0$ , and if  $C \neq \{St_C\}$ , then  $U_0^i(C - St_C) < 0$ . Therefore, any non-constant function  $f$  with  $E f(X) \leq y_b$  gives utility strictly less than  $W_i + y_j$ . ■

The residual claimant in an efficient contract is the ambiguity neutral one. The more ambiguity averse  $i$  becomes, the higher the premium that  $j$  collects for absorbing the ambiguity.<sup>3</sup>

**4.2. Role based efficiencies.** Knight also gave a role-based argument for who will be the residual claimant. If there is more ambiguity attached to the choices of one agent, then one expects that agent to be the residual claimant. There is a moral hazard argument in here.

My model-based version of this argument: Two players must choose their actions/effort levels simultaneously. The choices of one agent have larger effects on the ambiguity of the outcome. This structure defines the roles of the two players.

In general, tradeoffs between incentive effects and ambiguity effects are more complicated than the tradeoffs between incentive effects, and those are already very complicated. One can, however, see the patterns that lead to (something like) the residual claimant being the one whose actions have a large effect on the ambiguity of the outcome.

The basic simultaneous choice inefficiency when there are external effects. Turn into a random variable story. Add some  $K_0$ . Make one person's increased efforts reduce ambiguity, the other's have no effect. Note how you can improve efficiency by making one closer to a residual claimant.

## 5. GAMES

A Nash equilibrium without ambiguity is a vector,  $(\sigma_i^*)_{i \in I}$ , of beliefs which, when held by all players, belongs, componentwise, to the set of best responses to the beliefs. An equilibrium with ambiguity is a vector of sets of beliefs,  $(K_i^*)_{i \in I}$ , which, when held by all players, is a subset of, componentwise, the set of best responses to the set of beliefs.

---

<sup>3</sup>See [13] for an analysis of inefficient contracts in the face of two sided pessimistic preferences. Most economic activity, including most of the economic innovation, occurs in large organizations. Knight argues that large organizations can pool many ambiguous ventures, and by the regularity of laws of large numbers, reduce the ambiguity to (something closer to) risk. The variance calculations for set-valued random variables in [11] seem to be the appropriate tools to formalize this intuition.

We often think of equilibrium as a point where “things have settled down.” The present view of ambiguity is as a description of incomplete knowledge about the probabilities over consequences associated with different actions. There is a tension between “incomplete knowledge” and having “settled down.” This tension appears in the interpretations of the ambiguous equilibria of several of the games examined here.

**5.1. Definitions.**  $\Gamma(u) = (A_i, u_i)_{i \in I}$  defines a game with player set  $I = \{1, 2\}$ , strategy set  $A_i$  and expected utility preferences  $u_i$  for each  $i \in I$ . With some redundancy in the notation, a game with ambiguity is  $\Gamma_{Amb}(u, U) = (A_i, (u_i, U_i))_{i \in I}$ . The redundancy arises because each  $U_i$  is an extension of the corresponding  $u_i$ . Let  $K_i \in \mathbb{K}_i := \mathbb{K}(\Delta(A_i))$  be a non-empty, compact, convex subset of  $\Delta(A_i)$ .

In equilibrium, we assume that people have figured out what they are doing, that is, they have no ambiguity about their own actions. This means that for each fixed  $K = (K_i)_{i \in I}$ , the mapping  $\mu_i \mapsto U_i(\{\mu_i\} \times K_j)$  is affine.

Let  $K = (K_i)_{i \in I}$ .  $i$ 's best response sets to beliefs  $K$  with preferences  $U_i$  are  $Br_i(U_i, K)$ . By linearity, for all  $U_i$  and all  $K$ ,  $Br_i(U_i, K) \in \mathbb{K}_i$ .

**Definition 5.1.**  $K^* = (K_i^*)_{i \in I} \in \times_{i \in I} \mathbb{K}_i$  is an **equilibrium set** for  $\Gamma_{Amb}(u, U)$  if for all  $j \in I$ ,

$$K_j^* \subset Br_j(U_j, K^*).$$

The first observation about the definition is that Nash equilibria have minimal amounts of ambiguity.

**Lemma 5.2.**  $K^* = \{\sigma^*\}$  is an equilibrium for  $\Gamma_{Amb}(u, U)$  iff  $\sigma^*$  is an equilibrium for  $\Gamma(u)$ .

**Proof:**  $\{\sigma_j^*\} \subset Br_j(U_j, \{\sigma^*\})$  iff  $\sigma_j^* \in Br_j(u_j, \sigma^*)$ . ■

A second observation is that when  $K^*$  is not a singleton set, there is a great deal of wiggle room in the definition. This arises directly from the assumption that people know what they are doing. Since each  $K_j^*$  is a subset of  $Br_j(U_j, K^*)$ ,  $j$  may be playing any  $\sigma_j \in K_j^*$ . This is why it is called an equilibrium set. I tend to think that  $j$  plays the center of  $K_j^*$ , and others' guesses are centered around the truth.

The third observation is also true of Nash equilibria. Until now, we've used sets of probabilities as descriptions of what people know. Now, an equilibrating notion is providing bounds.<sup>4</sup> Some examples may make this clearer.

**5.2.  $U_i$ -dominance solvable games.** An action  $a_i$  is  $U_i$ -dominated if for all  $K$ , there exists a  $b_i$  such that  $U_i(K \setminus a_i) < U_i(K \setminus b_i)$ .

	Left	Right
Top	(8, 1)	(6, 0)
Bottom	(4, 0)	(2, 1)

There is a  $v_{\text{Top}}$  and a  $v_{\text{Bottom}}$  for 1's preferences. If preferences are balanced, then for all  $K_2$ ,

$$6 \leq U_1(\text{Top} \times K_2) \leq 8 \quad \text{and} \quad 2 \leq U_1(\text{Bottom} \times K_2) \leq 4.$$

This means that  $Br_1(U_1, K) \equiv \{\text{Top}\}$ , that Top is a  $U_1$ -dominant strategy. Therefore, in any equilibrium  $K_1^* = \{\text{Top}\}$ . If we are in an ambiguous equilibrium for this game, then 2 must be sure that 1 will play Top. Give 2's surety, 2's unique best response is Left, and the unique ambiguous equilibrium for this  $U_i$ -dominance solvable game is the unique Nash equilibrium.

To see the tension between ambiguity and equilibrium, consider the following informational situations. In the first situation, any sharp equilibrium prediction seems to be inappropriate. In the second, the equilibrium prediction may still seem to be inappropriately sharp.

1. You are involved in a strategic interaction with another player. You both have two choices.
2. More detail: You both know your preferences over sets of distributions over the four possibilities. You know nothing about the other person's preferences.
3. Yet more detail: You both know that 1 has a  $U_1$ -dominant strategy.

The point here is that the equilibrium requirement,  $K_j^* \subset Br_j(U_i, K^*)$ , has informational implications.

---

<sup>4</sup>I think that this is where the real tension between equilibrium and ambiguity arises.

5.3.  $u_i$ - but not  $U_i$ -dominance solvable games. The following game is  $u_i$ -dominance solvable, but not  $U_i$ -dominance solvable:

	Left	Right
Top	(8, 1)	(5, 0)
Bottom	(7, 0)	(4, 1)

For many different intervals  $[a, b]$ ,  $0 < a < b < 1$  there exist  $v_{\text{Top}} > v_{\text{Bottom}}$  such that 1's preferences are balanced and

$$U_1(\text{Top} \times [a, b]) = U_1(\text{Bottom} \times [a, b]).$$

Equal likelihood of Top and Bottom makes 2 indifferent. If  $v_{\text{Left}} = v_{\text{Right}}$ , then for any  $0 < s < \frac{1}{2}$ ,

$$K^* = ([\frac{1}{2} - s, \frac{1}{2} + s], [a, b])$$

is an ambiguous equilibrium.

The  $u_1$ -dominance argument for Top loses its force because 1's beliefs about 2 are a (widish) interval and this ambiguity is more tolerable when 1 chooses Bottom.

Attitudes toward ambiguity about other's actions may be dependent on own actions: 1 is a parent who could (Top) make sure their own child knew about the perils and pleasures of sex and drugs and rock and roll, or could (Bottom) try to protect them from such things by telling them about the evils inherent in worldly pleasures. 2 is a friend of the child who could (Left) go to a party with the sex and drugs and rock and roll crowd or (Right) with to a party with the chaperoned church crowd. The choices made by the friends of one's child(ren) affect a parent's utility.

Not knowing the distribution over what the child is doing is more tolerable when choosing Bottom because they have performed the requisite public displays of morality.

If you thought that the child would want to experience whatever was NOT taught, you could switch the 1's and 0's,

	Left	Right
Top	(8, 0)	(5, 1)
Bottom	(7, 1)	(4, 0)

**5.4. Ambiguity about monitoring.** Suppose that a monitoring agency has a budget devoted to monitoring for violations of worker safety laws. The budget is sufficient to monitor  $b$  of the  $M$  firms for violations in any given period. The fine for being caught is  $f$ , the net benefit from wrong-doing is  $\pi$ , firm  $m$  will violate the worker safety laws if

$$\pi(1 - p_m) - fp_m > 0 \text{ equivalently, } \pi/f > p_m/(1 - p_m),$$

where  $p_m$  is the probability that firm  $m$  is monitored.

Let  $p^*$  solve  $\pi/f = p/(1 - p)$ . If  $b/M > p^*$ , then random monitoring of all firms achieves complete compliance.<sup>5</sup> The solution for the monitoring agency, if they are interested in maximizing the amount of compliance, is to monitor  $M'$  of the firms at random where  $M'$  is the largest integer satisfying  $b/M' > p^*$ .

Another possibility is to introduce ambiguity about the size of  $b$  into the firms' decision problems. Not violating worker safety laws gets 0 for sure, with ambiguity about  $b$ , the probability of being monitored is in some interval  $\{c\} + [-r, +r]$ , and with ambiguity aversion, you can get deterrence with  $c < p^*$ , effectively stretching the budget of the monitoring agency. One does not expect this to be a permanent, or an equilibrium, solution.

**5.5. Ambiguity about punishment.** If people are risk averse, one can randomize over monetary punishments and get a larger deterrent effect on the risk averse. One can keep the legal system in a turmoil, say by regular Congressional re-writing of sentencing laws, so that the distribution of punishments associated with conviction is not known. This has a larger deterrent effect on the ambiguity averse.

A caveat: both randomization according to a stable (therefore learnable) distribution and the introduction of ambiguity to the punishments encourage criminal behavior among the more risk-tolerant and the more ambiguity tolerant. Perhaps they are the ones needing the least encouragement. More importantly, both proposals undercut the perceived legitimacy of the legal system, and it is surely this legitimacy that keeps many obeying the laws.

---

<sup>5</sup>There are horror stories about the small size of monitoring budgets for worker safety in the US.

## 6. SOME OTHER APPLICATIONS

**6.1. “Irrationality” in the face of small probabilities.** People in the United States often report being more afraid of dying in a plane crash, a terrorist bombing or a meteorite strike than by, say, the far more likely event of a car accident. Plane crashes, meteorite strikes, and terrorist bombings are outside the experience range of the circle of friends of most Americans, so there is little by way of commonly known evidence for the size of the probabilities. If ambiguity enters as the description of this lack of evidence, and the decision maker is ambiguity averse, they might well be made more unhappy by the ambiguous prospect of death by plane crash, bomb or meteor than by the risky prospect of death by car.

**6.2. Pessimistic and Optimistic Equilibrium Sets.** Suppose that pessimists play each other, what are the equilibrium sets with ambiguity? What about optimists?

**6.3. Portfolio Choice Theory.** We estimate the mean vector and variance-covariance matrix of returns and then pick optimal portfolio weights. We have a set of possible distributions over the true mean and v-cv matrix.

**6.4. Demand for ambiguity reduction.** Updating through events reduces ambiguity (by lowering dimensionality of any set  $A$  with non-empty interior). Other linear reductions are valuable. For a given linear  $U$  and set  $A$ , which reduction affects  $p_A$  the most?

## 7. EXTENSIONS

One of these looks easy, the other sometimes looks very easy and sometimes impossible.

**7.1. State spaces and Savage models.** If the state space,  $S$ , is finite, then  $U$  can be a continuous linear function on subsets of  $\Delta(S \times \mathfrak{C})$ , e.g. subsets that are graphs of functions. This also allows for state dependent preferences, as in [9].

**7.2. Infinite consequence spaces.** It is known [18] that there is no continuous extension of the Steiner point mapping to infinite dimensional Hilbert spaces. However,  $\Delta([0, M])$  is a compact convex subset of a Hilbert space. Maybe there

is a continuous extension to the compact subsets of  $\Delta([0, M])$ . This would depend on special properties of  $\Delta([0, M])$ .

## 8. CONCLUDING REMARKS

This present paper differs from the previous literature by changing the domain over which preferences are defined. Sets of probabilities are here part of the description of a choice problem. The previous literature derives the existence of sets of probabilities on a state space and a utility function on a space of consequences from axioms about preferences over functions from states to consequences. This difference makes the present theory relatively simple.

There is an analogy to textbook coverages of choice under uncertainty. The simplest approach is to suppose that probabilities are given, to invoke independence, deduce expected utility, and to study applications on the assumption of known probabilities. The second, more complicated but still textbook, approach deduces the existence of a subjective probability and an expected utility function from axioms about preferences over acts, modeled as functions from a state space to consequences. This paper is doing the first, not the second.

## REFERENCES

- [1] Allais, M. 1953, Le Comportement de l'Homme Rationnel Devant le Risque: Critique des Postulats et Axiomes de l'Ecole Americaine. *Econometrica* **21**, 503-546
- [2] Amarante, M. 2003, States, Models and Unitary Equivalence. Working paper, Department of Economics, Columbia University, New York.
- [3] Crawford, V. 1990, Equilibrium without Independence. *Journal of Economic Theory* **50**, 1, 127-54.
- [4] Dekel, E., B. L. Lipman, and A. Rustichini, 2001, Representing Preferences with a Unique Subjective State Space *Econometrica* **69**, 891-934.
- [5] Epstein, L. 1999, A Definition of Uncertainty Aversion. *Review of Economic Studies* **66**, 579-608.
- [6] Epstein, L. and J. Zhang, 2001, Subjective Probabilities on Subjectively Unambiguous Events. *Econometrica* **69**, 265-306.
- [7] Fenchel, W. and Jessen, B. 1938, Mengenfunktionen und konvexe Körper. *Danske Vid. Selskab. Mat.-fis. Medd.* **16**, 3.
- [8] Ghirardato, P. Marinacci, M. 2001, Range Convexity and Ambiguity Averse Preferences. *Economic Theory* **17**(3), 599-617.
- [9] Karni, E., Schmeidler, D., Vind, K. 1983, On State Dependent Preferences and Subjective Probabilities. *Econometrica* **51**(4), 1021-1032.
- [10] Knight, F. H. 1921, *Risk, uncertainty and profit*. Chicago, University of Chicago Press (1971, reprint of the 1921 ed.).

- [11] Körner, R. 1995, A Variance of Compact Convex Random Sets. Working paper, Institut für Stochastik, Bernhard-von-Cotta-Str. 2 09599 Freiberg.
- [12] Machina, M. 1982, 'Expected Utility' Analysis Without the Independence Axiom. *Econometrica* **50**, 277-323.
- [13] Mukerji, S. 1998, Ambiguity Aversion and Incompleteness of Contractual Form. *American Economic Review* **88**(5), 1207-31.
- [14] Mukerji S.; Tallon J-M. 2001, Ambiguity Aversion and Incompleteness of Financial Markets. *The Review of Economic Studies* **68**(4), 883-904.
- [15] Von Neumann, J. and Morgenstern, O. 1944, *Theory of games and economic behavior*. New York, Wiley, (1953, 3d ed.).
- [16] Savage, L. J. 1954, *The foundations of statistics*. New York, Wiley.
- [17] Schneider, R. 1993, *Convex Bodies: The Brunn-Minkowski Theory*, Cambridge, Cambridge University Press.
- [18] Vitale, R. 1985, Steiner Points in Infinite Dimensions, *Israel Journal of Mathematics*.

DEPARTMENT OF ECONOMICS, UNIVERSITY OF TEXAS, AUSTIN, TX 78712-0301 USA,  
 e-mail: maxwell@eco.utexas.edu