

Copyright
by
Martin Dumav
2011

The Report Committee for Martin Dumav
certifies that this is the approved version of the following report:

**The von Neumann/Morgenstern Approach to
Ambiguity**

APPROVED BY

SUPERVISING COMMITTEE:

Gordan Zitkovic, Supervisor

Mihai Sirbu

**The von Neumann/Morgenstern Approach to
Ambiguity**

by

Martin Dumav, B.S.; M.A.; M.S.

REPORT

Presented to the Faculty of the Graduate School of
The University of Texas at Austin
in Partial Fulfillment
of the Requirements
for the Degree of

MASTER OF ARTS

THE UNIVERSITY OF TEXAS AT AUSTIN

August 2011

Acknowledgments

Many thanks to Svetlana Boyarchenko, Takashi Hayashi, Mark Machina, Marcin Peski, and Dale Stahl for help with this paper. Special thanks to Maxwell Stinchcombe whose encouragement and patience has been invaluable. They should be held blameless.

The von Neumann/Morgenstern Approach to Ambiguity

Martin Dumav, M.A.
The University of Texas at Austin, 2011

Supervisor: Gordan Zitkovic

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 uncertain if the decision maker is choosing between distributions, and is ambiguous if the decision maker is choosing between sets of probabilities. The von Neumann/Morgenstern approach to uncertain choice problems uses a continuous linear function on probabilities. This paper develops the theory of ambiguous choice problems as a continuous, linear functions on closed convex *sets* of probabilities. This delivers: a framework encompassing most of the extant ambiguity averse preferences; a complete separation of attitudes towards risk and attitudes toward ambiguity; and generalizations of first and second order stochastic dominance rankings to ambiguous decision problem. Quasi-concave preferences on sets that satisfy a restricted betweenness property capture variational preferences.

Table of Contents

| | |
|--|-----------|
| Acknowledgments | iv |
| Abstract | v |
| Chapter 1. Theory | 1 |
| 1.1 Introduction | 1 |
| 1.1.1 Change of Variables in Risky Decision Problems | 2 |
| 1.1.2 Change of Variables in Ambiguous Decision Problems | 2 |
| 1.1.3 Continuous, Linear Functions on $\Delta_{\mathbf{X}}$ and $\mathbb{K}_{\Delta_{\mathbf{X}}}$ | 3 |
| 1.1.4 Domain Equivalence | 4 |
| 1.1.5 Nonlinear Theories | 5 |
| 1.1.6 Outline | 6 |
| 1.2 Two Consequences | 6 |
| 1.2.1 Urns and Interval Sets of Probabilities | 6 |
| 1.2.2 Representation, Separation, and Dominance | 8 |
| 1.2.3 Domain Equivalence | 11 |
| 1.2.4 Nonlinearities | 12 |
| 1.2.4.1 All closed sets | 12 |
| 1.2.4.2 All closed convex sets | 12 |
| 1.3 Representations of vNM Preferences | 12 |
| 1.3.1 Notation and Definitions | 13 |
| 1.3.2 Representations and Properties of vNM Preferences | 15 |
| 1.3.3 Separation of Risk and Ambiguity Attitudes | 22 |
| 1.3.3.1 Centers plus sets centered at 0 | 22 |
| 1.3.3.2 Via the representation theorem | 23 |
| 1.3.4 First and Second Order Dominance | 24 |
| 1.3.5 Domain Equivalence | 25 |
| 1.4 Finite Consequence Spaces | 29 |

| | | |
|-------------------|---|-----------|
| 1.4.1 | The Steiner Basis | 29 |
| 1.4.2 | Separating Risk and Ambiguity Attitudes | 32 |
| 1.4.3 | Extreme Set Representations | 33 |
| 1.4.4 | The Choice of a Center | 36 |
| 1.5 | Separating Risk and Ambiguity | 37 |
| 1.6 | First and Second Order Stochastic Dominance | 38 |
| Chapter 2. | Applications | 40 |
| 2.1 | Applications | 40 |
| 2.1.1 | First and Second Order Dominance, Diversification | 40 |
| 2.1.1.1 | A Functional Form | 40 |
| 2.1.1.2 | First Order Stochastic Dominance | 42 |
| 2.1.1.3 | Second Order Stochastic Dominance | 46 |
| | Bibliography | 49 |
| | Vita | 54 |

Chapter 1

Theory

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 [17])

1.1 Introduction

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. A risky decision problem is one in which the decision maker (DM) knows the probability distributions associated with their choices, an ambiguous decision problem is one in which the DM knows only the sets of probability distributions associated with their choices. Under study are ambiguous decision problems in which the DM's preferences are continuous linear functions on the class of sets of distributions over consequences.

1.1.1 Change of Variables in Risky Decision Problems

Savage [36] provides an axiomatic study of preferences over random variables, f, g , mapping a state space, Ω , to consequences, \mathbf{X} . Preferences satisfying his axioms can be represented by $f \succ g$ iff $\int_{\Omega} u(X(\omega)) dP(\omega) > \int_{\Omega} u(g(\omega)) dP(\omega)$ for a probability P on Ω and a (usually) continuous utility function $u : \mathbf{X} \rightarrow \mathbb{R}$. The change of variable involves defining $p = f(P)$ by $p(E) = P(f^{-1}(E))$ and $q = g(P)$ by $q(E) = P(g^{-1}(E))$ for $E \subset \mathbf{X}$. The change of variable theorem implies that we have $p \succ q$ iff $\int_{\mathbf{X}} u(x) dp(x) > \int_{\mathbf{X}} u(x) dq(x)$. This delivers preferences over $\Delta_{\mathbf{X}}$, distributions on consequences, which is the von Neumann-Morgenstern (vNM) framework.

Savage's axioms on preferences over random variables must deliver both the probability P and the utility function u . Because it does not need to do this double duty, the axiomatic treatment of vNM preferences is much simpler than Savage's treatment, reducing to assumptions of continuity and linearity on $\Delta_{\mathbf{X}}$. This paper performs the same change of variable for the theory of choice under ambiguity delivering the parallel simplification.

1.1.2 Change of Variables in Ambiguous Decision Problems

The first and most widely known model of ambiguous decision problems is due to Gilboa and Schmeidler [24]. Preferences satisfying their weakening Savage's axioms can be represented by $f \succ g$ iff $\min_{P \in S} \int_{\Omega} u(f(\omega)) dP(\omega) > \min_{P \in S} \int_{\Omega} u(g(\omega)) dP(\omega)$ for S a weakly closed, convex set of probabilities on Ω . More generally, Ghirardato, Maccheroni, and Marinacci [22] axiomatize

“biseparable” preferences representable by $f \succ g$ iff

$$\begin{aligned} \alpha \cdot \min_{P \in S} \int_{\Omega} u(f(\omega)) dP(\omega) + (1 - \alpha) \cdot \max_{Q \in S} \int_{\Omega} u(f(\omega)) dQ(\omega) > \\ \alpha \cdot \min_{P \in S} \int_{\Omega} u(g(\omega)) dP(\omega) + (1 - \alpha) \cdot \max_{Q \in S} \int_{\Omega} u(g(\omega)) dQ(\omega) \end{aligned} \quad (1.1)$$

where S is again a weakly closed, convex set of probabilities on Ω .

The change of variable delivers $A \succ B$ iff

$$\begin{aligned} \alpha \cdot \min_{p \in A} \int_{\mathbf{X}} u(x) dp(x) + (1 - \alpha) \cdot \max_{q \in A} \int_{\mathbf{X}} u(x) dq(a) > \\ \alpha \cdot \min_{p \in B} \int_{\mathbf{X}} u(x) dp(x) + (1 - \alpha) \cdot \max_{q \in B} \int_{\mathbf{X}} u(x) dq(a). \end{aligned} \quad (1.2)$$

Note that $A \mapsto \alpha \cdot \min_{p \in A} \int_{\mathbf{X}} u(x) dp(x) + (1 - \alpha) \cdot \max_{q \in A} \int_{\mathbf{X}} u(x) dq(a)$ is a continuous linear function on $\mathbb{K}_{\Delta_{\mathbf{X}}}$, the set of convex subsets of $\Delta_{\mathbf{X}}$.

1.1.3 Continuous, Linear Functions on $\Delta_{\mathbf{X}}$ and $\mathbb{K}_{\Delta_{\mathbf{X}}}$

The vNM weak* continuous linear functions on $\Delta_{\mathbf{X}}$ have an integral representation, delivering the utility function u . From properties of u , one derives properties of preferences, e.g. monotonicity and respect for first order dominance, monotonicity and concavity and respect for second order dominance. We develop the parallel integral representation result for weak* continuous linear functions on $\mathbb{K}_{\Delta_{\mathbf{X}}}$. Properties of this representation deliver a complete treatment of first and second order stochastic dominance in ambiguous decision problems as well as a complete separation between attitudes toward risk and attitudes toward ambiguity.

This separation result is somewhat surprising — it allows for the combination of any attitude toward risk to be combined with any attitude toward ambiguity. By contrast, in the preferences over random variables setting, comparisons between attitudes toward ambiguity are limited to comparisons between decision makers with the same attitude toward risk. The explanation lies in a comparison of how much is given and how much derived in the two approaches.

If one works with preferences over random variables to study ambiguous decision problems, from preferences one must not only derive which events are ambiguous, i.e. which events have multiple probabilities in the set S above, and which are risky, and to simultaneously identify attitudes toward both ambiguity and risk. By contrast, in the vNM approach one starts with sets of probabilities, including the singleton sets, and preferences over these sets of probabilities. From linear functions on sets, one needs only separate how the preferences act on the singleton sets, which gives the attitude toward risk, from how they act on the non-singleton sets, which gives the attitude toward ambiguity.

1.1.4 Domain Equivalence

In the random variable approach to ambiguous choice, preferences are typically taken to be state independent, that is, they only depend on the set of distributions over the consequences \mathbf{X} . This means that in modeling a decision maker facing any specific choice problem, one has many choices for

the model of the state space Ω and for the set of random variables over which the decision maker is picking. For risky problems, one will typically take the probability P on space Ω to be non-atomic. This yields a universality property for the model — all distributions, p , on any space of consequences, \mathbf{X} , in a wide class, e.g. any complete separable metric space, are of the form $f(P)$ for some random variable f . Modeling with random variables and modeling with induced distributions give the same domain for risky choice problems, the set of problems that can be modeled with the two approaches is the same.

We prove a parallel domain equivalence result for ambiguous choice problems: if the state space, Ω , supports a countably additive non-atomic distribution, then there is a universal closed convex set S of probabilities on Ω with the property that any (non-empty) closed convex set $A \subset \Delta_{\mathbf{X}}$ is of the form $X(S)$ for some random variable X .

The domain equivalence result has an additional implication — for state independent preferences in the random variable approach to ambiguous decision problems, there is no loss in assuming that there is only one set S of probabilities, and that set can only be identified up to having the universality property.

1.1.5 Nonlinear Theories

Machina [29] extended the range of vNM preferences from linear-in-probabilities to locally linear preferences, e.g. Dekel's [11] betweenness preferences. Some recently studied preferences for choice under ambiguity are not

linear-in-sets but locally linear. For example the variational preferences are quasi-concave, hence locally linear at most points.

1.1.6 Outline

The next section covers the main results of the paper in the case that there are just two consequences, $\#\mathbf{X} = 2$. The subsequent section covers the basic theory for the case of compact metric spaces of consequences: the representation theorem; domain equivalence; and stochastic dominance. The following section specializes to finite spaces, where the general results often have a more intuitive form. The penultimate section covers two extensions of the theory: the non-linear extensions; and the extension to more general spaces of consequences. The last section concludes and sketches applications that may be simpler with the vNM approach than with the random variable approach.

1.2 Two Consequences

Urn problems are a particularly clear and compelling way to explain the intuitions for preferences in the presence of ambiguity, and that is where we begin.

1.2.1 Urns and Interval Sets of Probabilities

An 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 preferences in experiments are

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

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

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

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

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

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 B&G 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 a probability p determines the chance that they win to knowing that the probability belongs to an interval with p at its center.

1.2.2 Representation, Separation, and Dominance

In this urn problem as in all two consequence problems, we take $\mathbf{X} = \{0, 1\}$, and $\Delta_{\mathbf{X}} \subset \mathbb{R}^{\{0,1\}}$ can be represented by $[0, 1]$ where $q \in [0, 1]$ corresponds to a probability q of receiving 1, the better outcome. Let \mathfrak{K} be the class of non-empty closed, convex subsets of the probabilities $[0, 1]$, that is, $\mathfrak{K} = \{[a, b] : 0 \leq a \leq b \leq 1\}$. In this case, our general representation theorem for continuous linear functions on convex sets of probabilities reduces to being of the form $U([a, b]) = u_1 a + u_2 b$ for $u_1, u_2 \in \mathbb{R}$.

An interval $[a, b]$ first order stochastically dominates another one, $[a', b']$ if $a \geq a'$ and $b \geq b'$. For the utility function $U(\cdot)$ to respect first order dominance,¹ we must have $U([a, b]) \geq U([a', b'])$ for all $a \geq a'$ and $b \geq b'$. This is equivalent to $u_1, u_2 \geq 0$, and non-triviality of the preferences requires at

¹With only two consequences, we cannot treat second order dominance here.

least one inequality strict, which we normalize with $u_1 + u_2 = 1$.

Restricted to singleton sets of probabilities, U is a vNM utility function on $\{0, 1\}$. Since intervals with no width correspond to risky choices, the normalization gives $U([p, p]) = u_1 p + u_2 p = p$, e.g. $U([0, 0]) = 0$ and $U([1, 1]) = 1$. From this, the vNM utility function on $\mathbf{X} = \{0, 1\}$ is $u(0) = 0$ and $u(1) = 1$, which leads to the connection to Ghirardato, Maccheroni, and Marinacci's [22] α -minmaxEU preferences (see eqn. (1.1) above). These are preferences over sets of probabilities given by $U(A) = \alpha \cdot (\min_{p \in A} \int u dp) + (1 - \alpha) \cdot (\max_{q \in A} \int u dq)$. In the present context, the utility of an interval $[a, b]$ is

$$U([a, b]) = u_1 \cdot \left(\min_{p \in [a, b]} \int_{\mathbf{X}} u(x) dp(x) \right) + u_2 \cdot \left(\max_{q \in [a, b]} \int_{\mathbf{X}} u(x) dq(x) \right),$$

where $\mathbf{X} = \{0, 1\}$ and u_1 plays the role of α .

In the α -minmaxEU preferences, $\alpha > \frac{1}{2}$, here $u_1 > u_2$, corresponds to ambiguity aversion, and a change of basis in $\mathfrak{K}(\Delta(\{0, 1\}))$ allows us to see why this should be so. A set $[a, b]$ can be viewed as $[p - r, p + r]$ where $p = (a + b)/2$ and $r = (b - a)/2$. With the normalization, $U([p - r, p + r]) = (u_1 + u_2)p - (u_1 - u_2)r$, conveniently re-written as $U([p - r, p + r]) = p - vr$. Here, $v = (u_1 - u_2) > 0$ corresponds to disliking expansions of the set of probabilities about the center p , corresponding to ambiguity aversion, and we say that the preferences are **ambiguity averse** if $v > 0$, **ambiguity neutral** if $v = 0$, and **ambiguity loving** if $v < 0$.

It is notable that by working with preferences over $\mathfrak{K}(\Delta(\{0, 1\}))$, we have completely separated the attitude toward risk, the p part of the util-

ity function, and the attitude toward ambiguity, the vr part of the utility function.² Given the extraordinary lengths to which one must go to achieve even partial separation between attitudes towards risk and ambiguity in the preferences over random variables framework, it is perhaps puzzling that the separation is so immediate. The reason is that the question is simpler here: in the vNM approach, one distinguishes between attitudes toward singleton sets of probabilities and *given* larger sets of probabilities; in the Savage approach, one must, from preferences over random variables, deduce both which random variables are ambiguous and compare risk and ambiguity.

After the normalization, the property that $|v| \leq 1$ corresponds to preferences that respect first order stochastic dominance, which also has an informative alternate interpretation in the ambiguous choice situations under study here. The tradeoff between risk and ambiguity seems unbalanced if knowing that one faces the worst risky option in a set is preferred to facing the set, $\{p - r\} \succ [p - r, p + r]$, or if the set is preferred to the best risky option in the set $[p - r, p + r] \succ \{p + r\}$. In the two consequence case, the risk-ambiguity tradeoff is **balanced** iff $|v| \leq 1$, i.e. iff the preferences respect first order stochastic dominance.

²The parameter v measures the tradeoff between risk and ambiguity. It can be elicited by giving people a choice between risky and ambiguous urns.

1.2.3 Domain Equivalence

For most modeling of random variables, one can take the probability space to be the unit interval with the uniform distribution, λ . This is because every probability distribution, p , on a wide class of spaces (including every complete separable metric space) is the image measure, $X_p(\lambda)$, for an appropriately chosen random variable, X_p . This is a domain equivalence result, it means that one can study random phenomena by studying distributions or by studying random variables, the choice is a matter of convenience. Further, the unit interval can be replaced by any probability space that supports a countably additive non-atomic distribution.

In the case that the consequence space is $\mathbf{X} = \{0, 1\}$, one can explicitly give a closed convex set S of probability distributions on $[0, 1]$ with the property that for every $A \in \mathfrak{R}(\Delta(\{0, 1\}))$, there is a random variable X_A such that $X_A(S) = A$. This is a domain equivalence result that says that for modeling choice problems under ambiguity, one can either work with random variables or with sets of distributions. The set S can be taken to be the set of distributions having density (with respect to λ) equal to $2 - x$ on $[0, \frac{1}{2}]$ and equal to x on $(\frac{1}{2}, 1]$, $0 \leq x \leq 2$. A random variable $X(\omega) = 1_{[0,r)}(\omega) + 1_{[\frac{1}{2}, \frac{1}{2}+s)}(\omega)$, $r, s \leq \frac{1}{2}$, has the property that $P_x(X = 1) = r \cdot (2 - x) + s \cdot x = 2r + (s - r)x$. For $r \geq s$, this gives the interval $[2s, 2r]$, for $r < s$, the interval $[2r, 2s]$.

1.2.4 Nonlinearities

One can consider non-linearities in preferences over all closed convex subsets of $\Delta_{\mathbf{X}}$ or over all closed subsets of $\Delta_{\mathbf{X}}$.

1.2.4.1 All closed sets

Continuity and linearity of preferences means that there is no loss in restricting preferences to the closed convex subsets of $\Delta(\{0, 1\})$. Continuity means that if the (Hausdorff) distance between two sets is 0, then they are indifferent, and the distance between a set and its closure is 0. Linearity means that, taking A to be any closed subset of $\Delta(\{0, 1\})$, $U(\frac{1}{2}A + \frac{1}{2}A) = U(A) = U(\sum_{i \leq n} \frac{1}{n}A)$. Since $\sum_{i \leq n} \frac{1}{n}A \rightarrow \text{co}(A)$ where $\text{co}(A)$ is the convex hull of A , we have $U(A) = U(\text{co}(A))$. By contrast, quasi-concavity, as in variational preferences over random variables, would lead to the conclusion that $\text{co}(A) \succeq A$,

1.2.4.2 All closed convex sets

Preferences over $\mathfrak{R}(\Delta(\mathbf{X}))$ are quasi-concave if $[A \sim B] \Rightarrow [\alpha A + (1 - \alpha)B] \succeq A$. In the special case that $\mathbf{X} = \{0, 1\}$, the corresponds to being less sensitive to ambiguity as the center of the interval increases.

1.3 Representations of vNM Preferences

We develop the theory when the space of consequences, \mathbf{X} is a compact metric space. Generalizations to more general spaces of consequences are

indicated in the appendix.

1.3.1 Notation and Definitions

\mathbf{X} is a compact metric space of consequences, with metric $d(\cdot, \cdot)$, open balls $B_\epsilon(x) = \{y \in \mathbf{X} : d(x, y) < \epsilon\}$. The Borel σ -field, \mathcal{X} , is the smallest σ -field of subsets of \mathbf{X} containing all of the open balls $B_\epsilon(x)$. $C_{\mathbf{X}} (= C_b(\mathbf{X}))$ is the set of continuous (necessarily bounded) \mathbb{R} -valued functions on \mathbf{X} with the sup norm and the associated Borel σ -field.

$\Delta_{\mathbf{X}}$ is the set of countably additive probabilities on \mathcal{X} with the weak* topology. Letting K be a compact, convex, balanced subset of the unit ball in $C_{\mathbf{X}}$ with the property that $E_K := \{\lambda f : f \in K, \lambda \geq 0\}$ is dense in $C_{\mathbf{X}}$, the weak* topology can be metrized by

$$\rho_K(p, q) = \sup_{f \in K} |\langle f, p \rangle - \langle f, q \rangle|. \quad (1.3)$$

With the metric ρ_K , $\Delta_{\mathbf{X}}$ is also a compact metric space.

We will often regard $\Delta_{\mathbf{X}}$ as a subset of $\mathbf{ca}(\mathbf{X})$, the set of finite, signed, countably additive measures on \mathcal{X} . The Riesz representation theorem tells us that $C_{\mathbf{X}}$ is the predual of $\mathbf{ca}(\mathbf{X})$, i.e. $\mathbf{ca}(\mathbf{X}) = C_{\mathbf{X}}^*$. This implies that a linear $\ell : \Delta \rightarrow \mathbb{R}$ is ρ_K -continuous (i.e. weak*-continuous) iff it has a representation as $\ell(p) = \int_{\mathbf{X}} u(x) dp(x)$ for some $u \in C_{\mathbf{X}}$, and the u is unique. Properties of ℓ are therefore determined by the simpler object u .

$\mathbb{K}_{\Delta_{\mathbf{X}}}$ denotes the set of non-empty closed convex subsets of $\Delta_{\mathbf{X}}$ with

the weak* Hausdorff metric,

$$d_H(A, B) = \inf\{\epsilon > 0 : A \subset B^\epsilon, B \subset A^\epsilon\}.$$

With the metric d_H , $\mathbb{K}_{\Delta_{\mathbf{X}}}$ is also a compact metric space. As needed, we regard $\Delta_{\mathbf{X}}$ as a closed convex subset of $\mathbf{ca}(\mathbf{X})$, the set of finite, signed, countably additive measures on \mathbf{X} with the weak* topology. For $x \in \mathbf{X}$, point mass on x is the element of $\Delta_{\mathbf{X}}$ defined and by denoted $\delta_x(E) = 1_E(x)$.

For $p, q \in \mathbf{ca}(\mathbf{X})$ and $\alpha, \beta \in \mathbb{R}$, $\alpha p + \beta q \in \mathbf{ca}(\mathbf{X})$ is defined by $(\alpha p + \beta q)(E) = \alpha p(E) + \beta q(E)$, and (Minkowski) addition is defined by $\alpha A + \beta B := \{\alpha p + \beta q : p \in A, q \in B\}$. If $\beta = 1 - \alpha$ and $A, B \in \mathbb{K}_{\Delta_{\mathbf{X}}}$, then $\alpha A + \beta B \in \mathfrak{K}(\Delta)$.

For each $A \in \mathbb{K}_{\Delta_{\mathbf{X}}}$ and $f \in C_{\mathbf{X}}$, define $\bar{\mu}_A(f) = \max_{p \in A} \langle f, p \rangle$ and $\underline{\mu}_A(f) = \min_{p \in A} \langle f, p \rangle$. The functions $\bar{\mu}_A(\cdot)$ and $\underline{\mu}_A(\cdot)$ have Lipschitz constant 1, and since K is compact and the $\bar{\mu}_A$ are concave, the lattice version of the Stone-Weierstrass theorem tells us that $\overline{\text{span}}(\{\bar{\mu}_A, \underline{\mu}_A : A \in \mathfrak{K}(\mathbf{ca}(\mathbf{X}))\})$ is dense in $C(K)$. Further, $\max_{f \in K} |\bar{\mu}_A(f) - \bar{\mu}_B(f)| = d_H(A, B)$, so that the mapping $A \leftrightarrow \bar{\mu}_A$ is a linear isometry between $\mathbb{K}_{\Delta_{\mathbf{X}}}$ and $C(K)$ with the sup norm.

$\partial U_{\mathbf{X}}$ denotes the boundary of the unit ball in $C_{\mathbf{X}}$. With the sup norm metric, it is a complete separable metric space. It also has useful geometric properties.

Definition 1. A subset E of a vector space \mathfrak{X} is **radial** if for all $x \in E$, $x \neq 0$, $r \cdot x \in E$ iff $r = 1$. A function $\psi : \mathfrak{X} \rightarrow \mathbb{R}$ is **radial** iff for $x \in \mathfrak{X}$ and all $r \geq 0$, $\psi(r \cdot x) = r \cdot \psi(x)$.

Note that subsets of $\partial U_{\mathbf{X}}$ are radial sets, for all $A \in \mathbb{K}_{\Delta_{\mathbf{X}}}$, the functions $\bar{\mu}_A$ and $\underline{\mu}_A$ are radial, and that all radial functions satisfy $\psi(0) = 0$.

The functions $\bar{\mu}_A$ and $\underline{\mu}_A$ have Lipschitz constant 1, and Lipschitz functions play a large role in the development. Let $\text{Lip}(\partial U_{\mathbf{X}})$ denote the bounded Lipschitz functions on $\partial U_{\mathbf{X}}$ with the bounded Lipschitz norm given by $\|\psi\|_{BL} := \max\{\|\psi\|_{\infty}, \sup_{f \neq g} \frac{|\psi(f) - \psi(g)|}{\|f - g\|_{\infty}}\}$. With the supnorm, $\partial U_{\mathbf{X}}$ is a metric space with diameter 2, adjoining the 0 function gives us a “pointed” metric space, $(\partial U_{\mathbf{X}} \cup \{0\})$ with the supnorm. There is a canonical identification of bounded Lipschitz functions on $\partial U_{\mathbf{X}}$ and the Lipschitz functions on $(\partial U_{\mathbf{X}} \cup \{0\})$: for $\psi \in \text{Lip}(\partial U_{\mathbf{X}})$, define $\hat{\psi} : \partial U_{\mathbf{X}} \cup \{0\} \rightarrow \mathbb{R}$ by $\hat{\psi}(f) = \psi(f)$ if $f \neq 0$ and $\hat{\psi}(0) = 0$. Let $\text{Lip}_0(\partial U_{\mathbf{X}}) = \{\hat{\psi} : \psi \in \text{Lip}(\partial U_{\mathbf{X}})\}$ with the norm $\|\hat{\psi}\|_{\text{Lip}_0} = \sup_{f \neq g} \frac{|\hat{\psi}(f) - \hat{\psi}(g)|}{\|f - g\|_{\infty}}$. The mapping $\psi \leftrightarrow \hat{\psi}$ is an isometric isomorphism, and it identifies the two Banach spaces $(\text{Lip}(\partial U_{\mathbf{X}}), \|\cdot\|_{BL})$ and $(\text{Lip}_0(\partial U_{\mathbf{X}}), \|\cdot\|_{\text{Lip}_0})$.

1.3.2 Representations and Properties of vNM Preferences

A weak* continuous rational preference relation on $\mathbb{K}_{\Delta_{\mathbf{X}}}$ is a complete, transitive relation, \succeq , on $\mathbb{K}_{\Delta_{\mathbf{X}}}$ such that for all $B \in \mathbb{K}_{\Delta_{\mathbf{X}}}$, the sets $\{A : A \succ B\}$ and $\{A : B \succ A\}$ are open. We will always assume that preferences are continuous, and except when discussing deviations from linearity, we will also assume that the following holds.

Axiom 1 (Independence). For all $A, B, C \in \mathbb{K}_{\Delta_{\mathbf{X}}}$ and all $\alpha \in (0, 1)$, $A \succeq B$ if and only if $\alpha A + (1 - \alpha)C \succeq \alpha B + (1 - \alpha)C$.

The proof of the following is an easy adaptation of standard arguments.

Theorem 1.3.1. *A continuous rational preference relation on $\mathbb{K}_{\Delta_{\mathbf{X}}}$ satisfies the Independence Axiom if and only if it can be represented by a continuous linear function.*

We now turn to an integral representation of continuous linear functions on $\mathbb{K}_{\Delta_{\mathbf{X}}}$.

Theorem 1.3.2. *For every $\eta \in \mathbf{ca}(\partial U_{\mathbf{X}})$, the functional $L_{\eta} : \text{Lip}(\partial U_{\mathbf{X}}) \rightarrow \mathbb{R}$ defined by $L_{\eta}(\psi) = \int_{\partial U_{\mathbf{X}}} \psi(f) d\eta(f)$ is in the predual of $\text{Lip}(\partial U_{\mathbf{X}})$. Further, if $\eta' \neq \eta''$ in $\mathbf{ca}(\partial U_{\mathbf{X}})$, then for some $A \in \mathbb{K}_{\Delta_{\mathbf{X}}}$, $L_{\eta'}(\mu_A) \neq L_{\eta''}(\mu_A)$.*

Some comments are appropriate.

1. The proof of Theorem 1.3.2 proceeds by noting that the μ_A are Lipschitz functions on $\partial U_{\mathbf{X}}$, that the L_{η} belong to the closure of a predual of $\text{Lip}(\partial U_{\mathbf{X}})$, and that the $\bar{\mu}_A$ and $\underline{\mu}_A$ are a determining class of Lipschitz functions for the L_{η} 's. According to Weaver [42, §2.2, p. 40], whether or not the predual of Lip (or Lip_0) is unique up to isometric isomorphism is not known. Thus, it is possible that there are additional continuous linear functionals not contained of the form L_{η} .
1. $\mu_A(f) := \max_{p \in A} \langle p, f \rangle$ defines the support function for $A \in \mathbb{K}_{\Delta_{\mathbf{X}}}$, so this result is a direct generalization of the representation theorem for the case that \mathbf{X} is a finite space.

2. We believe that the class of L_η 's exhausts the set of continuous linear functions, but have been unable to prove this.
3. The measure η has a Hahn-Jordan decomposition $\eta = \eta_+ - \eta_-$ where η_+ and η_- are non-negative measures with disjoint carriers. Since $\max_{p \in A} \langle p, f \rangle = -\min_{p \in A} \langle p, -f \rangle$ for all $f \in \partial U$, L_η can be re-written as

$$L_\eta(A) = \int_{\partial U} \min_{p \in A} \langle p, f \rangle d\eta_{\min}(f) + \int_{\partial U} \max_{q \in A} \langle p, g \rangle d\eta_{\max}(g) \quad (1.4)$$

where $\eta_{\min} = \eta_-$ and $\eta_{\max} = \eta_+$.

4. Since η_{\min} and η_{\max} are both non-negative and at least one of them is non-null when L is non-trivial, the normalization $\eta_{\min}(\partial U) + \eta_{\max}(\partial U) = 1$ is harmless, and maintained from now on.
5. Taking $\eta_{\min} = \alpha \cdot \delta_u$ and $\eta_{\max} = (1 - \alpha) \cdot \delta_u$ delivers the α -minmaxEU preferences as a special case.

Proof of Theorem 1.3.2. Let FinSupp denote the set of finitely supported measures on $\partial U_{\mathbf{X}}$, i.e. $\text{FinSupp} = \{\sum_{i=1}^n \beta_i \delta_{f_i} : n \in \mathbb{N}, \beta_i \in \mathbb{R}, f_i \in \partial U_{\mathbf{X}}, \delta_{f_i}(E) = 1_E(f_i)\}$. From Weaver [42, Theorem 2.2.2 and Corollary 2.2.3(a), p. 39-40], $\text{Lip}(\partial U_{\mathbf{X}})$ is isometrically isomorphic to FinSupp^* , the dual of FinSupp when FinSupp has the dual norm $\|\eta\|_{\text{Lip}^*} := \sup\{|\langle \psi, \eta \rangle| : \|\psi\|_{\text{Lip}} \leq 1\}$. As a Banach space, the predual of $\text{Lip}(\partial U_{\mathbf{X}})$ is a closed, hence complete, vector subspace of its second dual, $\text{Lip}(\partial U_{\mathbf{X}})^*$. Therefore, the predual of $\text{Lip}(\partial U_{\mathbf{X}})$ is the completion of FinSupp in the norm $\|\cdot\|_{\text{Lip}^*}$. Standard weak* approximation of

measures by finitely supported measures immediately imply that FinSupp is $\|\cdot\|_{\text{Lip}^*}$ -dense in $\mathbf{ca}(\partial U_{\mathbf{X}})$, showing that every L_η belongs to the predual.

The uniqueness proof is somewhat more involved. Suppose that for some $\eta' \neq \eta''$, $L_{\eta' - \eta''}(\mu_A) \equiv 0$. Let $\eta = \eta' - \eta''$. We will show, through a series of Lemmas, that $L_\eta(\mu_A) \equiv 0$ implies that $\eta = 0$, contradicting $\eta' \neq \eta''$.

Lemma 1.3.3. *If $\eta \in \mathbf{ca}(\partial U_{\mathbf{X}})$, then there exists a countable sequence of disjoint compact sets K_n with the property that $\sum_n |\eta|(K_n) = |\eta|(\partial U_{\mathbf{X}})$, $|\eta|(K_1) > \frac{1}{2}|\eta|(\partial U_{\mathbf{X}})$, and $|\eta|(K_{n+1}) > \frac{1}{2}|\eta|(\partial U_{\mathbf{X}} \setminus (\cup_{i=1}^n K_i))$.*

Proof. Since $\partial U_{\mathbf{X}}$ is a complete separable metric space, $|\eta|$ is tight, that is, for every $\epsilon > 0$, there exists a compact K_ϵ such that $|\eta|(K_\epsilon) \geq (1 - \epsilon)|\eta|(\partial U_{\mathbf{X}})$. Pick K_1 for $\epsilon = \frac{1}{2}$, pick $K_2 \subset (\partial U_{\mathbf{X}} \setminus K_1)$ using the tightness of $|\eta|$ restricted to $(\partial U_{\mathbf{X}} \setminus K_1)$, and proceed inductively. \square

Lemma 1.3.4. *If K_n is a sequence of disjoint, compact subsets of $\partial U_{\mathbf{X}}$ and $r_n \downarrow 0$ a strictly decreasing sequence, then $F := \text{cl}(\cup_n r_n \cdot K_n) = \cup_n r_n \cdot K_n \cup \{0\}$, and F is compact and radial.*

Proof. The disjointness of the K_n imply that F is radial. To prove that $\text{cl}(\cup_n r_n \cdot K_n) = \cup_n r_n \cdot K_n \cup \{0\}$, it is sufficient to show that $\cup_n r_n \cdot K_n \cup \{0\}$ is compact, which completes the proof. If f_k is a sequence in $\cup_n r_n \cdot K_n \cup \{0\}$, then either each $r_n \cdot K_n$ contains at most finitely many f_k , or for some n' , infinitely many f_k belong to $r_{n'} \cdot K_{n'}$. In the first case, $f_k \rightarrow 0$, in the second, f_k has a convergent subsequence in $r_{n'} K_{n'}$. \square

Lemma 1.3.5. *For any $\eta \in \mathbf{ca}(\partial U_{\mathbf{X}})$ there exists a compact radial set F and an $\eta' \in \mathbf{ca}(F)$ such that $\int_{\partial U_{\mathbf{X}}} \psi(f) d\eta(f) = \int_F \psi(f) d\eta'(f)$ for all radial functions ψ . Further, $\eta' = 0$ iff $\eta = 0$, and F can be chosen so that the span of $\{\bar{\mu}_A, \underline{\mu}_A : A \in Kdx\}$ is dense in \mathcal{C}_0 , the set of continuous functions on F satisfying $\psi(0) = 0$.*

Proof. If $|\eta| = 0$, there is nothing to prove. Otherwise, we normalize η by dividing it by $|\eta|(\partial U_{\mathbf{X}})$ so that $|\eta|$ is a probability. From Lemma 1.3.3, we can pick a disjoint sequence K_n such that $|\eta|(K_n) < 1/2^n$. Define $K'_n = \frac{1}{\log(n+1)} \cdot K_n$. By Lemma 1.3.4, $F := \cup_n K'_n$ is compact and radial. For $E \subset F$, define $\eta'(E) = \sum_n \eta(\log(n+1) \cdot (E \cap K'_n))$ so that $|\eta'|(F) = \sum_n \frac{\log(n+1)}{2^n} < \infty$, implying that $\eta' \in \mathbf{ca}(F)$. By construction, $\eta' = 0$ iff $\eta = 0$. If ψ is a radial function on $C_{\mathbf{X}}$, then the equality $\int_{\partial U_{\mathbf{X}}} \psi(f) d\eta(f) = \int_F \psi(f) d\eta'(f)$ is immediate any ψ which is a simple function on $\partial U_{\mathbf{X}}$, and dominated convergence completes the argument.

For the second part of the proof, we appeal to the following well-known Corollary of the Stone-Weierstrass Theorem:

If F is a compact Hausdorff space with at least two points and L is a lattice in $C(F)$, then a function ψ in $C(F)$ belongs to the closure of L iff for each pair of distinct points x and y in F and for each $\epsilon > 0$, there exists some f in L for which $|f(x) - \psi(x)| < \epsilon$ and $|f(y) - \psi(y)| < \epsilon$.

The span of $\{\bar{\mu}_A, \underline{\mu}_A : A \in \mathbb{K}_{\Delta_{\mathbf{x}}}\}$ is a vector subspace of \mathcal{C}_0 and is a lattice. Let ψ be an arbitrary element of \mathcal{C}_0 , and let f and g be arbitrary points in F . If one of them, say f , is equal to 0, then for any $A \in \mathbb{K}_{\Delta_{\mathbf{x}}}$ and some $\beta = \psi(g)/\bar{\mu}_A(g) \in \mathbb{R}$, $\beta\bar{\mu}_A(f) = \psi(f) = 0$ and $\beta\bar{\mu}_A(g) = \psi(g)$. If neither f nor g is 0, then they belong to K'_{n_f} and K'_{n_g} respectively. Restricted to the compact $K'_{n_f} \cup K'_{n_g}$, the regular vector lattice version of the Stone-Weierstrass theorem implies that there exists f in the span with $|f(x) - \psi(x)| < \epsilon$ and $|f(y) - \psi(y)| < \epsilon$. \square

Lemma 1.3.6. *If M is a metric space, x a point in M , \mathcal{G}_x is a sup norm dense subset of $\mathcal{C}_x := \{f \in C_b(M) : f(x) = 0\}$, and $\eta \in \mathbf{ca}(M)$ has the property that $|\eta|(\{x\}) = 0$, then $\eta = 0$ iff for all $g \in \mathcal{G}_x$, $\langle g, \eta \rangle = 0$.*

Proof. If $\eta = 0$, then for all $f \in C_b(M)$, $\int f d\eta = 0$ and $\mathcal{G}_x \subset C_b(M)$.

Now suppose that for all $g \in \mathcal{G}_x$, $\int g d\eta = 0$ but that $\eta \neq 0$. Pick an arbitrary $\epsilon > 0$. We will show that there exists a $\delta' > 0$ such that for all $\delta \in (0, \delta')$, $|\eta|(M \setminus B_\delta(x)) < \epsilon$ where $B_\delta(x)$ is the δ -ball around x . By countable additivity, this implies that for all $\epsilon > 0$, $|\eta|(M \setminus \{x\}) \leq \epsilon$. Since ϵ was arbitrary, this in turn implies that $|\eta|(M \setminus \{x\}) = 0$. Since $|\eta|(\{x\}) = 0$, we have a contradiction.

Since $|\eta|(\{x\}) = 0$, countable additivity implies that $|\eta|(B_\delta(x)) \downarrow 0$ as $\delta \downarrow 0$. Pick δ' such that $|\eta|(B_{\delta'}(x)) < \epsilon/3$. Pick arbitrary $\delta \in (0, \delta')$. The basic extension theorems for continuous functions on closed sets tell us that for any $f \in C_b(M)$, not necessarily in \mathcal{C}_x , with $\|f\|_\infty = 1$, there exists $\hat{f} \in \mathcal{C}_x$ that

agrees with f on the closed set $M \setminus B_\delta(x)$ and has $\|\hat{f}\|_\infty \leq 1$. Pick $g \in \mathfrak{G}_x$ such that $\|g - \hat{f}\|_\infty < \epsilon/3|\eta|(M)$. By choice of δ' and g ,

$$\left| \int g d\eta - \int_{M \setminus B_\delta(x)} f d\eta \right| \leq \quad (1.5)$$

$$\left| \int f d\eta - \int_{M \setminus B_\delta(x)} f d\eta \right| + \left| \int \hat{f} d\eta - \int f d\eta \right| \left| \int g d\eta - \int \hat{f} d\eta \right| < \epsilon/3 + \epsilon/3 + \epsilon/3. \quad (1.6)$$

Since $\int g d\eta = 0$, we know that for all $f \in C_b(M)$ with $\|f\|_\infty \leq 1$, $\left| \int_{M \setminus B_\delta(x)} f d\eta \right| < \epsilon$. This directly implies that $|\eta|(M \setminus B_\delta(x)) < \epsilon$. \square

Returning now to the proof of uniqueness, suppose that $L_\eta(\bar{\mu}_A) \equiv 0$. We must show that $\eta = 0$. By Lemmas 1.3.3, Lemma 1.3.4, and the first part of Lemma 1.3.5, we know that there is an η' on a compact, radial set F such that $\int_F \bar{\mu}_A(f') d\eta'(f') = \int_{\partial U_{\mathbf{X}}} \bar{\mu}_A(f) d\eta(f)$ and $\int_F \underline{\mu}_A(f') d\eta'(f') = \int_{\partial U_{\mathbf{X}}} \underline{\mu}_A(f) d\eta(f)$ for all $A \in \mathbb{K}_{\Delta_{\mathbf{X}}}$. By the second part of Lemma 1.3.5 and Lemma 1.3.6, we know that this implies that $\eta' = 0$. But Lemma 1.3.5 tells us that $\eta' = 0$ iff $\eta = 0$. \square

Recall from (1.3) that we let K be a compact, convex, balanced subset of the closed unit ball in $C_{\mathbf{X}}$ with the property that $\overline{\text{span}}(K) = C_{\mathbf{X}}$. For $r, s \in \mathbf{ca}(\mathbf{X})$, we define the metric $\rho_K(r, s) = \sup_{f \in K} |\langle f, r \rangle - \langle f, s \rangle|$ and the corresponding open balls $B_K(r, \epsilon) = \{s : \rho_K(r, s) < \epsilon\}$. Because $\overline{\text{span}}(K) = C_{\mathbf{X}}$, ρ_K is a metric, and it is easy to establish that for a norm bounded sequence r_n , $\rho_K(r_n, r) \rightarrow 0$ iff r_n weak* converges to r , and this implies that $(\Delta_{\mathbf{X}}, \rho_K)$ is compact. The corresponding (Hausdorff) metric on $\mathfrak{K}(\mathbf{ca}(\mathbf{X}))$ is $d_H(A, B) =$

$\inf\{\epsilon > 0 : A \subset B^\epsilon, B \subset A^\epsilon\}$ where $C^\epsilon := \cup_{r \in C} B_K(r, \epsilon)$ is the open ball around $C \subset \mathbf{ca}(\mathbf{X})$.

Every continuous function on the compact set $\mathbb{K}_{\Delta_{\mathbf{X}}}$ is uniformly continuous. Further, $\mathfrak{K}(\mathbf{ca}(\mathbf{X})) = \mathbb{R}_+ \cdot \mathbb{K}_{\Delta_{\mathbf{X}}} - \mathbb{R}_+ \cdot \mathbb{K}_{\Delta_{\mathbf{X}}}$. Therefore, every continuous linear function on $\mathbb{K}_{\Delta_{\mathbf{X}}}$ has a uniformly continuous extension to $\mathfrak{K}(\mathbf{ca}(\mathbf{X}))$. Since $\overline{\text{span}}(\{\mu_A : A \in \mathfrak{K}(\mathbf{ca}(\mathbf{X}))\})$ is dense in $C(K)$, the continuous linear functions on $\mathbb{K}_{\Delta_{\mathbf{X}}}$ and the continuous linear functions on $C(K)$ are identical. This prepares us for the basic representation theorem.

Let \mathcal{L} denote the set of continuous linear functions on $\mathbb{K}_{\Delta_{\mathbf{X}}}$, let $\text{extr}(K)$ denote the set of extreme points in K , and for $\eta \in \mathbf{ca}(\text{extr}(K))$, define the function $L_\eta : \mathbb{K}_{\Delta_{\mathbf{X}}} \rightarrow \mathbb{R}$ by

$$L_\eta(A) = \int_{\text{extr}(K)} \max_{p \in A} \langle p, f \rangle d\eta(f) = \int_{\text{extr}(K)} \bar{\mu}_A(f) d\eta(f). \quad (1.7)$$

1.3.3 Separation of Risk and Ambiguity Attitudes

There are two ways to understand how continuous linear preferences on sets delivers a complete separation of attitudes towards risk and attitudes toward ambiguity. The first expresses sets as sum of their center plus a set centered at 0, the second works through the representation theorem just given.

1.3.3.1 Centers plus sets centered at 0

As noted above in the two outcome case, a set $[a, b]$ can be expressed as $\{p\} + [-r, +r]$ where $p = (a + b)/2$, $r = (b - a)/2$. This expresses every

element of $\mathfrak{R}(\Delta(\{0, 1\}))$ in terms of a singleton set and a set centered at 0. The continuous linear L can be expressed as its action on the singleton sets plus its action on the sets centered at 0. The action of L on singleton sets is continuous and linear, hence is a vNM utility function, u , for risky choice situations. The function u can be replaced by any function v and added back to the action of L on the sets centered at 0. A detailed look at this approach can be found in the section on finite consequence spaces.

1.3.3.2 Via the representation theorem

Restricting L to the singleton sets gives a particular vNM utility function. Specifically, $L(\{p\}) = \int_{\partial U_{\mathbf{x}}} \langle p, f \rangle d\eta_{\min}(f) + \int_{\partial U_{\mathbf{x}}} \langle p, g \rangle d\eta_{\max}(g)$ for all p . This immediately implies that $L(\{p\}) = \int_{\mathbf{X}} u(x) dp(x)$ where $u = \int_{\partial U_{\mathbf{x}}} f d\eta(f)$, $\eta := \eta_{\min} + \eta_{\max}$. In other words, the vNM part of L corresponds to the average of the functions distributed according to η .

By shifting the distributions η_{\min} and η_{\max} by $v - u$, we arrive at continuous linear preferences with vNM preferences given by v , but the rest of the properties of the preferences have not changed.³ In this way, any attitude toward risk can be combined with any attitude toward ambiguity.

³Shifting the distributions η_{\min} and η_{\max} by $v - u$ means “convoluting the distributions with a point mass on the function $v - u$.”

1.3.4 First and Second Order Dominance

In expected utility analysis with the consequence space being an interval subset of \mathbb{R} , e.g. $\mathbf{X} = [0, M]$, first and second order stochastic dominance play a central role. For $p, q \in \Delta([0, M])$: p **first order dominates** q , written $p \succeq_F q$, if for all non-decreasing $u \in C([0, M])$, $\langle u, p \rangle \geq \langle u, q \rangle$; and p **second order dominates** q , written $p \succeq_S q$, if $p \succeq_F q$ and for all non-decreasing concave $u \in C([0, M])$, $\langle u, p \rangle \geq \langle u, q \rangle$. Thus $p \succeq_F q$ iff every expected utility maximizer with monotonic preferences prefers p to q , and $p \succeq_S q$ if in addition, every risk-averse expected utility maximizer with monotonic preferences prefers p to q . Expected utility (vNM) preferences on Δ **respect first order dominance** if $[p \succeq_F q] \Rightarrow [p \succeq q]$, and this is equivalent to the expected utility function, u , being non-decreasing.

Definition 1. For sets $A, B \in \mathfrak{R}(\Delta([0, M]))$, we say that A **first (resp. second) order dominates** B , written $A \succeq_F B$ (resp. $A \succeq_S B$), if for all non-decreasing (resp. non-decreasing and concave) $u \in C([0, M])$,

$$\max_{p \in A} \langle u, p \rangle \geq \max_{q \in B} \langle u, q \rangle, \text{ and } \min_{p \in A} \langle u, p \rangle \geq \min_{q \in B} \langle u, q \rangle. \quad (1.8)$$

vNM preferences on $\mathfrak{R}(\Delta([0, M]))$ **respect first (resp. second) order dominance** if $[A \succeq_F B] \Rightarrow [L(A) \geq L(B)]$ (resp. $[A \succeq_S B] \Rightarrow [L(A) \geq L(B)]$).

Theorem 1.3.7. *A continuous linear preference relation on $\mathfrak{R}(\Delta([0, M]))$ respects first order stochastic dominance iff η_{\min} and η_{\max} are supported on the non-decreasing elements of $\partial U_{\mathbf{X}}$, and supports second order stochastic domi-*

nance iff η_{\min} and η_{\max} are supported on the non-decreasing concave elements of $\partial U_{\mathbf{x}}$,

1.3.5 Domain Equivalence

We now show that one can study either sets of distributions on a state space and random variables or sets of distributions on spaces of consequences.

Let (Ω, \mathcal{F}) be a measure space, and let $\Delta(\mathcal{F})$ denote the set of countably additive probabilities on \mathcal{F} .

Definition 2. A set $S \subset \Delta(\mathcal{F})$ is **universal** if for all complete separable metric spaces (M, d) with Borel σ -field \mathcal{M} , and all closed convex $A \in \mathfrak{K}(\Delta(\mathcal{M}))$, there exists a measurable $X_A : \Omega \rightarrow M$ such that $X_A(S) = A$.

Theorem 1.3.8. *If $\Delta(\mathcal{F})$ contains a countably additive non-atomic probability, then there exists a universal $S \subset \Delta(\mathcal{F})$.*

There are two implications of this result provided that the state space supports a countably additive non-atomic probability: first, for any of the common spaces of consequences, M , preferences over $\mathfrak{K}(\Delta(M))$ are equivalent to probabilistically sophisticated preferences over (measurable) functions $f : \Omega \rightarrow M$; second, when preferences are probabilistically sophisticated and depend on the image measures $f(T)$ for some $T \subset \Delta(\mathcal{F})$, we can replace T by a universal S and suffer no loss of situations that can be modeled.

Proof. We will first prove the result in the special case that (Ω, \mathcal{F}) is $([0, 1]^{\mathbb{N}}, \mathcal{B}^{\mathbb{N}})$, the countable product of the unit interval with the product topology and the

corresponding Borel (product) σ -field, and (M, \mathcal{M}) is $([0, 1], \mathcal{B})$, the unit interval with the (usual) Borel σ -field. The general case will follow easily from this. The typical element of Ω is denoted $\omega = (\omega_1, \omega_2, \dots) = (\omega_n)_{n \in \mathbb{N}}$.

The proof of the special case has two steps: (A) produces a set $S \subset \Delta(\mathcal{F})$ with $K(S) := \{X(S) : \text{such that } X : \Omega \rightarrow M \text{ is measurable}\}$ dense in $\mathfrak{K}(\Delta(M))$, and (B) shows that $K(S)$ is closed.

(A) Let λ denote the uniform distribution on $[0, 1]$ and let $\lambda^\infty = *_{n \in \mathbb{N}} \lambda$ in $\Delta(\mathcal{F})$ be the countable convolution of λ with itself, i.e. the distribution on (Ω, \mathcal{F}) defined by the property that the canonical projections, $(\text{projection}_n)_{n \in \mathbb{N}}$, are i.i.d. uniformly distributed random variables. Let $\{A_n : n \in \mathbb{N}\}$ be a dense subset of $\mathfrak{K}(\Delta(\mathcal{M}))$ with each A_n being the convex hull of finitely many distributions having continuous densities with respect to $\lambda \in \Delta(\mathcal{M})$.

Our first step is to specify a set S and a set of $X_n : \Omega \rightarrow M$ such that $X_n(S) = A_n$ for each $n \in \mathbb{N}$. We begin with the X_n and use them to define S .

- (a) Specifying the X_n : For each n , define $X_n : \Omega \rightarrow M$ by $X_n(\omega) = \omega_n$ (these are the canonical projection maps).
- (b) Specifying S : Let \mathcal{F}_n denote the minimal sub- σ -field of \mathcal{F} making X_n measurable. For a sequence $\nu_n \in \Delta(\mathcal{F}_n)$, define the convolution $*_{n \in \mathbb{N}} \nu_n \in \Delta(\mathcal{F})$ to be the probability agreeing with ν_n on each \mathcal{F}_n and having the random variables $\{X_n : n \in \mathbb{N}\}$ independent. For

each sequence $\mu = (\mu_n)_{n \in \mathbb{N}} \in \times_n A_n$, define the sequence $(\nu_n)_{n \in \mathbb{N}}$ by $\nu_n(\mu_n) = X_n^{-1}(\mu_n) \in \Delta(\mathcal{F}_n)$, and define $\nu(\mu)$ as $*_{n \in \mathbb{N}} \nu_n(\mu_n)$. Let $S' = \{\nu(\mu) : \mu \in \times_n A_n\}$ and let $S = \overline{\text{co}}(S')$ denote its closed convex hull.⁴

It is immediate that for each $n \in \mathbb{N}$, $X_n(S) = A_n$, which implies that $K(S) := \{X(S) : \text{such that } X : \Omega \rightarrow M \text{ is measurable}\}$ is dense in $\mathfrak{K}(\Delta(M))$. Using this observation, we now show that $K(S) = \mathfrak{K}(\Delta(M))$.

(B) Pick arbitrary $A \in \mathfrak{K}(\Delta(M))$ and let $A_n = X_n(S)$ be a sequence converging to A . It is sufficient to show that there exists a measurable $X : \Omega \rightarrow M$ such that for all $\nu \in S$, $X_{n'}(\nu) \rightarrow X(\nu)$, for some subsequence n' , and this is what we now do.

Each measurable $X : \Omega \rightarrow [0, 1]$ *a fortiori* belongs to $L^2 := L^2(\Omega, \mathcal{F}, \lambda^\infty)$, and the set of $[0, 1]$ -valued measurable X is sequentially compact in the weak topology on L^2 (Alaoglu's theorem). Therefore we can pick a subsequence $X_{n'}$ and a $[0, 1]$ -valued measurable X such that for all $Y \in L^2$, $\int_\Omega X_{n'} Y d\nu \rightarrow \int_\Omega X Y d\nu$. By the choice of the A_n , each $\nu \in S$ has a continuous density, ψ_ν , with respect to λ^∞ . Since $[0, 1]^\infty$ is compact, ψ_ν is bounded, hence belongs to L^2 . Therefore, for all $E \in \mathcal{F}$, $\int X_{n'} \psi_\nu 1_E d\lambda^\infty \rightarrow \int X \psi_\nu 1_E d\lambda^\infty = \int_E X d\nu$. Taking E 's of the form $X^{-1}([a, b])$, $0 \leq a \leq b \leq 1$, shows that $X_{n'}(\nu) \rightarrow X(\nu)$.

⁴As we are closing a convex set, the closure can be taken in the weak* topology, the norm topology, or any intermediate topology, e.g. the weak topology.

We now turn to deducing the general case from the special case. We first show that (C), we can replace $[0, 1]$ by an arbitrary complete separable metric space (M, d) , then (D) turn to showing that we can replace $([0, 1]^\infty, \mathcal{B})$ by

(C) If $M = \{x_n : n \in \mathbb{N}\}$ is countably infinite, partition $[0, 1]$ into non-degenerate intervals $\{[a_n, b_n] : n \in \mathbb{N}\}$ and define $\varphi : [a_n, b_n] \rightarrow x_n$. φ is one-to-one and constant on the intervals. For any $\mu \in \Delta(M)$, define $p(\mu) \in \Delta([0, 1])$ by $p(\{a_n\}) = \mu(\{x_n\})$ for all $n \in \mathbb{N}$. For any $A \in \mathfrak{K}(\Delta(M))$, $p(A) \in \mathfrak{K}(\Delta([0, 1]))$. By the special case, $(\varphi \circ X_{p(A)})(S) = A$. The case of finite M is entirely similar. Finally, suppose that M is a complete separable metric space. By the Borel isomorphism theorem, there exists a $\varphi : [0, 1] \leftrightarrow M$ which is one-to-one, onto, and bi-measurable, i.e. which has both φ and φ^{-1} measurable. For each $\mu \in \Delta(M)$, define $p(\mu) = \varphi^{-1}(\mu)$. By the special case, $(\varphi \circ X_{p(A)})(S) = A$.

(D) Now suppose that (Ω, \mathcal{F}) is a measure space and that P is a countably additive, non-atomic element of $\Delta(\mathcal{F})$. Because $[0, 1]$ and $[0, 1]^\mathbb{N}$ are both uncountable, complete separable metric spaces, there exists $\varphi : [0, 1] \leftrightarrow [0, 1]^\mathbb{N}$ which is one-to-one, onto, and bi-measurable. Let q be the countably additive non-atomic probability $\varphi^{-1}(\lambda^\infty)$. It is easy to construct a measurable $\Psi : \Omega \rightarrow [0, 1]$ such that $\Psi(P) = q$. Let T denote the set of countably additive, extensions of $\{\Psi^{-1}(\nu) : \nu \in S\}$. For any $A \in \mathfrak{K}(\Delta(M))$, $(X_A \circ \Psi)(T) = A$. \square

One implication of Theorem 1.3.8 is that all of the modeling of ambiguous choice situations can be done using a single set S of probability distributions. A further consequence is that for any universal S , the set of S -unambiguous events contains a σ -field on which each $\nu \in S$ is non-atomic. To see why, let $M = [0, 1]$ and $A = \{\lambda\}$, and pick an $X_A : \Omega \rightarrow M$ such that $X_A(S) = A$. On the σ -field $X_A^{-1}(\mathcal{B})$, every $\nu \in S$ must agree, and since λ is non-atomic, so is each ν .

1.4 Finite Consequence Spaces

Suppose that \mathbf{X} has ℓ elements, so that $\Delta(\mathbf{X})$ is $\ell - 1$ -dimensional. As in the two consequence case, we translate by a $p \in \Delta$ so that $\Delta(\mathbf{X})$ is a subset of the $\ell - 1$ dimensional vector subspace perpendicular to the vector of 1's, also denoted $\mathbb{R}^{\ell-1}$. $\mathfrak{K}(\mathbb{R}^{\ell-1})$ is the cone of non-empty, compact, convex subsets of $\mathbb{R}^{\ell-1}$. $\mathfrak{K}(\Delta) \subset \mathfrak{K}(\mathbb{R}^{\ell-1})$ is the set of non-empty, compact, convex subsets of Δ , with typical elements A, B, C, D .

1.4.1 The Steiner Basis

In the two outcome case, we made a change of basis to give a direct sum decomposition of every $[q, s] \in \mathfrak{K}(\mathbb{R}^1)$ 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. This expresses every set as a center plus a set centered at 0. The key to giving a similar direct sum decomposition with larger consequence spaces is a continuous linear definition of the center of a set.

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

Definition 2. The **Steiner point** of $A \in \mathfrak{K}(\mathbb{R}^{\ell-1})$ is the vector-valued integral

$$St(A) = \int_{S^{\ell-1}} \arg_A(\mathbf{v}) d\lambda(\mathbf{v}).$$

For example, if the distances between point masses on distinct consequences are equal, then $St(\Delta)$ is the distribution putting equal mass on all elements of \mathbf{X} because each vertex of Δ is $\arg_{\Delta}(u)$ for $1/\#\mathbf{X}$ of the surface of $S^{\ell-1}$.

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 $A \mapsto St(A)$. For a proof of the following, see e.g. [38, Theorem 3.4.2, p. 167].

Theorem 1.4.1 (Steiner). *The mapping $A \mapsto St(A)$ is the unique continuous linear selection that commutes with rigid motions.*⁵

The barycenter/center-of-gravity is not continuous when dimensionality changes.

⁵Replacing λ with any full support probability, ν , on $S^{\ell-1}$ would give a continuous linear selection. However, combining continuity and linearity with the rotation invariance of λ shows that Steiner point commutes with rigid motions, that is, if $R : \mathbb{R}^n \rightarrow \mathbb{R}^n$ is a rigid motion, then $R(St(A)) = St(R(A))$.

Example 1. Suppose there are three consequences, $\mathbf{X} = \{a, b, c\}$. 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, $St(A_\epsilon) \rightarrow St(A_0)$ 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 $\epsilon \downarrow 0$.

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

Lemma 1.4.2. *$St^{-1}(0) \oplus \mathbb{R}^{\ell-1}$ is a direct sum decomposition of the cone $\mathfrak{K}(\mathbb{R}^{\ell-1})$ 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, $(A - St(A))$.

1.4.2 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 vNM utility function. This shows that linear preferences are the linear extension of expected utility preferences.

Theorem 1.4.3. *If $U : \mathfrak{K}(\Delta) \rightarrow \mathbb{R}$ 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 $u \in \mathbb{R}^{\ell-1}$.*

The function U^0 will typically pay attention to both the shape and the orientation of the set $A - St(A)$. For example, even if $d(r, s) = d(r', s')$ and $c := \frac{1}{2}(r + s) = \frac{1}{2}(r' + s')$, one expects that $U^0([r, s]) \neq U^0([r', s'])$ if the orientation of the two line segments $[r, s]$ and $[r', s']$ differ.

Proof. If U is the restriction of a linear $U : \mathfrak{K}(\mathbb{R}^{\ell-1}) \rightarrow \mathbb{R}$, 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 $\mathfrak{K}(\Delta)$ iff it is the restriction of a linear U on $\mathfrak{K}(\mathbb{R}^{\ell-1})$. This follows from the observation that every $A \in \mathfrak{K}(\mathbb{R}^{\ell-1})$ is of the form $A = r \cdot (B - St(B)) + St(A)$ for some $B \in \mathfrak{K}(\Delta)$ and $r \geq 0$. \square

One implication of this result is that any linear U^0 can be combined with any vNM u . This allows complete separation of attitudes towards ambiguity, encoded in U^0 , and attitudes toward risk, encoded in u . Attitudes towards ambiguity are reflected in attitudes toward expansions and contractions of a set around its center.

Definition 3. Preferences on $\mathfrak{R}(\Delta)$ represented by a continuous 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 loving** if $U^0(St^{-1}(0)) = \mathbb{R}_+$.

Continuous linear $U_1 = (U_1^0, u_1)$ is more ambiguity averse than $U_2 = (U_2^0, u_2)$ if, after normalizing both so that $U_1(\mathfrak{R}(\Delta(\mathbf{X}))) = U_2(\mathfrak{R}(\Delta(\mathbf{X}))) = [0, 1]$, $U_1^0(A - St(A)) \leq U_2^0(A - St(A))$ for all A .

Preferences over random variables are probabilistically sophisticated if they depend only on the distribution(s) over consequences that they induce. The present consequentialist approach starts from this assumption. As a result, Epstein's [16] and Ghirardato and Marinacci's [21] definitions of ambiguity aversion agree. As we will see later, both definitions of ambiguity aversion match the present one in the consequentialist setting, but either definition of comparative ambiguity aversion requires identical attitudes toward risk.

1.4.3 Extreme Set Representations

A vNM utility function maps $\Delta(\mathbf{X})$ to \mathbb{R} . Being continuous and linear, it is determined by its values on the extreme points of $\Delta(\mathbf{X})$, i.e. by its values on the point masses $\delta_c, c \in \mathbf{X}$. So starting from a continuous linear $u : \Delta(\mathbf{X}) \rightarrow \mathbb{R}$, we abuse notation slightly by defining the vNM utility function $u : \mathbf{X} \rightarrow \mathbb{R}$ by

setting $u(c)$ equal to $u(\delta_c)$. In the same fashion, a continuous linear function on $\mathfrak{K}(\Delta(\mathbf{X}))$ is determined by its value on the extreme *sets* in $\mathfrak{K}(\Delta(\mathbf{X}))$, providing an alternative to the Steiner direct sum decomposition. For $A, B \in \mathfrak{K}(\Delta(\mathbf{X}))$, we denote the set $\{\alpha A + (1 - \alpha)B : \alpha \in [0, 1]\}$ as a line segment, $\|A, B\|$.

Definition 4. $A \in \mathfrak{K}(\Delta(\mathbf{X}))$ is an **extreme set** if for all line segments $\|B, C\|$ containing A , either $A = B$ or $A = C$. The class of extreme sets is denoted $\text{extr}(\mathfrak{K})$.

The vNM part of a continuous linear U is given by the value of U on the extreme sets $\{\delta_c\}$, $c \in \mathbf{X}$. The “rest” of the utility function is specified by its value on the remaining extreme sets. In the two consequence case, there is only one other extreme set, $[0, 1]$, and its value, $-v$ is what appears in the representation $U([p - r, p + r]) = p - vr$. In the three outcome case, there are infinitely many other extreme sets, and one can represent U as the sum of its vNM part and how it behaves on the other sets.

Example 2 (Three consequences). In the case $\mathbf{X} = \{a, b, c\}$, $\Delta(\mathbf{X})$ is 2-dimensional, $\mathfrak{K}(\Delta(\mathbf{X}))$ is infinite dimensional, hence the class of extreme sets must be infinite. They are either of the form $\{\delta_a\}$, $\{\delta_b\}$, $\{\delta_c\}$, or are the closed convex hull of either two or three points, specifically $\text{co}(p, q, r)$, where $p, q, r \in \Delta(\mathbf{X})$ are of the form

$$\begin{aligned} p &= (p', (1 - p'), 0), \\ q &= (0, q', (1 - q')), \text{ and} \\ r &= ((1 - r'), 0, r'), \end{aligned}$$

where $p', q', r' \in [0, 1]$, and a triple (x, y, z) denotes $x\delta_a + y\delta_b + z\delta_c$. Special cases include:

1. (a) if $p' = q' = r' = 1$, then $\mathbf{co}(p, q, r) = \Delta(\mathbf{X})$;
2. if $q' = r' = 0$, then for all p , $\mathbf{co}(p, q, r) = \Delta(\{a, b\})$;
3. if $p' = q' = r' = \frac{1}{2}$, then $\mathbf{co}(p, q, r)$ is the triangle joining the mid-points of the faces of $\Delta(\mathbf{X})$;
4. if $p' = 0$ and $q' = 1$, then $\mathbf{co}(p, q, r)$ is the line joining δ_b and r .

The vNM part of U gives the value of $U(\{\delta_a\})$, $U(\{\delta_b\})$, and $U(\{\delta_c\})$. The rest of U can be specified by a function $(p', q', r') \mapsto f(p', q', r')$ that has $U(\mathbf{co}(p, q, r)) = f(p', q', r')$ Hausdorff continuous. This requires continuity plus: the value of p' not mattering much if q' and r' are close to 0; the value of q' not mattering much if p' and r' are close to 1; and the value of r' not mattering much if $1 - p'$ and q' are close to 0. Further, if $a \prec b \prec c$, then one would expect $f(\cdot, q', r')$ to be decreasing while $f(p', \cdot, r')$ and $f(p', q', \cdot)$ are increasing.

Let α be the interior angle of triangle $\mathbf{co}(p, q, r)$ at the vertex p , β the interior angle at q , and γ the interior angle at r . The cosine law delivers α , β , and γ from $\|p - q\|$, $\|p - r\|$, and $\|q - r\|$. Letting α', β', γ' be the complementary angles (i.e. $\alpha + \alpha' = 180$), the Steiner point of $\mathbf{co}(p, q, r)$ is $\frac{1}{360}[\alpha'p + \beta'q + \gamma'r]$.

1.4.4 The Choice of a Center

There are many continuous linear selections that can be used as the definition of the center of a compact convex set (e.g. by changing λ to some full support ν on $S^{\ell-1}$ in the Steiner point). Calculations using linear preferences on $\mathfrak{K}(\Delta(\mathbf{X}))$ cannot be affected by such a change in basis, except perhaps in their ease or difficulty. However, each different center gives rise to a different direct sum decomposition, hence labels different sets of linear preferences as ambiguity averse/neutral/seeking. In this sense, the decomposition of attitudes toward risk and attitudes toward ambiguity is dependent on the choice of the Steiner basis for the direct sum decomposition of $\mathfrak{K}(\Delta(\mathbf{X}))$.

Interest therefore centers on finding cases in which the choice of the Steiner point as the center has a more solid claim to special status. For example, when it agrees with the barycenter, then it commutes with linear homeomorphisms $T : \Delta(\mathbf{X}) \rightarrow \mathbb{H}$ with \mathbb{H} a Hilbert space. A rich class of sets that will often appear in the applications below have this property.

1. The Steiner point of any line $[p, q] \subset \Delta(\mathbf{X})$ is $\frac{1}{2}p + \frac{1}{2}q$, and the Steiner point of $T([p, q])$ is $T(\frac{1}{2}p + \frac{1}{2}q) = \frac{1}{2}T(p) + \frac{1}{2}T(q)$.
2. If $A = \{p \in \Delta : (p - q_0)^T M (p - q_0) = c\}$, M a positive definite matrix, the Steiner point is q_0 , and the Steiner point of $T(A)$ is $T(q_0)$.
3. By linearity, if $A = \sum_{i \leq I} \alpha_i [p_i, q_i]$, the α_i a convex set of weights, then A is a parallelipiped, its Steiner point is $\sum_i \alpha_i c_i$ where $c_i = \frac{1}{2}p_i + \frac{1}{2}q_i$, and the Steiner point of $T(A)$ is $T(\sum_i \alpha_i c_i)$.

4. In a similar fashion, if A is the convex sum of ellipses and parallelipeds, the Steiner point of $T(A)$ is $T(St(A))$.

This points to a variety of possible alternatives to the definition of ambiguity aversion given in Definition 3.

1. (a) One can define preferences to be ambiguity averse relative to a class of sets $\mathbb{A} \subset \mathfrak{R}(\Delta(\mathbf{X}))$ if expansions around the Steiner points of these sets reduce utility. In this case one might focus on the class \mathbb{A} of sets for which the barycenter and the Steiner point agree.
2. One can change basis in $\mathfrak{R}(\Delta(\mathbf{X}))$ by choosing a different center and ask for ambiguity aversion relative to expansions around that center.
3. One can combine different centers with different classes of sets.

Implicitly, the first of these approaches is being followed in the applications for which the barycenter and the Steiner point agree. The applications have not, as of yet, provided a compelling reason to pursue either the second or third approach.

1.5 Separating Risk and Ambiguity

Generalizing the finite consequence separation of attitudes toward risk and ambiguity requires picking a continuous linear center. In the infinite-dimensional context of $\Delta(\mathbf{X})$ when \mathbf{X} is an infinite set, there can be no rotation invariant continuous selection of a center [41].

Let ν be any Gaussian probability distribution with support equal to $\partial U \subset C_b(\mathbf{X})$. For $A \in \mathfrak{K}(\Delta(\mathbf{X}))$, define the ν -center of A as

$$C_\nu(A) = \int_{\partial U} \operatorname{argmax}_{p \in A} \langle p, f \rangle d\nu(f). \quad (1.9)$$

Since the function $f \mapsto \max_{p \in A} \langle p, f \rangle$ is concave and bounded, its non-differentiability points are a Gauss null set (see Benyamini and Lindenstrauss for this). Thus for only a ν -null set does the argmax contain more than 1 point, hence the integral is well-defined. Since A is closed and convex and ν is a probability, $C_\nu(A) \in A$. The continuity of $C_\nu(\cdot)$ follows from dominated convergence, linearity is immediate.

Every $A \in \mathfrak{K}(\Delta(\mathbf{X}))$ has a unique expression as $A = (A - C_\nu(A)) + \{C_\nu(A)\}$, and $C_\nu(A - C_\nu(A)) = 0$. Therefore

$$L(A) = L((A - C_\nu(A))) + L(\{C_\nu(A)\}), \text{ and } C_\nu(A - C_\nu(A)) = 0. \quad (1.10)$$

Restricted to singleton sets, L is an expected utility function, call it ℓ . For any expected utility function, ℓ' , we can define

$$L'(A) = L((A - C_\nu(A))) + \ell'(\{C_\nu(A)\}). \quad (1.11)$$

The attitude toward risk is contained in either ℓ or ℓ' , and is independent of the attitude toward ambiguity, contained in the restriction of L to the sets having ν -center equal to 0.

1.6 First and Second Order Stochastic Dominance

Suppose now that \mathbf{X} is a compact interval, $[0, M]$.

Definition 3. For sets $A, B \in \mathfrak{K}(\Delta(\mathbf{X}))$, we say that A **first order dominates** B , written $A \succeq_F B$, if for all non-decreasing $u \in C(\mathbf{X})$,

$$\max_{p \in A} \langle u, p \rangle \geq \max_{q \in B} \langle u, q \rangle, \text{ and } \min_{p \in A} \langle u, p \rangle \geq \min_{q \in B} \langle u, q \rangle. \quad (1.12)$$

For second order dominance, we require that the inequalities in (1.12) also hold for non-decreasing concave u .

$A \succeq_F B$ iff every expected utility maximizer with monotonic preferences prefers their favorite element of A to their favorite element of B and simultaneously prefers their least favorite element of A to their least favorite element of B .

Definition 4. Preferences \succeq on $\mathfrak{K}(\Delta)$ respect first order dominance if $[A \succeq_F B] \Rightarrow [L(A) \geq L(B)]$.

Theorem 1.6.1. *If η_{\min} and η_{\max} are non-negative and supported on the non-decreasing elements of ∂U , then the associated L respects first order dominance. If L respects first order dominance, then it can be represented by a pair of non-negative η_{\min} and η_{\max} that are supported on the non-decreasing elements of ∂U .*

Proof. The first part follows from the monotonicity of integrals. The second part follows (**open problem here**) from generalizing examples in the finite case. □

Chapter 2

Applications

2.1 Applications

This section covers the beginnings of portfolio choice theory for ambiguity, giving the ambiguous versions of first and second order stochastic dominance.

2.1.1 First and Second Order Dominance, Diversification

First is a function form that makes clearer some of the basics: first and second order dominance, comparative ambiguity aversion, and risk equivalents.

2.1.1.1 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 parallelepipeds spanned by intervals of point masses. We normalize so that $f(0, 0) = 0$ and $f(M, M) = 1$.

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

$$f(a, b) = \frac{1}{2}u(a) + \frac{1}{2}u(b) - \frac{r}{2}|v(b) - v(a)|^h, \text{ with } h \in [1, \infty). \quad (2.1)$$

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 as in the following table. The α -MMEU preferences have kinks, and it is the presence of these kinks that leads e.g. to Dow and Werlang's [13] results that there are ranges of prices at which DM's will neither buy nor sell a stock.

| | 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) | $Df = \begin{pmatrix} \frac{1}{2}u'(a) + r(u(b) - u(a)) \\ \frac{1}{2}u'(b) - r(u(b) - u(a)) \end{pmatrix}$ | $Df = \begin{pmatrix} \frac{1}{2}u'(a) - r(u(a) - u(b)) \\ \frac{1}{2}u'(b) + r(u(a) - u(b)) \end{pmatrix}$ |

2.1.1.2 First Order Stochastic Dominance

For F_1 and F_2 closed subsets of $[0, M]$, F_1 **dominates** F_2 , written $F_1 \succeq_D F_2$, 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 $\{c, d\}$, $a \leq b$, $c \leq d$, $F_1 \succeq_D F_2$ requires $a \geq c$ and $b \geq d$.

Definition 5. A linear U on $\mathfrak{K}(\Delta)$ is **set monotonic** if $[F_1 \succeq_D F_2] \Rightarrow [U(A_{F_1}) \geq U(A_{F_2})]$,

When F_1 and F_2 are singleton sets, set monotonicity is equivalent to the monotonicity of the vNM 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 $p \succeq_{FOSD} q$, 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 6. For $A, B \in \mathfrak{K}(\Delta)$, A **dominates** B , $A \succeq_D B$, if for all $p \in A$, there is a q in B such that $p \succeq_{FOSD} q$ and for all $q \in B$, there is a $p \in A$ such that $p \succeq_{FOSD} q$.

Note that this is *not* the lattice ordering for sets inherited from the lattice $(\Delta[0, M], \succeq_{FOSD})$, 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 $A \succeq_{SD} B$ 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 $A \in \mathfrak{K}(\Delta)$ is not a lattice subset of Δ , then it is not the case that $A \succeq_{SD} A$ while $A \succeq_D A$ as a triviality since every p satisfies $p \succeq_{FOSD} p$. More

substantively, let $A = \{p \in \Delta : p[M/3, M] = 1, p \text{ has a density w.r.t. Lebesgue measure}\}$ and $B = \{q \in \Delta : q[0, 2M/3] = 1\}$. $A \succeq_D B$ 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 $A \succeq_{SD} B$ — 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 2.1.1. *For closed $F_1, F_2 \subset [0, M]$, $F_1 \succeq_D F_2$ iff $A_{F_1} \succeq_D A_{F_2}$.*

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

Definition 7. A function U on $\mathfrak{K}(\Delta)$ **respects dominance** if $[A \succeq_D B] \Rightarrow [U(A) \geq U(B)]$.

Theorem 2.1.2. *If a linear U respects dominance, then it is set monotonic.*

Proof. If $F_1 \succeq_D F_2$, then $A_{F_1} \succeq_D A_{F_2}$, 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 $\mathfrak{K}(\Delta)$ 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 $U : \mathfrak{K}(\Delta) \rightarrow \mathbb{R}$ are determined by its values on the extreme points in the space of sets $\mathfrak{K}(\Delta)$. However, not all extreme sets are of the form A_F , F a closed subset of $[0, M]$.

Lemma 2.1.3. $\mathcal{A}_F := \{A_F : F \subset [0, M], F \text{ closed, non-empty}\}$ does not exhaust the class of extreme sets of $\mathfrak{K}(\Delta)$.

Proof. Suppose first that we replace the consequence space $[0, M]$ by a three point set, $\mathbf{X} = \{a, b, c\}$. 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. $\mathfrak{K}(\mathbf{X})$ 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 $S(\mathbf{X})$ 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, $F_n, F, [d(F_n, F) \rightarrow 0] \Rightarrow [A_{F_n} \rightarrow A_F]$. 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 $B_n \rightarrow B$ 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 . \square

2.1.1.3 Second Order Stochastic Dominance

A continuous linear utility on probabilities respects second order stochastic dominance iff the associated vNM utility function is concave. The parallel concept for a linear U on $\mathfrak{R}(\Delta)$ 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 8. 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 $A, B \in \mathfrak{R}(\Delta)$, 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 9. A linear U on $\mathfrak{R}(\Delta)$ is **horizontally concave** if for all $A, B \in \mathfrak{R}(\Delta)$, $U(\alpha A \oplus_{hz} (1 - \alpha) B) \geq \alpha U(A) + (1 - \alpha) U(B)$.

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 vNM 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, onconvity of $u(\cdot)$ and $v(\cdot)$ in $f(a, b) = \frac{1}{2} u(a) + \frac{1}{2} u(b) - \frac{\tau}{2} |v(b) - v(a)|^h$ from eqn. (2.1) 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 parallelipeds based on point masses.

It would be nice to know if horizontal concavity has implications for yet broader classes of $A, B \in \mathfrak{K}(\Delta)$ 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.

Bibliography

- [1] Ahn, D. S. (2008), Ambiguity Without a State Space, *Review of Economic Studies* **75**, 3-28.
- [2] Allais, M. 1953, Le Comportement de l'Homme Rationnel Devant le Risque: Critique des Postulates et Axiomes de l'Ecole Americaine. *Econometrica* **21**, 503-546
- [3] Amarante, M. 2003, States, Models and Unitary Equivalence. Working paper, Department of Economics, Columbia University, New York.
- [4] Amarante, M. 2006, thin sets, definitions of ambiguity.
- [5] Anscombe, F.J. and R. J. Aumann (1963), A Definition of Subjective Probability, *Annals of Mathematical Statistics* **34**, 199-205.
- [6] Blackwell, D. 1951, Comparison of experiments, *Proceedings of the Second Berkeley Symposium on Mathematical Statistics and Probability*, 1950 (Berkeley and Los Angeles), University of California Press, pp. 93-102.
- [7] Blackwell, D. 1953, Equivalent comparisons of experiments, *Ann. Math. Statistics* **24**, 265-272.
- [8] Chambers, C. 2008, Proper Scoring Rules for General Decision Models, *Games and Economic Behavior* **63**, 32-40.

- [9] Corbae, D., M. Stinchcombe, and J. Zeeman, 2009, *An Intro etc.*. Princeton University Press.
- [10] Crawford, V. 1990, Equilibrium without Independence. *Journal of Economic Theory* **50**, 1, 127-54.
- [11] Dekel, E. 1986, An axiomatic characterization of preferences under uncertainty: Weakening the independence axiom, *Journal of Economic Theory* **40**, 304-318.
- [12] Dekel, E., B. L. Lipman, and A. Rustichini, 2001, Representing Preferences with a Unique Subjective State Space *Econometrica* **69**, 891-934.
- [13] Dow, J. and S. Werlang, 1992, Uncertainty Aversion, Risk Aversion, and the Optimal Choice of Portfolio, *Econometrica* **60**(1), 197-204.
- [14] Dow, J. and S. Werlang, 1994, Nash Equilibrium under Knightian Uncertainty: Breaking Down Backward Induction, *Journal of Economic Theory* **64**, 305-324.
- [15] Ellsberg, D. 1961, Risk, Ambiguity, and the Savage Axioms, *Quarterly Journal of Economics* **75**(4), 643-669.
- [16] Epstein, L. 1999, A Definition of Uncertainty Aversion. *Review of Economic Studies* **66**, 579-608.
- [17] Epstein, L. and J. Zhang, 2001, Subjective Probabilities on Subjectively Unambiguous Events. *Econometrica* **69**, 265-306.

- [18] Fellner, W. 1961, Distortion of Subjective Probabilities as a Reaction to Uncertainty, *Quarterly Journal of Economics* **75**(4), 670-689.
- [19] Fenchel, W. and Jessen, B. 1938, Mengenfunktionen und konvexe Körper. *Danske Vid. Selskab. Mat.-fis. Medd.* **16**, 3.
- [20] Ghirardato, P. Marinacci, M. 2001, Range Convexity and Ambiguity Averse Preferences. *Economic Theory* **17**(3), 599-617.
- [21] Ghirardato, P. Marinacci, M. 2002, Ambiguity Made Precise: a Comparative Foundation *Journal of Economic Theory* **102**, 251-289.
- [22] Ghirardato, P. F. Maccheroni and M. Marinacci, 2004, Differentiating ambiguity and ambiguity attitude *Journal of Economic Theory* **118**, 133-173.
- [23] Gilboa, I. 1987, Expected Utility with Purely Subjective Non-Additive Priors, *Journal of Mathematical Economics* **16**, 279-304.
- [24] Gilboa, I. and D. Schmeidler (1989), Maxmin expected utility with non-unique priors, *Journal of Mathematical Economics* **18**, 141-153.
- [25] Karni, E., Schmeidler, D., Vind, K. 1983, On State Dependent Preferences and Subjective Probabilities. *Econometrica* **51**(4), 1021-1032.
- [26] Knight, F. H. 1921, *Risk, uncertainty and profit*. Chicago, University of Chicago Press (1971, reprint of the 1921 ed.).

- [27] Körner, R. 1995, A Variance of Compact Convex Random Sets. Working paper, Institut für Stochastik, Bernhard-von-Cotta-Str. 2 09599 Freiberg.
- [28] Maccheroni, F. M. Marinacci, and A. Rustichini, 2006, Ambiguity Aversion, Robustness, and the Variational Representation of Preferences. *Econometrica* **74**(6) 1447-1498.
- [29] Machina, M. 1982, 'Expected Utility' Analysis Without the Independence Axiom. *Econometrica* **50**, 277-323.
- [30] Marinacci, M. 2002, Probabilistic Sophistication and Multiple Priors, *Econometrica*, 70, 755-764.
- [31] Mukerji, S. 1998, Ambiguity Aversion and Incompleteness of Contractual Form. *American Economic Review* **88**(5), 1207-31.
- [32] Mukerji S.; Tallon J-M. 2001, Ambiguity Aversion and Incompleteness of Financial Markets. *The Review of Economic Studies* **68**(4), 883-904.
- [33] Olszewski, W. 2007, Preferences Over Sets of Lotteries, *Review of Economic Studies* **74**(2), 567-595.
- [34] Von Neumann, J. and Morgenstern, O. 1944, *Theory of games and economic behavior*. New York, Wiley, (1953, 3d ed.).
- [35] Rådstrøm, H. 1952, An Imbedding Theorem for Spaces of Convex Sets, *Proceedings of the American Mathematical Society* **3**(1), 165-169.
- [36] Savage, L. J. 1954, *The foundations of statistics*. New York, Wiley.

- [37] Schmeidler, D. (1989), Subjective Probability and Expected Utility without Additivity, *Econometrica* **57**(3), 571-587.
- [38] Schneider, R. 1993, *Convex Bodies: The Brunn-Minkowski Theory*, Cambridge, Cambridge University Press.
- [39] Siniscalchi, M. 2006, A Behavioral Characterization of Plausible Priors, *Journal of Economic Theory* **128**, 91-135.
- [40] Starr, R. 1969, Quasi-equilibria for Economies with Non-Convex Preferences, *Econometrica* **37**, 25-38.
- [41] Vitale, R. 1985, Steiner Points in Infinite Dimensions, *Israel Journal of Mathematics*.
- [42] Weaver, N. 1999, *Lipschitz Algebras*. New Jersey, World Scientific.

Vita

Martin Dumav was born in Istanbul, Turkey on 11 April 1982. He received the Bachelor of Science degree in Engineering from the Istanbul Technical University in 2004. He attended and received the Master of Arts degree from Emory University in 2006. After then he applied to the University of Texas at Austin for enrollment in their economics program. He was accepted and started graduate studies in August, 2006. While studying economics, he took graduate mathematic courses to support his research program in economics.

Permanent address: Department of Economics
University of Texas
Austin, TX 78712-0301 USA
e-mail: mdumav@gmail.com

This report was typeset with L^AT_EX[†] by the author.

[†]L^AT_EX is a document preparation system developed by Leslie Lamport as a special version of Donald Knuth's T_EX Program.