CONSEQUENTIALIST AMBIGUITY: CHOICE UNDER AMBIGUITY AS CHOICE BETWEEN 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 the decision maker is choosing between incompletely described outcomes, that is, between 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.

1. Introduction

> 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 [7])

This paper takes Epstein and Zhang's rough distinction as the defining difference between risky choice problems and ambiguous choice problems, and takes the "relevant events" to be sets of consequences. If the decision maker (DM) knows the probability distributions associated with their choices, it is a risky problem. If the DM knows only that the probability distributions associated with their choices belong to some set, it is ambiguous.

Under study are continuous linear preferences over *sets* of distributions over consequences. Restricted to singleton sets, these are von Neumann-Morgenstern (vN-M) preferences, which contain the DM's attitude toward risk. Expansions away from and contractions towards the center of a set increase and decrease the ambiguity in a decision problem. The utility effects of these expansions or contractions characterize the DM's attitude toward ambiguity.

Date: First version: February 12, 2003. This version: April 19, 2007
Key words and phrases. Ambiguity, decision theory, information, options, equilibria.
Many thanks to the Thursday theorists, Svetlana Boyarchenko, Takashi Hayashi, and Dale Stahl for frequent help with this paper.

These preferences simplify modeling because they allow for a complete separation of attitudes toward risk and attitudes toward ambiguity. The major cost of this simplification is the need for a more detailed description of the choice problem.

The next section develops the theory, first for the two outcome case, then for the finite outcome case, and then for an interval set of consequences. The following section (begins) the applications.

2. Theory

This section begins by treating the basic idea behind the separation of attitudes toward risk and attitudes toward ambiguity in the simplest setting, where the set of consequences, $\mathbf{C}$, contains only two points. The essential device is a change of basis, expressing every interval of probabilities, $[q, s]$ as $[p - r, p + r]$ where $p = (q + s)/2$ is the center and $r = (s - q)/2$ is radius. Equivalently, this views the set $[q, r]$ as the set $\{p\} + [-r, +r]$, a point plus a set centered at 0. Linear utility functions on intervals are, after re-scaling if necessary, of the form $p - vr$, and v measures the tradeoff between risk and ambiguity.

Following this is the theory for $\mathbf{C}$ being a finite set. Again, the essential idea is to write sets of probabilities as a centerpoint plus a set centered at 0. Attitudes toward risk are carried in the utility of the centerpoint, attitudes toward ambiguity in the part of the utility that depends on the set centered at 0. Here the tradeoff between risk and ambiguity depends on the shape of the set centered at 0.

Extending this type of analysis of continuous linear preferences on sets of probabilities to the case where the set of consequences is $\mathbf{C} = [0, M]$ is more difficult. The main problem is the non-existence of a continuous notion of the center of an infinite dimensional set of probabilities. However, over restricted domains in $\Delta([0, M])$, it is possible to study the analogues of first and second order stochastic dominance, and the risk equivalents of ambiguous choices.

2.1. The Basic Idea.

The best-known ambiguous choice problem is due to Ellsberg [5]. It presents a paradox for any choice theory based on beliefs about distributions over consequences. If this is relevant to descriptions of behavior, then preferences over probability distributions are descriptively invalid for large classes of choice problems.

2.1.1. *Ellsberg Urns.*

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

(1) Choices between single tickets:
   (a) The choice between the Red ticket or the Green ticket.

(b) The choice between the Red ticket or the Blue ticket.

(2) Choices between pairs of tickets:
    (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 DM gets $\$1,000$, otherwise they get nothing, a two-point set of consequences. Modal preferencesare

$$R \succ G \text{ and } R \succ B, \text{ as well as}$$
$$R\&B \prec G\&B \text{ and } R\&G \prec B\&G.$$

If people like higher probabilities of better outcomes and assigned probabilities to these events, then we have

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

The probability that the Red ticket wins is $\frac{1}{3}$. That is, the action "choose Red" is risky, with the known 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, $[\frac{1}{3}, 1]$. The preferences $R \succ G$ and $R \succ B$ correspond to $\{\frac{1}{3}\} \succ [0, \frac{2}{3}]$, while the preferences $R\&B \prec G\&B$ and $R\&G \prec B\&G$ correspond to $[\frac{1}{3}, 1] \prec \{\frac{2}{3}\}$. In these terms, a summary of the Ellsberg paradox is that people prefer knowing that the uncertain outcome is distributed according to the center of the interval to the interval itself.

This paper studies preferences over sets of probabilities as an explanation of ambiguity aversion. An alternative explanation, due to Fellner [8], is that people may believe that some potentially mischievious force, e.g. experimentalists, cheat by changing the number of balls in the urn after the ticket is chosen. As will be shown, there is little to distinguish between these explanations.

2.1.2. *Separation of Risk and Ambiguity Attitudes.* When there are but two consequences, say 0 and 1 (thousand dollars), linear preferences, $\succsim$, on, $\mathbb{K} = \mathbb{K}([0,1])$, the closed, convex subsets of the probabilities $[0,1]$, have an informative representation. Intervals $[q, s]$, $0 \leq q \leq s \leq 1$ can be represented as vectors $(q, s) \in \mathbb{R}^2$ with $q \leq s$. A change of basis is $[q, s] = [p - r, p + r]$ where where $p := \frac{1}{2}q + \frac{1}{2}s \in [0, 1]$ is the center of the interval and $r := \frac{1}{2}(s - q) \in [0, \frac{1}{2}]$ is its radius. With $\{p\} = [p, p]$ and the normalization $U([0, 0]) = 0$, $U([1, 1]) = 1$, the following is immediate.

Lemma 1. *The unique normalized representation of continuous, affine preferences on* $\mathbb{K}$ *is* $U([p-r, p+r]) = p - vr$.

This gives a complete separation between the attitude toward risk, the p part of the utility function, and the attitude toward ambiguity, the vr part of the utility function. The cost is that the amount of ambiguity, r, must be specified as part of the detail of the model of the decision problem.

The parameter v measures the tradeoff between ambiguity and risk, and characterizes the preferences when there are only two outcomes. 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 $v > 0$, and v can be elicited.

2.1.3. *Elicitation.* Eliciting the v's that characterize different decision makers can be done by offering choices between ambiguous choices and risky choices. In the Ellsberg case, introduce a second urn with q 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) Choices between single tickets:
   (a) The choice between the Green or the Yellow ticket.
   (b) The choice between the Blue or the Yellow ticket.
(2) Choices between pairs of tickets:
   (a) The choice between the Red-Green pair of tickets or the White ticket.
   (b) The choice between the 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.

With $A = [0, \frac{2}{3}]$, the **risk equivalent** p_A that makes the Green (or the Blue) ticket indifferent to the Yellow ticket satisfies $p_A = \frac{1}{3} - v \cdot \frac{1}{3}$ so that $v = (1 - 3p_A)$. Positive v's correspond to $p_A < \frac{1}{3}$. With $B = [\frac{1}{3}, 1]$, the p_B that makes the Red-Green (or the Red-Blue) pair of tickets indifferent to the White ticket indifferent satisfies $(1 - p_B) = \frac{2}{3} - v \cdot \frac{1}{3}$ so that $v = 3p_B - 1$. In this case, positive v's correspond to $p_B > \frac{1}{3}$, that is, $1 - p_B < \frac{2}{3}$.

Inspection shows that $-1 \leq v \leq +1$ is necessary and sufficient to keep risk equivalents, p_A and p_B, in the interval $[0, 1]$. This is not the only reason for such a restriction on v, and we will return to this below.

2.1.4. *Choice of a Center.* The key step in the analysis is to express an $A \in \mathbb{K}$ as $[A - s(A)] + s(A)$ where $s(A)$, the center of A, depends continuously and linearly on A. This allows the separation of U into a sum of its action on sets centered at 0, and its action on probabilities. Candidates for the continuous linear 'center' of $[a, b]$ are $s_\gamma([a, b]) := \gamma a + (1-\gamma)b$, $\gamma \in [0, 1]$.

Use of a different weighted center is a change of basis in the representation of the linear function, U. This means that the definition of ambiguity neutrality — indifference to expansions and/or contractions around the center or $v = 0$ — is specific to this choice of center. In more detail, using $\gamma = \frac{1}{2}$ as above defines ambiguity neutrality by indifference between $[p - r_1, p + r_1]$ and $[p - r_2, p + r_2]$ for all p and all $r_1, r_2 \geq 0$. By contrast, if the center of each $[a, b]$ is defined as $\frac{2}{3}a + \frac{1}{3}b$, ambiguity neutrality is defined by $[p - r_1, p + 2 \cdot r_1] \sim [p - r_2, p + 2 \cdot r_2]$ for all p and all $r_1, r_2 \geq 0$, and ambiguity aversion by $[p - r_1, p + 2 \cdot r_1] \succ [p - r_2, p + 2 \cdot r_2]$ for all p and all $r_2 > r_1 \geq 0$.

The choice of $\gamma = \frac{1}{2}$ is the unique choice of continuous linear center that respects (i.e. commutes with) rigid motions and translations — think of the rigid motion and translation that takes $[a, b]$ to $[1 - b, 1 - a]$. More generally, for finitely many consequences, there is a unique continuous linear definition of the center of a compact convex set that respects rigid motions and translations, called the Steiner point. As in the two outcome case, choosing the Steiner point as center has implications for what it means to model someone as being ambiguity averse.

2.2. **Finite Consequence Spaces.** We now discuss continuous, linear preferences on the set of compact, convex subsets of distributions when the "relevant events" are subsets of a finite set of consequences.

2.2.1. *Notation.* The finite space of consequences is $\mathbf{C}$, which has ℓ elements, typically denoted a, b, c, x, y. $\Delta = \Delta(\mathbf{C}) = \{p \in \mathbb{R}^{\mathbf{C}}_+ : \sum_x p_x = 1\}$ is the set of probabilities on $\mathbf{C}$, with typical elements p, q, r, s. By convention, we analyze Δ within its affine hull, which, by translation, is taken to be $\mathbb{R}^{\ell-1}$.

$\mathbb{K}(\mathbb{R}^{\ell-1})$ is the cone of non-empty, compact, convex subsets of $\mathbb{R}^{\ell-1}$. $\mathbb{K}(\Delta) \subset \mathbb{K}(\mathbb{R}^{\ell-1})$ is the set of non-empty, compact, convex subsets of Δ, with typical elements A, B, C, D. Define $\alpha A + \beta B = \{\alpha x + \beta y : x \in A,\ y \in B\}$ for $A, B \in \mathbb{K}(\mathbb{R}^{\ell-1})$ and $\alpha, \beta \in \mathbb{R}_+$. $\mathbb{K}(\Delta)$ is a convex subset of $\mathbb{K}(\mathbb{R}^{\ell-1})$, and is compact in the usual metric.[1]

A function $U : \mathbb{K}(\mathbb{R}^{\ell-1}) \to \mathbb{R}$ is **linear** if for all $\alpha, \beta \geq 0$ and all $A, B \in \mathbb{K}(\mathbb{R}^{\ell-1})$, $U(\alpha A + \beta B) = \alpha U(A) + \beta U(B)$. As with linear functions, $U(0) = 0$. If $U : \mathbb{K}(\Delta) \to \mathbb{R}$, it is **linear** if $U(\alpha A + (1-\alpha)B) = \alpha U(A) + (1-\alpha)U(B)$ for $\alpha \in [0, 1]$ and $A, B \in \mathbb{K}(\Delta)$. We will see that a linear U on $\mathbb{K}(\mathbb{R}^{\ell-1})$ is determined by its values on $\mathbb{K}(\Delta)$

[1]This is not a vector space because $A + \{0\} = A$ for all A, but $A - A \neq \{0\}$ unless A is a singleton set. For $x \in \mathbb{R}^{\ell-1}$, $B_\epsilon(x)$ denotes the usual (Euclidean) ϵ-ball around x, and for $A \in \mathbb{K}(\mathbb{R}^{\ell-1})$, $A^\epsilon := \cup_{p \in A} B_\epsilon(p)$ is the ϵ-ball around A. The (Hausdorff) metric on $\mathbb{K}(\mathbb{R}^{\mathbf{C}}_+)$ is defined by $d_H(A, B) := \inf\{\epsilon > 0 : A \subset B^\epsilon, \text{ and } B \subset A^\epsilon\}$. See [18] for the isometrically isomorphic embedding of this space of convex sets in a Banach space.

The unit sphere is $\boldsymbol{U} = \{\boldsymbol{u} \in \mathbb{R}^{\ell-1} : \boldsymbol{u} \cdot \boldsymbol{u} = 1\}$, and λ is the uniform (probability) distribution on $\boldsymbol{U}$. The support function of $A \in \mathbb{K}(\mathbb{R}^{\ell-1})$ is defined by $\mu_A(\boldsymbol{u}) = \max_{a\in A} \boldsymbol{u} \cdot a$ for $\boldsymbol{u} \in \boldsymbol{U}$. The corresponding argmax function as $\arg_A(\boldsymbol{u}) := \{a \in A : \boldsymbol{u} \cdot a = \mu_A(\boldsymbol{u})\}$. For all but a λ null set of $\boldsymbol{U}$, $\arg_A$ is a singleton set.

2.2.2. *The Independence Axiom.* An extension of compound lottery logic interprets $\alpha A + (1 - \alpha)B$ for $A, B \in \mathbb{K}(\Delta)$, $\alpha \in [0, 1]$. This, and a the set version of the Independence Axiom of von Neumann-Morgenstern (vN-M) expected utility theory delivers continuous, linear preferences. We begin with

Definition 1. *A* ***continuous rational preference relation on*** $\mathbb{K}(\Delta)$ *is a complete, transitive relation,* $\succsim$*, on* $\mathbb{K}(\Delta)$ *such that for all* $B \in \mathbb{K}(\Delta)$*, the sets* $\{A : A \succ B\}$ *and* $\{A : B \succ A\}$ *are open.*

If A and B are singleton sets, then $\alpha A + (1 - \alpha)B$ has a compound lottery interpretation as a initial lottery over lotteries. The inital lottery picks the lottery A with probability α and picks the lottery B with probability $(1-\alpha)$. There is a similar interpretation for convex combinations of ambiguous outcomes. Indifference between the timing of the resolution of different types of uncertainty is the content of the following.

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

Example 1. *Suppose that there are only two consequences, a good one and a bad one.* $A = [0, \frac{1}{3}]$ *corresponds to being told (only) that the probability of the bad consequence is at least twice as large as the good one. For* $B = [\frac{2}{3}, 1]$*, the role of the good and bad consequences are reversed.*

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 $\frac{1}{2}A + \frac{1}{2}B = [\frac{1}{3}, \frac{2}{3}]$*. One arrives at the same interval by tossing a weighted coin with the relative odds of the good and the bad outcome being anyplace between* $1 : 2$ *and* $2 : 1$*.*

A preference relation on Δ satisfies continuity and independence if and only if it can be represented by a continuous linear function on Δ. The same is true for preferences on $\mathbb{K}(\Delta)$. The omitted proof of the following result directly parallels textbook treatments of expected utility preferences.

Theorem 1. *A continuous rational preference relation on* $\mathbb{K}(\Delta)$ *satisfies the Independence Axiom if and only if it can be represented by a continuous posilinear function.*

2.2.3. *Separating Risk and Ambiguity.* In the two outcome case, there is a direct sum decomposition of every $[q, s] \in \mathbb{K}(\mathbb{R}^1)$ as $\{p\} + [-r, +r]$ where $p = \frac{1}{2}q + \frac{1}{2}s$ is the center and $r = \frac{1}{2}(s - q)$ is the radius. Letting $St^{-1}(0)$ denote the set of intervals centered at 0, we can go a bit further. Every element of $\mathbb{K}(\mathbb{R}^1)$ has a unique representation as $A_0 + \{p\}$ with $A_0 \in St^{-1}(0)$ and $p \in \mathbb{R}^1$. We write this kind of direct sum decomposition as $\mathbb{K}(\mathbb{R}^1) = St^{-1}(0) \oplus \mathbb{R}^1$. The key to giving a similar direct sum decomposition with larger consequence spaces is a continuous linear definition of the center of a set. We use

Definition 1. *The* ***Steiner point of*** $A \in \mathbb{K}(\mathbb{R}^{\ell-1})$ *is the vector-valued integral*

$$St(A) = \int_{\partial \boldsymbol{U}} \arg_A(\boldsymbol{u})\, d\lambda(\boldsymbol{u}).$$

For example, $St(\Delta)$ is the distribution putting equal mass on all elements of $\mathbf{C}$ because each vertex of Δ is $\arg_\Delta(u)$ for $1/\#\mathbf{C}$ of the surface of $\boldsymbol{U}$. The following compares the center of gravity to the Steiner point.

Example 2. *Suppose there are three consequences,* $\mathbf{C} = \{a, b, c\}$*. Consider the set* A *described by "a and b are close to equally likely." Interpreting "close to" as "the ratio of the probabilities is within* ϵ *of* $\frac{1}{2}$*, in the simplex, the set* A_ϵ *is the convex hull of the points* $(\frac{1}{2} + \epsilon, \frac{1}{2} - \epsilon, 0)$*,* $(\frac{1}{2} - \epsilon, \frac{1}{2} + \epsilon, 0)$*, and* $(0, 0, 1)$*. As* $\epsilon \downarrow 0$*, the two dimensional* A_ϵ *converge to* A_0*, the one dimensional convex hull of* $(\frac{1}{2}, \frac{1}{2}, 0)$ *and* $(0, 0, 1)$*. The center of gravity of each* A_ϵ *is* $(\frac{1}{2}(1 - x), \frac{1}{2}(1 - x), x)$ *where* $x = 1 - 1/\sqrt{2} < \frac{1}{2}$*, while the center of gravity of* A_0 *is* $(\frac{1}{2}(1 - y), \frac{1}{2}(1 - y), y)$ *where* $y = \frac{1}{2}$*. On the other hand,* $St(A_\epsilon) \to St(A_0)$ *because* $St(A_\epsilon) = (\frac{1}{2}(1 - (y - \delta)), \frac{1}{2}(1 - (y - \delta)), (y - \delta))$ *where* $y = \frac{1}{2}$ *and* $\delta \downarrow 0$ *as* $\epsilon \downarrow 0$*.*

Since A is convex and the mass of λ is 1, the $St(A)$ is in (the relative interior of) A so that $A \mapsto St(A)$ is a selection. An easy application of the Theorem of the Maximum delivers the continuity and linearity of $A \mapsto St(A)$. Replacing λ with any full support probability, ν, on $\boldsymbol{U}$ would give a selection with these properties. However, combining these properties with the rotation invariance of λ shows that Steiner point is equivariant under rigid motions, that is, if $R : \mathbb{R}^n \to \mathbb{R}^n$ is a rigid motion, then $R(St(A)) = St(R(A))$. Requiring equivariance has the effect of forcing $\nu = \lambda$. For a proof of the following, see e.g. [20, Theorem 3.4.2, p. 167].

Theorem 2 (Steiner)**.** *The mapping* $A \mapsto St(A)$ *is the unique continuous linear selection that is equivariant under rigid motions.*

Note that $St(A - St(A)) = 0$ so that every $A \in \mathbb{K}(\mathbb{R}^{\ell-1})$ is the translate of a set having 0 for its Steiner point. Further, $St^{-1}(0)$ is a cone within the cone $\mathbb{K}(\mathbb{R}^{\ell-1})$, and every $A \in \mathbb{K}(\mathbb{R}^{\ell-1})$ is a translate of an element of this cone. Put another way,

Lemma 2. *$St^{-1}(0) \oplus \mathbb{R}^{\ell-1}$ is a direct sum decomposition of the cone $\mathbb{K}(\mathbb{R}^{\ell-1})$ into a cone and a vector space.*

For a set A of distributions over consequences, the direct sum decomposition gives a representation of the risk, $St(A)$, and of the ambiguity, $(A - St(A))$.

2.2.4. *Separating Risk and Ambiguity Attitudes.* The direct sum decomposition allows us to isolate the action of a linear U on risk from its action on ambiguity. The function u in the following is a vN-M utility function. This shows that linear preferences are the linear extension of expected utility preferences.

Theorem 3. *If $U : \mathbb{K}(\Delta) \to \mathbb{R}$ is linear if and only if it is of the form $U(A) = U^0(A - St(A)) + u \cdot St(A)$ for some linear U^0 on the cone $St^{-1}(0)$ and some $u \in \mathbb{R}^{\ell-1}$.*

Proof: If U is the restriction of a linear $U : \mathbb{K}(\mathbb{R}^{\ell-1}) \to \mathbb{R}$, then $U^0 := U_{|St^{-1}(0)}$ and u defined by $u \cdot p = U(\{p\})$ delivers the equivalence. It is therefore sufficient to show that U is linear on $\mathbb{K}(\Delta)$ iff it is the restriction of a linear U on $\mathbb{K}(\mathbb{R}^{\ell-1})$. This follows from the observation that every $A \in \mathbb{K}(\mathbb{R}^{\ell-1})$ is of the form $A = r \cdot (B - St(B)) + St(A)$ for some $B \in \mathbb{K}(\Delta)$ and $r \geq 0$. ∎

One implication of this result is that any linear U^0 can be combined with any vN-M u. This allows complete separation of attitudes towards ambiguity, encoded in U^0, and attitudes toward risk, encoded in u.

Definition 2. *Preferences on $\mathbb{K}(\Delta)$ represented by a linear $U = (U^0, u)$ are*

(1) ***ambiguity averse*** *if $U^0(St^{-1}(0)) = \mathbb{R}_-$,*
(2) ***ambiguity neutral*** *if $U^0(St^{-1}(0)) = \{0\}$, and*
(3) ***ambiguity seeking*** *if $U^0(St^{-1}(0)) = \mathbb{R}_+$.*

Attitudes towards ambiguity are reflected in attitudes toward expansions and contractions of a set around its center.

2.2.5. *The Choice of a Center.* As noted above, there are many continuous linear selections that can be used as the definition of the center of a compact convex set. Each gives rise to a different direct sum decomposition, hence labels a different set of linear preferences as ambiguity averse/neutral/seeking. There are, however, many cases in which the choice of the Steiner point as center seems natural.

For example, the use of the Steiner point implies, *inter alia*, that the center of any line, $[r, s]$, in Δ is $p = \frac{1}{2}r + \frac{1}{2}s$, regardless of the orientation of the line. This does **not** mean that the orientation of the lines of the same length with center c are indifferent, if $A = [r, s]$, $B = [r', s']$ with $\frac{1}{2}r + \frac{1}{2}s = \frac{1}{2}r' + \frac{1}{2}s' = p$, one can still have $U(A) \neq U(B)$.

If one accepts the correctness of the choice of center for lines, then linearity, that is, the Independence Axiom, implies that the center of all parallelograms in Δ should be the Steiner point. Applying linearity again implies that the center of a convex combinations of parallelograms with different orientations should be the Steiner point. Continuing in this fashion yields a rich class of sets, containing e.g. all parallelipipeds, on which the Steiner point is the correct center.

2.3. **An Interval of Consequences.** We now discuss s continuous, linear preferences on the set of compact, convex subsets of distributions when the "relevant events" are subsets of a closed and bounded interval subset of $\mathbb{R}$, taken to be $[0, M]$ here.

2.3.1. *A Quick Comparison with Expected Utility Preferences.* A linear function on a compact convex set is determined by its values on the extreme points of the set. For $\Delta[0, M]$, the set of (countably additive Borel) probabilities on $[0, M]$, the set of extreme points is the set of point masses, δ_a, $a \in [0, M]$, defined by $\delta_a(E) = 1_E(a)$. For a linear $L : \Delta[0, M] \to 0$, this gives rise to the von Neumann-Morgenstern utility function, defined by $u(x) = L(\delta_x)$.

It is the fact that $L : \Delta[0, M] \to \mathbb{R}$ is determined by $u : [0, M] \to \mathbb{R}$ that makes analysis of properties of u so fruitful: the vN-M utility function u is monotonic, if and only if the preferences represented by L respect first order stochastic dominance; u is concave if and only if the preferences represented by L respect second order stochastic dominance. There are parallel properties for linear functions on $\mathbb{K}(\Delta)$ but they are somewhat less useful because the set of extreme points of $\mathbb{K}(\Delta)$ is more difficult to charaterize and work with.

Finally, one can define risk equivalents of sets of probabilities that play the same role that certainty equivalents play for comparing degrees of risk aversion amongst expected utility preferences.

2.3.2. *Notation.* Here $\mathbf{C} = [0, M]$, and $\Delta = \Delta[0, M]$ is the set of countably additive Borel probabilities on $\mathbf{C}$. An element p of Δ is identified with its cdf, $F_p(x) = p(-\infty, x]$, and distance is $d(p, q) = \left[\int_0^M (F_p(x) - F_q(x))^2 \, dx\right]^{\frac{1}{2}}$. Note that

(1) $d(p_n, p) \to 0$ iff p_n converges weakly to p,
(2) for $a \leq b$ $d(\delta_a, \delta_b) = \left[\int_a^b 1^2 \, dx\right]^{\frac{1}{2}} = \sqrt{|a - b|}$, and
(3) for all p, q, $d(p, \alpha p + (1 - \alpha)q) = (1 - \alpha)d(p, q)$.

$\mathbb{K}(\Delta)$ is the set of non-empty, compact, convex subsets of Δ with the associated Hausdorff metric. For $n = 0, 1, \ldots$, $\mathbb{K}^n(\Delta)$ is the class of n-dimensional subsets of

$\mathbb{K}(\Delta)$. The class of finite dimensional subsets of $\mathbb{K}(\Delta)$ is $\mathbb{K}^{fd} := \cup_{n=0}^{\infty}\mathbb{K}^n(\Delta)$, which is dense in $\mathbb{K}(\Delta)$. Mention of the Steiner point of a set A automatically entails the assumption that $A \in \mathbb{K}^{fd}$.

For $A, B \in \mathbb{K}(\Delta)$, the line segment $[\![A, B]\!]$ is the set $\{\alpha A + (1-\alpha)B : \alpha \in [0,1]\}$. $A \in \mathbb{K}(\Delta)$ is an **extreme set** if for all line segments $[\![B, C]\!]$ in $\mathbb{K}(\Delta)$, if $A \in [\![B, C]\!]$, then either $A = B$ or $A = C$. A useful class of extreme sets are of the form $A_F = \{p \in \Delta : p(F) = 1\}$, F a closed, non-empty subset of $[0, M]$.

Lemma 3. *If F is a closed subset of $[0, M]$, then A_F is an extreme set.*

Proof: A_F is convex because $\alpha p(F) + (1-\alpha)q(F) = 1$ for $p, q \in A_F$. A_F is closed because for a sequence p^n in A_F, $[p_n \to p] \Rightarrow [p(F) \geq \limsup_n p_n(F) = 1]$. Suppose that $A_F \in [\![A, B]\!]$ and that $A_F \neq A$. There are two possibilities, either $A \subsetneq A_F$ or else A contains probabilities p with $p(F) < 1$. In the either case, for any $\alpha > 0$, $\alpha A + (1-\alpha)B \neq A_F$, so that $A_F = B$. ∎

The subclass $\{A_F : F \text{ is finite}\}$ is dense in this class of extreme sets. Specifying interesting classes of U on this dense set is often relatively easy.

2.3.3. *A Functional Form.* To get at the intuition for what respecting first and second order stochastic dominance with sets of probabilities entails, we will use a particular class of functions. This will also give intuition about relative ambiguity aversion.

For $a, b \in [0, M]$, $a \leq b$, linearity implies that specifying the function $f(a, b) := U([\![\delta_a, \delta_b]\!])$ specifies U for all parallelipeds spanned by intervals of point masses. We normalize so that $f(0, 0) = 0$ and $f(M, M) = 1$.

For a vN-M utility function $u : [0, M] \to [0, 1]$ with $u(0) = 0$ and $u(M) = 1$, and a function $v : [0, M] \to \mathbb{R}$, the following class of functions $f(\cdot, \cdot)$ is useful:

$$f(a, b) = \tfrac{1}{2}u(a) + \tfrac{1}{2}u(b) - \tfrac{r}{2}|v(b) - v(a)|^h, \text{ with } h \in [1, \infty). \tag{1}$$

The $\frac{1}{2}u(a) + \frac{1}{2}u(b)$ part arises because the Steiner point of $[\![\delta_a, \delta_b]\!]$ is $\frac{1}{2}\delta_a + \frac{1}{2}\delta_b$, and $\frac{1}{2}u(a) + \frac{1}{2}u(b)$ is the corresponding expected utility of the Steiner point. For $B = \{\alpha\delta_a + (1-\alpha)\delta_b : \alpha \in [s-t, s+t]\}$, the corresponding Steiner point is $s\delta_a + (1-s)\delta_b$, and the corresponding linear utility is $U(B) = su(a) + (1-s)u(b) - rt|v(b) - v(a)|^h$. If $r > 0$ and v is increasing, we have ambiguity aversion, if $r < 0$, ambiguity seeking. If $r \neq 0$ and v has regions of increase and regions of decrease, we have both.

Within this class, the following special cases represent risk and ambiguity averse preferences that satisfy set monotonicity, respecting first order stochastic dominance, horizontal concavity, respecting second order stochastic dominance. For the given range of values of r, the preferences also have bounded risk-ambiguity tradeoffs, a condition that will guarantee, for an ambiguous set A, the existence of a risk equivalent lottery over the best and the worst outcome. In these examples, we take u to be concave and $r \geq 0$, giving risk and ambiguity averse preferences.

(1) $f(a,b) = \frac{1}{2}u(a) + \frac{1}{2}u(b) - \frac{r}{2}|b-a|$, $0 \leq r \leq u'(M)$;
(2) $f(a,b) = \frac{1}{2}u(a) + \frac{1}{2}u(b) - \frac{r}{2}|u(b)-u(a)|$, $0 \leq r \leq 1$;
(3) $f(a,b) = \frac{1}{2}u(a) + \frac{1}{2}u(b) - \frac{r}{2}|b-a|^2$, $0 \leq r \leq \frac{1}{2M}u'(M)$; and
(4) $f(a,b) = \frac{1}{2}u(a) + \frac{1}{2}u(b) - \frac{r}{2}|u(b)-u(a)|^2$, $0 \leq r \leq \frac{1}{2}$.

Supposing that u is smooth, we find that the first two cases have kinked indifference curves at the diagonal in that a,b-plane, while the second two will accept small fair "bets." One sees this by calculating the gradients and checking whether or not the slopes of the indifference curves match as the (a,b) pair approach equality.

	Gradient if $a < b$	Gradient if $a > b$
(1)	$Df = \begin{pmatrix} \frac{1}{2}(u'(a)+r) \\ \frac{1}{2}(u'(b)-r) \end{pmatrix}$	$Df = \begin{pmatrix} \frac{1}{2}(u'(a)-r) \\ \frac{1}{2}(u'(b)+r) \end{pmatrix}$
(2)	$Df = \begin{pmatrix} \frac{1}{2}u'(a)(1+r) \\ \frac{1}{2}u'(b)(1-r) \end{pmatrix}$	$Df = \begin{pmatrix} \frac{1}{2}u'(a)(1-r) \\ \frac{1}{2}u'(b)(1+r) \end{pmatrix}$
(3)	$Df = \begin{pmatrix} \frac{1}{2}u'(a)+r(b-a) \\ \frac{1}{2}u'(b)-r(b-a) \end{pmatrix}$	$Df = \begin{pmatrix} \frac{1}{2}u'(a)-r(a-b) \\ \frac{1}{2}u'(b)+r(a-b) \end{pmatrix}$
(4)	$Df = \begin{pmatrix} \frac{1}{2}u'(a)+r(u(b)-u(a)) \\ \frac{1}{2}u'(b)-r(u(b)-u(a)) \end{pmatrix}$	$Df = \begin{pmatrix} \frac{1}{2}u'(a)-r(u(a)-u(b)) \\ \frac{1}{2}u'(b)+r(u(a)-u(b)) \end{pmatrix}$

2.3.4. *First Order Stochastic Dominance.* For F_1 and F_2 closed subsets of $[0,M]$, F_1 **dominates** F_2, written $F_1 \succsim_D F_2$, if for all $x_1 \in F_1$, there exists an $x_2 \in F_2$ with $x_2 \leq x_1$, and for all $x_2 \in F_2$, there exists an $x_1 \in F_1$ with $x_2 \leq x_1$. If F_1 and F_2 are the two point sets $\{a,b\}$ and $\{c,d\}$, $a \leq b$, $c \leq d$, $F_1 \succsim_D F_2$ requires $a \geq c$ and $b \geq d$.

Definition 3. *A linear U on $\mathbb{K}(\Delta)$ is* ***set monotonic*** *if* $[F_1 \succsim_D F_2] \Rightarrow [U(A_{F_1}) \geq U(A_{F_2})]$,

When F_1 and F_2 are singleton sets, set monotonicity is equivalent to the monotonicity of the vN-M part of the utility function. When F_1 and F_2 are two point sets, set monotonicity is equivalent to the monotonicity of the function $f(a,b) := U([\![\delta_a, \delta_b]\!])$.

A probability $p \in \Delta$ first order stochastically dominates q, written $p \succsim_{FOSD} q$, if for all consequences $c \in [0,M]$, $p(c,M] \geq q(c,M]$, equivalently, if $F_p(c) \leq F_q(c)$ for all c.

Definition 4. *For $A, B \in \mathbb{K}(\Delta)$, A* ***dominates*** *B, $A \succsim_D B$, if for all $p \in A$, there is a q in B such that $p \succsim_{FOSD} q$ and for all $q \in B$, there is a $p \in A$ such that $p \succsim_{FOSD} q$.*

Note that this is *not* the lattice ordering for sets inherited from the lattice $(\Delta[0,M], \succsim_{FOSD}$), which would define $s = p \vee q$ as having the cdf $F_s(x) = \min\{F_p(x), F_q(x)\}$, $s = p \wedge q$

as having the cdf $F_s(x) = \max\{F_p(x), F_q(x)\}$, and define **strong dominance** by $A \succsim_{SD} B$ if for all $(p, q) \in A \times B$, $p \vee q \in A$ and $p \wedge q \in B$.

Example 3. *Strong dominance implies dominance, but the reverse is not true. If $A \in \mathbb{K}(\Delta)$ is not a lattice subset of Δ, then it is not the case that $A \succsim_{SD} A$ while $A \succsim_D A$ as a triviality since every p satisfies $p \succsim_{FOSD} p$. More substantively, let $A = \{p \in \Delta : p[M/3, M] = 1,\ p$ has a density w.r.t. Lebesgue measure$\}$ and $B = \{q \in \Delta : q[0, 2M/3] = 1\}$. $A \succsim_D B$ because all $p \in A$ stochastically dominate $\delta_{M/3} \in B$ and all $q \in B$ are stochastically dominated by $p = U[2M/3, M]$, the uniform distribution on the interval $[2M/3, M]$.*

It is not the case that $A \succsim_{SD} B$ — if $p = U[2M/3, M] \in A$ so that $F_p(M/2) = \frac{1}{4}$ and $q = \frac{1}{4}\delta_{M/2} + \frac{3}{4}\delta_{2M/3} \in B$, then $p \vee q \notin A$.

Lemma 4. *For closed $F_1, F_2 \subset [0, M]$, $F_1 \succsim_D F_2$ iff $A_{F_1} \succsim_D A_{F_2}$.*

Proof: Consider the points masses in A_{F_1} and A_{F_2}. ∎

Definition 5. *A function U on $\mathbb{K}(\Delta)$* ***respects dominance*** *if $[A \succsim_D B] \Rightarrow [U(A) \geq U(B)]$.*

Theorem 4. *If a linear U respects dominance, then it is set monotonic.*

Proof: If $F_1 \succsim_D F_2$, then $A_{F_1} \succsim_D A_{F_2}$, so that respecting dominance yields $U(A_{F_1}) \geq U(A_{F_2})$, that is, U is set monotonic. ∎

At this point, the theory of linear preferences on $\mathbb{K}(\Delta)$ becomes significantly richer than the theory of linear preferences on Δ. This means that it is correspondingly more difficult to work with in full generality.

Values of a linear L on Δ are determined by the easily identifiable extreme points, $\{\delta_x : x \in [0, M]\}$. Values of a linear $U : \mathbb{K}(\Delta) \to \mathbb{R}$ are determined by its values on the extreme points in the space of sets $\mathbb{K}(\Delta)$. However, not all extreme sets are of the form A_F, F a closed subset of $[0, M]$.

Lemma 5. *$\mathcal{A}_F := \{A_F : F \subset [0, M],\ F$ closed, non-empty$\}$ does not exhaust the class of extreme sets of $\mathbb{K}(\Delta)$.*

Proof: Suppose first that we replace the consequence space $[0, M]$ by a three point set, $\mathbf{C} = \{a, b, c\}$. Here, $\{A_F : \emptyset \neq F \subset \{a, b, c\}\}$ has seven elements, $A_{\{a\}}$, $A_{\{b\}}$, $A_{\{c\}}$, $A_{\{a,b\}}$, and so on. $\mathbb{K}(\mathbf{C})$ is an infinite dimensional compact set, hence, by the Krein-Milman Theorem, must have infinitely many extreme points. More specifically, all elements in the span of $S(\mathbf{C})$ have at most six sides, and all of them are parallel to the outer boundaries of Δ. Such figures stay a strictly positive distance away from B when B is any circular subset of $\Delta(\{a, b, c\})$. All that remains is to show that B is not the limit of any sequence of sets B_n which are in the span of $\mathcal{A}_F$.

For closed sets, F_n, F, $[d(F_n, F) \to 0] \Rightarrow [A_{F_n} \to A_F]$. Since the finite subsets of $[0, M]$ are dense in the closed sets, this shows that any limit of finite convex combinations $\sum_F \alpha_F A_F$ is a limit of finite convex combinations where the F are restricted to be finite.

Suppose now that B_n is a sequence of finite convex combinations of sets of the form $\sum_k \alpha_{k,n} A_{F_{k,n}}$, $F_{k,n}$ finite, and that $B_n \to B$ for the circular $B \subset \Delta(\{a, b, c\})$ described above.

For $\delta > 0$ let $C_\delta \subset [0, M]$ be the δ-ball around $\{a, b, c\}$. For any $\delta > 0$, for large n, $\sum_{\{k,n:F_{k,n} \subset C_\delta\}} \alpha_{k,n} > 1 - \delta$. By taking δ smaller than the minimum distance between a, b, and c, we see that B_n can only converge to something in the span of $S(\{a, b, c\})$, that is, it cannot converge to B. ∎

2.3.5. *Second Order Stochastic Dominance.* A continuous linear utility on probabilities respects second order stochastic dominance iff the associated vN-M utility function is concave. The parallel concept for a linear U on $\mathbb{K}(\Delta)$ is a set version of 'horizontal' concavity.

Treating probabilities as functions from classes of sets to $[0, 1]$, the usual way to take convex combinations of probabilities yields 'vertical' combinations of their cdfs, $F_{\alpha p+(1-\alpha)q}(x) = \alpha F_p(x) + (1 - \alpha)F_q(x)$. Since the domain, $[0, M]$, is convex, one could also take combinations of the cdfs by shifting them horizontally.

Example 4. *For $p = U[a, b]$ and $q = U[c, d]$ being uniform distributions on the indicated intervals, define $\alpha p \oplus_{hz} (1-\alpha)q$ to be the uniform distribution $U[\alpha a+(1-\alpha)c, \alpha b+(1 - \alpha)d]$. Here, for any height $s \in (0, 1)$, we identify the points x_p with $F_p(x_p) = s$ and x_q with $F_q(x_q) = s$ and define the cdf of $\alpha p \oplus_{hz} (1 - \alpha)q$ to take the value s at the point $\alpha x_p + (1 - \alpha)x_q$.*

To generalize the example to probabilities with point masses and flat regions at different heights requires a bit of notation. For a given cdf F, define the correspondence $\widehat{F}(x) = [F(x-), F(x)]$ (which is equal to $\{F(x)\}$ when $p(x) = 0$). For each $s \in [0, 1]$, the lower inverse of $\widehat{F}$ at s is $\widehat{F}^-(s) = \{x \in [0, M] : s \in \widehat{F}(x)\}$ is a non-empty, closed interval subset of $[0, M]$. A cdf F is uniquely identified by the values of the correspondence $\widehat{F}^-(s)$ for $s \in (0, 1)$.

Definition 6. *The* ***horizontal convex combination of*** p ***and*** q *is the distribution $r = \alpha p \oplus_{hz} (1 - \alpha)q$ with the cdf identified by the correspondence $s \mapsto \alpha \widehat{F}_p^-(s) + (1 - \alpha)\widehat{F}_q^-(s)$, $s \in [0, 1]$. For $A, B \in \mathbb{K}(\Delta)$, the* ***horizontal convex combination of*** A ***and*** B *is the set $\alpha A \oplus_{hz} (1 - \alpha)B = \{\alpha p \oplus_{hz} (1 - \alpha)q : p \in A,\ q \in B\}$.*

Vertical and horizontal convex combinations are quite different:

(1) $\frac{1}{2}\delta_0 \oplus_{hz} \frac{1}{2}\delta_1 = \delta_{\frac{1}{2}}$, a point mass, while $\frac{1}{2}\delta_0 + \frac{1}{2}\delta_1$ is the distribution of a Bernoulli($\frac{1}{2}$) random variable;

(2) with $p = \frac{1}{2}\delta_0 + \frac{1}{2}\delta_1$ and $q = U[1,2]$, $\frac{1}{2}p \oplus_{hz} \frac{1}{2}q = \frac{1}{2}U[0, \frac{3}{4}] + \frac{1}{2}U[1\frac{1}{4}, 2]$ while the cdf of $\frac{1}{2}p + \frac{1}{2}q$ has jumps of size $\frac{1}{4}$ at 0 and 1 and is a straight line on the interval $[1,2]$;

(3) $\alpha A_{F_1} \oplus_{hz} (1-\alpha) A_{F_2} = \{p : p(\alpha F_1 + (1-\alpha)F_2) = 1\} = A_{\alpha F_1 + (1-\alpha)F_2}$ is an extreme set while $\alpha A_{F_1} + (1-\alpha)A_{F_2} = \{p : p(F_1) = \alpha,\ p(F_2) = (1-\alpha)\}$ is not an extreme set.

Definition 7. *A linear U on $\mathbb{K}(\Delta)$ is* ***horizontally concave*** *if for all $A, B \in \mathbb{K}(\Delta)$, $U(\alpha A \oplus_{hz} (1-\alpha)B) \geq \alpha U(A) + (1-\alpha)U(B)$.*

If $F = \{x\}$ is a singleton set, then A_F contains only δ_x, point mass on x. Restricted to the singleton sets $F_1 = \{x_1\}$ and $F_2 = \{x_2\}$, the vN-M part of a horizontally concave U has satisfies $u(\alpha x_1 + (1-\alpha)x_2) \geq \alpha u(x_1) + (1-\alpha)u(x_2)$. Slightly more generally, oncavity of $u(\cdot)$ and $v(\cdot)$ in $f(a,b) = \frac{1}{2}u(a) + \frac{1}{2}u(b) - \frac{r}{2}|v(b) - v(a)|^h$ from eqn. (1) gives horizontal concavity on the class of sets of the form A_F, F finite. Continuity and linearity extend this to the span of the class $\{A_F : F \text{ closed}\}$, which includes the parallelipipeds based on point masses.

It would be nice to know if horizontal concavity has implications for yet broader classes of $A, B \in \mathbb{K}(\Delta)$ that are comparable using second order stochastic dominance. In some special cases they are, but general results seem difficult to find. However, horizontal concavity is the essential ingredient for addressing portfolio problems because it requires that convexifying the set of values that a set of probabilities takes on is better than having a lottery over the sets.

2.3.6. *Comparative Ambiguity Aversion.* Working with the Steiner points of convex sets requires that they be finite dimensional. The present analysis focuses, therefore, on the behavior of linear preferences $\mathbb{K}^{fd}$, the dense subset of finite dimensional elements of $\mathbb{K}(\Delta)$. When the space of consequences is finite, $\mathbb{K}^{fd} = \mathbb{K}(\Delta)$, and there is no loss.

The one-sided derivative of U^0 in the direction A at 0 is defined by $D_A^+ U^0(0) := \lim_{r \downarrow 0}(U^0(r \cdot A) - U^0(0))/r$. When there are only two consequences, the derivative in the direction $A = [-\frac{1}{2}, \frac{1}{2}] \in St^{-1}(0)$ is the v in the representation $U([p-r, p+r]) = p - vr$ that appears in Lemma 1.

From Definition 2, linear preferences are ambiguity averse iff for all $A \in St^{-1}(0)$, $D_A^+ U^0(0) \leq 0$ and for some B, $D_B^+ U^0(0) < 0$. Ambiguity seeking preferences reverse the signs, and ambiguity neutrality requires for all $A \in St^{-1}(0)$, $D_A^+ U^0(0) = 0$.

If $U^0(St^{-1}(0)) = \mathbb{R}$, then the preferences prefer increases of some forms of ambiguity and dislike increases of other forms of ambiguity. This involves $D_A^+ U^0(0) > 0$ for some A and $D_B^+ U^0(0) < 0$ for some B. For example, one can have a preference for not knowing the distribution of one's likely time of death, while disliking ambiguity about the distributions associated with different retirement portfolio choices. By the intermediate value theorem, there is an $\alpha \in (0,1)$ such preferences are ambiguity neutral with respect to $\alpha A + (1-\alpha)B$, that is, $D^+_{\alpha A+(1-\alpha)B} U^0(0) = 0$.

The cone $\mathcal{N} = \{A \in St^{-1}(0) : D^+_A U^0(0) = 0\}$ is the set of ambiguous outcomes over which the preferences are neutral. The cone contains only $\{0\}$ when preferences are strictly ambiguity averse or seeking, but is otherwise non-trivial. To compare degrees of ambiguity aversion requires normalization. For any non-trivial linear $U = (U^0, u)$, define $N_U = \max_{A \in \mathbb{K}(\Delta)} U(A) - \min_{B \in \mathbb{K}(\Delta)} U(B)$.

Definition 8. *For any non-trivial pair $U = (U^0, u)$ and $V = (V^0, v)$ of linear utilities, U is* ***more ambiguity averse than*** *V if for all $A \in St^{-1}(0)$, $U^0(A)/N_U \leq V^0(A)/N_V$, equivalently, if $D_A U^0(0)/N_U \leq D_A V^0(0)/N_{U'}$.*

2.3.7. *Risk Equivalents.* Certainty equivalents are a tool that can be used to compare degrees of risk aversion. In a similar fashion, risk equivalents facilitate the comparison of degrees of ambiguity aversion. Rather than carry around the N_U and N_V terms, in this part of the analysis, all linear U's are normalized so that $\min_A U(A) = 0$ and $\max_B U(B) = 1$.

The existence of risk equivalents requires a balance between attitudes towards risk and attitudes toward ambiguity.

Definition 9. *For a risk and ambiguity averse U, the risk-ambiguity tradeoff is* ***balanced*** *if $E\,u(X_A) - u(a) \geq |U^0(A - St(A))|$ for all $a \in [0, M]$ and all $A \subset \Delta[a, M]$ where X_A is a random variable with distribution $St(A)$.*

From the concavity of u, a sufficient condition for balanced tradeoff is that for all $a \in [0, M]$ and all $A \subset [a, M]$, $\frac{1}{M-a}[E\,X_A - a] \geq |U^0(A - St(A))|$.

Lemma 6. *For a risk and ambiguity averse U, the risk-ambiguity tradeoff is balanced iff for all $[a, b] \subset [0, M]$ and for all $A \subset \Delta[a, b]$, $u(a) \leq U(A) \leq u(b)$.*

Proof: Easy. ∎

From this result, one sees that **un**balanced tradeoffs require that a DM facing outcomes limited to an interval $[a, b]$ either: if $U(A) < u(a)$, be willing to pay to reduce their ambiguity about distributions on $[a, b]$ to certainty of the worst outcome, a; or, if $u(b) < U(A)$, be willing to pay to substitute ambiguity about distributions on $[a, b]$ for certainty of the best outcome, b.

If $u(\cdot)$ is concave and the risk-ambiguity tradeoffs are balanced, then $u(0) = 0$ and $u(M) = 1$. This implies that for all $x \in [0, M]$, $u(x) \geq \frac{x}{M}$. The risk premium for a distribution $p \in \Delta$ measure the horizontal distance between the expectation under p and the certainty equivalent of p. The utility premium measures the corresponding vertical distance. Specifically, the distance between the utility of getting the expectation of p for sure and the expected utility of getting a random reward distributed as p.

Definition 10. *For $u : [0, M] \to [0, 1]$, the **utility premium of** $p \in \Delta$ is* $\mathrm{Prem}^u(p) := \int_0^M [u(x) - \frac{x}{M}]\, dp(x)$*. Equivalently, if X_p is a random variable with distribution p, then* $\mathrm{Prem}^u(p) = E\, u(X_p) - \frac{1}{M} E\, X_p$.

Certainty equivalents allow us to reduce risk to certain prospects. Risk equivalents allow us to reduce ambiguity to a risky bet over the best and the worst outcome.

Definition 11. *The **risk equivalent** of $A \in \mathbb{K}(\Delta)$ for the preferences $U = (U^0, u)$ is* $\mathrm{Reqv}_A^U = \rho$ *satisfying* $\rho u(M) + (1-\rho)u(0) = U(A) - \mathrm{Prem}^u(St(A))$.

If the risk-ambiguity tradeoff is balanced, Reqv_A^U is a number in $[0, 1]$.

Theorem 5. *For risk and ambiguity averse U and V with balanced risk-ambiguity tradeoffs, U is more ambiguity averse than V iff for all $A \in \mathbb{K}(\Delta)$,* $\mathrm{Reqv}_A^U \geq \mathrm{Reqv}_A^V$.

Proof: For $A \in \mathbb{K}^{fd}$, let X_A be a random variable with law $St(A)$. By definition, $\mathrm{Reqv}_A^U = U(A) - \mathrm{Prem}^u(St(A)) = U^0(A - St(A)) + E\, u(X_A) - [E\, u(X_A) - \frac{1}{M} E\, X_A]$, and this yields $\mathrm{Reqv}_A^U = U^0(A - St(A)) - \frac{1}{M} E\, X_A$. Applied to V, this yields $\mathrm{Reqv}_A^V = V^0(A - St(A)) - \frac{1}{M} E\, X_A$. Therefore, $\mathrm{Reqv}_A^U \leq \mathrm{Reqv}_A^V$ iff $U^0(A - St(A)) \leq V_0(A - St(A))$. ∎

2.3.8. *Risk Equivalents: the Two Outcome Case.* Another way to phrase the balanced tradeoff assumption is as a kind of betweenness — for $b > a$ and $C \subset \Delta[a, b]$, $U(\{\delta_a\}) \leq U(C) \leq U(\{\delta_b\})$. The following example implies that this does not generalize even to $U(A) < U(B)$ implying that $U(A) \leq U(\mathbf{co}\,(A \cup B)) \leq U(B)$.

Example 5. *If $p \neq q$ are non-degenerate distributions, $A = \{p\}$ and $B = \{q\}$, and $U(A) = U(B)$, then $\mathbf{co}\,(A \cup B) = [\![p, q]\!]$. With strict ambiguity aversion in the direction of the line segment $[\![p, q]\!]$, $U(\mathbf{co}\,(A \cup B)) < U(A) = U(B)$.*

However, this kind of betweenness does appear in the two outcome case. Let $\mathbf{C} = \{w, b\}$ with w the worst outcome and b the best. In this case, the utility premium is always 0 so that the risk equivalent of A is that number $\mathrm{Reqv}_A = \rho \in \mathbb{R}$ that satisfies $\rho U(\{b\}) + (1-\rho)U(\{w\}) = U(A)$. Recall that in the two outcome case, utility can be normalized to $U([p - r, p + r]) = p - vr$.

Theorem 6. *In the normalized two outcome case, the following are equivalent:*

(1) *for all A, $0 \leq \mathrm{Reqv}_A \leq 1$,*
(2) *the risk-ambiguity tradeoff is balanced,*
(3) *for all A, B with $U(A) < U(B)$, $U(A) \leq U(\mathbf{co}\,(A \cup B)) \leq U(B)$, and*

(4) $-1 \leq v \leq 1$.

Proof: The equivalence of the first with $-1 \leq v \leq 1$ is immediate.

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,

$$a \leq U([a,b]) = \frac{a+b}{2} - v\frac{b-a}{2} \leq b. \tag{2}$$

Rearrangement shows that $-1 \leq v \leq +1$ iff for all $0 \leq a \leq b \leq 1$, (2) holds.

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

$$U([a,b]) \geq U(\mathbf{co}\,([a,b] \cup [c,d])) \geq U([c,d]). \tag{3}$$

If $[a,b] \subset [c,d]$ or $[c,d] \subset [a,b]$, both inequalities in (3) 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, $[a,b] \vee [c,d] := \mathbf{co}\,([a,b] \cup [c,d]) = [c,b]$. The first inequality in (3) 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 (3) 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. ∎

2.4. **Comparing Consequentialist and State Space Approaches.** vN-M worked with linear preferences on $\Delta(\mathbf{C})$, representable by

$$V_{vN-M}(p) = \int_{\mathbf{C}} u(c)\, dp(c) \tag{4}$$

for a utility function $u : \mathbf{C} \to \mathbb{R}$.

Savage worked with preferences over mappings, f, from a state space, Ω, to $\mathbf{C}$. Savage's preferences are representable by a probability P on Ω and a utility function u on $\mathbf{C}$, specifically,

$$V_{Savage}(f) = \int_{\Omega} u(f(\omega))\, dP(\omega). \tag{5}$$

Since u depends on ω only through $f(\omega)$, if we define $p = Pf^{-1}$ as the image law of P under the mapping f, we have, by change of variable, $W(f) = V(Pf^{-1})$.

The two formalizations are often interpreted very differently, objectively and subjectively. vN-M have given p's, usually assumed to be related to some notion of the "true" or "objective" distribution. Savage deduces the existence of the P used to evaluate random variables, so that P is subjective, and Pf^{-1} need not be the "true" distribution.

Extending Gilboa and Schmeidler, Ghirardato and Marinacci have "bi-separable" preferences over sets of f's. A leading case is the set of α-MEU preferences. These are specified by an $\alpha \in [0,1]$, a convex set $S \subset \Delta(\Omega)$, and a function $u : \mathbf{C} \to \mathbb{R}$. They are representable by

$$(6) \qquad V_{MEU}(f) = \alpha \min_{P \in S} \int_\Omega u(f(\omega))\, dP(\omega) + (1-\alpha) \max_{P \in S} \int_\Omega u(f(\omega))\, dP(\omega).$$

If $\alpha = 1$, we have Gilboa and Schmeidler's original preferences.

If one is willing to ask that subjective probabilities match objective probabilities, then the change of variable device that makes Savage preferences and vN-M preferences interchangeable makes α-MEU a special case of linear preferences over sets of probabilities:

(1) Since S is convex, $A := \{Pf^{-1} : P \in S\}$ is a convex subset of $\Delta(\mathbf{C})$;
(2) The functions $\min_{p \in A} \int_{\mathbf{C}} u(c)\, dp(c)$ and $\max_{p \in A} \int_{\mathbf{C}} u(c)\, dp(c)$ are continuous and linear in the set A.

Define the u-utility possibility set of $A \subset \Delta(\mathbf{C})$ as the interval

$$(7) \qquad I_u(A) = [\min_{p \in A} \textstyle\int_{\mathbf{C}} u(c)\, dp(c), \max_{p \in A} \int_{\mathbf{C}} u(c)\, dp(c)] \subset \mathbb{R}.$$

Representing α-MEU preferences in terms of the utility functions considered here requires that $I_u(A) = I_u(B)$ iff $A \sim B$.

With $I_u(A) = [u', u'']$, define $c_u = \frac{1}{2}u' + \frac{1}{2}u''$ as the center of the utility interval, and define $r_u = \frac{1}{2}(u'' - u')$ as the radius of the utility interval. Direct calculation yields

$$(8) \qquad \alpha u' + (1-\alpha)u'' = c_u + (1-2\alpha) r_u.$$

Here, $\alpha > \frac{1}{2}$ corresponds to ambiguity aversion, and shrinking A towards its centerpoint only matters through the shrinking of r_u. Put differently, that the only aspect of the spread of an ambiguous set of probabilities that matters is how big its spread is in the u-direction.

For $S \subset \Delta(\Omega)$ one can define the unambiguous events as the set $\{E : \forall P, Q \in S : P(E) = Q(E)\}$. Such classes of events are not, generally, fields. A random variable, $f : \Omega \to \mathbf{C}$ is unambiguous if for all $P, Q \in S$, $Pf^{-1} = Qf^{-1}$. It is a subtle and difficult issue to go from assumptions on a preference ordering over f's to comparisons of risk and ambiguity attitudes. It is far easier here, in the vN-M style framework, because one assumes that, as part of the problem, one is given the description of the sets of probabilities. class of unambiguous events,

3. Application(s)

The first bit is a treatment of ambiguous equilibria in non-cooperative games, it parallels Crawford [**?**]. After this, there are a number of themes under development: interpretations of ambiguity aversion as the reduced form of more complicated dynamic games; interpretations of ambiguous equilibria as disequilibrium points along some learning dynamic; relations to the subjective, state-space approaches to ambiguity preferences, some of which can be interpretationally subsumed in the present approach, some of which cannot.

3.1. **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.

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.

3.1.1. *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 2. *$K^* = (K_i^*)_{i\in I} \in \times_{i\in I}\mathbb{K}_i$ 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 7. *$K^* = \{\sigma^*\}$ is an equilibrium for $\Gamma_{Amb}(u, U)$ iff σ^* 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.[2] Some examples may make this clearer.

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

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

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

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

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.

[2] I think that this is where the real tension between equilibrium and ambiguity arises.

3.1.3. *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 $0 < s < \frac{1}{2}$,

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

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 worldy 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	$(7,1)$	$(4,0)$

3.1.4. *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 π, firm m will violate the worker safety laws if

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

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.[3] 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.

3.1.5. *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.

3.2. **Incomplete Learning.** One interpretation of ambiguity is that it refers to situations in which we do not know the distribution, that is, we have not yet learned the distribution. Ambiguity aversion may represent some reactions to some forms of this ...

3.3. **Ambiguity Aversion as a Reaction to a Hostile Universe.** Suppose that one is considering an unfamiliar trade. It is a bad signal if the person offering the trade is unwilling to be specific about what is involved.[4] In this way, ambiguity aversion can capture utility in a situation where one believes that the rest of the universe is hostile [8]. An extreme version of the hostile universe story appears in *minimax* statistics, where one evaluates a statistical estimator by supposing that after the estimator is picked, Nature picks the distribution that maximizes the statisticians loss from making an error.

[3] There are horror stories about the small size of monitoring budgets for worker safety in the US.

[4] It is an even worse signal if they are trying to baffle you with bullshit.

3.4. **Efficient Allocations of Risk and Uncertainty.** When risk aversion and ambiguity aversion go hand in hand, efficient allocations of risk and uncertainty are, essentially, allocations of risk with vN-M utility functions modified to be more risk averse.

Two people have a random endowment, either x_i or y_i, so that $\mathbf{x} = (x_1, x_2)$ or $\mathbf{y} = (y_1, y_2)$, in $\mathbb{R}^2_+$. The goal is to transfer the allocation between the two so as to efficiently allocate the risk and uncertainty. A transfer is a function (t_x, t_y), and the resulting allocations are $(x_1 + t_x, y_1 + t_y)$ for the first person, and $(x_1 - t_x, y_2 - t_y)$ for the second. Person i's U_i is given by $U_i([\![\delta_a, \delta_b]\!]) = \frac{1}{2}u_i(a) + \frac{1}{2}u_i(b) - \frac{r}{2}|v_i(b) - v_i(a)|^h$ where u_i and v_i are increasing concave functions on $[0, M]$, and u_i is normalized by $u_i(0) = 0$, $u_i(M) = 1$.

The efficient, individually rational allocations are the solutions to

$$\max_{(t_x, t_y)} \tag{9}$$

When there is both risk and uncertainty to be allocated, the efficient solutions that arise are the same as the ones that arise in the efficient allocations ...

3.5. **Market Completeness and Demand for Options.** Building on [15], [16], we have ...

References

[1] Allais, M. 1953, Le Comportment de l'Homme Rationnel Devant le Risque: Critique des Postulates 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] Ellsberg, D. 1961, Risk, Ambiguity, and the Savage Axioms, *Quarterly Journal of Economics* **75**(4), 643-669.

[6] Epstein, L. 1999, A Definition of Uncertainty Aversion. *Review of Economic Studies* **66**, 579-608.

[7] Epstein, L. and J. Zhang, 2001, Subjective Probabilities on Subjectively Unambiguous Events. *Econometrica* **69**, 265-306.

[8] Fellner, W. 1961, Distortion of Subjective Probabilities as a Reaction to Uncertainty, *Quarterly Journal of Economics* **75**(4), 670-689.

[9] Fenchel, W. and Jessen, B. 1938, Mengenfunktionen und konvexe Körper. *Danske Vid. Selskab. Mat.-fis. Medd.* **16**, 3.

[10] Ghirardato, P. Marinacci, M. 2001, Range Convexity and Ambiguity Averse Preferences. *Economic Theory* **17**(3), 599-617.

[11] Karni, E., Schmeidler, D., Vind, K. 1983, On State Dependent Preferences and Subjective Probabilities. *Econometrica* **51**(4), 1021-1032.

[12] Knight, F. H. 1921, *Risk, uncertainty and profit.* Chicago, University of Chicago Press (1971, reprint of the 1921 ed.).
[13] Körner, R. 1995, A Variance of Compact Convex Random Sets. Working paper, Institut fr Stochastik, Bernhard-von-Cotta-Str. 2 09599 Freiberg.
[14] Machina, M. 1982, 'Expected Utility' Analysis Without the Independence Axiom. *Econometrica* **50**, 277-323.
[15] Mukerji, S. 1998, Ambiguity Aversion and Incompleteness of Contractual Form. *American Economic Review* **88**(5), 1207-31.
[16] Mukerji S.; Tallon J-M. 2001, Ambiguity Aversion and Incompleteness of Financial Markets. *The Review of Economic Studies* **68**(4), 883-904.
[17] Von Neumann, J. and Morgenstern, O. 1944, *Theory of games and economic behavior.* New York, Wiley, (1953, 3d ed.).
[18] Røadström, H. 1952, An Imbedding Theorem for Spaces of Convex Sets, *Proceedings of the American Mathematical Society* **3**(1), 165-169.
[19] Savage, L. J. 1954, *The foundations of statistics.* New York, Wiley.
[20] Schneider, R. 1993, *Convex Bodies: The Brunn-Minkowski Theory*, Cambridge, Cambridge University Press.
[21] 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