Random Expected Utility†

Faruk Gul

and

Wolfgang Pesendorfer

Princeton University

August 2004

Abstract

We develop and analyze a model of random choice and random expected utility. A *decision problem* is a finite set of lotteries describing the feasible choices. A *random choice rule* associates with each decision problem a probability measure over choices. A *random utility function* is a probability measure over von Neumann-Morgenstern utility functions. We show that a random choice rule maximizes some random utility function if and only if it is *mixture continuous*, *monotone* (the probability that a lottery is chosen does not increase when other lotteries are added to the decision problem), *extreme* (lotteries that are not extreme points of the decision problem are chosen with probability zero), and *linear* (satisfies the independence axiom).

† This research was supported by grants SES0236882 and SES0214050 from the National Science Foundation. We thank the editor and three anonymous referees for their suggestions and comments.

1. Introduction

In this paper, we develop and analyze a model of random choice and random expected utility. Modeling behavior as stochastic is a useful and often necessary device in the econometric analysis of demand. The choice behavior of a group of subjects with identical characteristics each facing the same decision problem presents the observer with a frequency distribution over outcomes. Typically, such data is interpreted as the outcome of independent random choice by a group of identical individuals. Even when repeated decisions of a single individual are observed, choice behavior may exhibit variation and therefore suggest random choice by the individual.

Let Y be a set of choice objects. A finite subset D of Y represents a decision problem. The individual's behavior is described by a random choice rule ρ which assigns to each decision problem a probability distribution over feasible choices. The probability that the agent chooses $x \in D$ is denoted $\rho^D(x)$. A random utility function is a probability measure μ on some set of utility functions $U \subset \{u : Y \to I\!R\}$. The random choice rule ρ maximizes the random utility function μ if $\rho^D(x)$ is equal to the $\mu-$probability of choosing some utility function u that attains its maximum in D at x (for all D, x).

Modeling random choice as a consequence of maximizing a random utility function is common practice in both empirical and theoretical work. When the frequency distribution of choices describes the behavior of a group of individuals, the corresponding random utility function is interpreted as a random draw of a member of the group (and hence of his utility function). When the data refers to the choices of a single individual, the realization of the individual's utility function can be interpreted as the realization of the individual's private information. In the analysis of preference for flexibility (Kreps (1979), Dekel, Lipman and Rustichini's (2001)) the realization of the agent's random utility function corresponds the realization of his subjective (emotional) state.

In all these cases, the random utility function is observable only through the resulting choice behavior. Hence, testable hypotheses must be formulated with respect to the random choice rule ρ. A central objective of the random choice literature has been to identify those random choice rules that are consistent with random utility maximization.

This amounts to answering the following question: what conditions on ρ are necessary and sufficient for there to exist a random utility function μ that is maximized by ρ?

We answer this question for random expected utility maximization. Hence, the set U consists of all von Neumann-Morgenstern utility functions. In many applications, economic agents choose among risky prospects. For example, understanding random choice in the context of the portfolio choice problem requires interpreting choice behavior as a stochastic version of a particular theory of behavior under risk. Our theorem enables us to relate random choice to the simplest theory of choice under uncertainty; expected utility theory.

The choice objects in our model are lotteries over a finite set of prizes. We identify four properties of random choice rules that ensure its consistency with random expected utility maximization. These properties are (i) monotonicity, (ii) mixture continuity, (iii) linearity, and (iv) extremeness.

A random choice rule is *monotone* if the probability of choosing x from D is at least as high as the probability of choosing x from $D \cup \{y\}$. Thus, monotonicity requires that the probability of choosing x cannot increase as more alternatives are added to the choice problem.[1]

A random choice rule is *mixture continuous* if it satisfies a stochastic analogue of the von Neumann-Morgenstern continuity assumption. We also use a stronger continuity assumption (*continuity*) which requires that the random choice rule is a continuous function of the decision problem.

A random choice rule is *linear* if the probability of choosing x from D is the same as the probability of choosing $\lambda x + (1-\lambda)y$ from $\lambda D + (1-\lambda)\{y\}$. Linearity is the analogue of the independence axiom in a random choice setting.

A random choice rule is *extreme* if, with probability one, the chosen lottery is an extreme point of the decision problem. Extreme points are those elements of the choice problem that are unique optima for *some* von Neumann-Morgenstern utility function.

A *regular* random utility function is one where in any decision problem, with probability 1, the realized utility function has a unique maximizer.[2] Hence, for a regular random utility function ties are 0-probability events.

[1] Monotonicity is a well known implication of maximization of a random utility function. It is often referred to as *regularity* (Luce and Suppes (1965)).

[2] In the psychology literature this property is referred to as *non-coincident*, see Falmagne (1983).

Theorem 2 is our main result. It says that *a random choice rule maximizes some regular (finitely additive) random utility function if and only if the random choice rule is monotone, mixture continuous, linear and extreme.* Hence, monotonicity, mixture continuity, linearity, and extremeness are the only implications of random expected utility maximization.

Theorem 2 permits random utility functions that are not countably additive. In Theorem 3 we characterize the behavior generated by countably additive random utility functions. Theorem 3 says that *a random choice rule maximizes some regular, countably additive, random utility function if and only if the random choice rule is monotone, continuous, linear and extreme.* Hence, if we add the requirement that the random utility function is countably additive then the maximizing random choice rule is continuous (rather than mixture continuous). Conversely, if we add the requirement that the random choice rule is continuous (which implies mixture continuity) then the corresponding random utility function is countably additive.

A deterministic utility function is a special case of a random utility function. Clearly, it is not regular since there are choice problems for which ties occur with positive probability. The difficulty with non-regular random utility functions is that the associated choice behavior is ambiguous. In section 5, we show that this difficulty can be overcome by adding a tie-breaking rule. Theorem 5 demonstrates that our tie-breaking rules lead to well defined random choice rules that are monotone, mixture continuous, linear and extreme. Hence, the maximizers of non-regular random utilities are identified by the same conditions as the maximizers of regular random utilities. In this sense, the restriction to regular random utilities is without loss of generality. Put differently, for any maximizer ρ of a non-regular random utility function there is a *regular* random utility function μ' such that ρ also maximizes μ' (Theorem 7). To achieve this generality it is essential that we allow random utilities that are not countably additive: Theorem 8 shows that - except for trivial cases - imposing a tie-breaker on a non-regular random utility yields a regular but only finitely additive random utility function.

Studies that investigate the empirical validity of expected utility theory predominantly use a random choice setting. For example, the studies described in Kahneman and Tversky

(1979) report frequency distributions of the choices among lotteries by groups of individuals. Their tests of expected utility theory focus on the independence axiom. In particular, the version of the independence axiom tested in their experiments corresponds exactly to our linearity axiom. It requires that choice frequencies stay unchanged when each alternative is combined with some fixed lottery. Our theorems identify all of implications of random expected utility maximization that are relevant for the typical experimental setting.

The related literature on stochastic choice (McFadden and Richter (1991), Falmagne (1978), Clark (1995)) has focused on deterministic alternatives rather than lotteries. In the supplement to this paper (Gul and Pesendorfer (2004)) we adapt the analysis found in the literature to our setting to facilitate a precise comparison of the results.[3]

McFadden and Richter (1991) consider a setting with a finite set of alternatives.[4] They introduce the axiom of revealed stochastic preference (ARSP) - a stochastic analogue of the strong axiom of revealed preference - and show that it is necessary and sufficient for a random choice rule to maximize some regular random utility function. In the supplement (Gul and Pesendorfer (2004)) we provide an appropriate version of ARSP for the setting considered in this paper. It is immediate that this axiom is necessary for a random choice rule to maximize some random utility function. We show that it implies monotonicity, linearity, extremeness and mixture continuity. Hence, our main theorem implies that ARSP applied to our setting is equivalent to monotonicity, linearity, extremeness and mixture continuity.[5]

Clark (1995) considers a setting that allows for an arbitrary collection of utility functions and decision problems. He introduces an axiom termed "coherency" and shows that it is necessary and sufficient for a random choice rule to maximize a regular random utility function. Coherency is closely related to a theorem of De Finetti's which provides a necessary and sufficient condition for a function defined on a collection of subsets to have

[3] In addition to the literature discussed below, there is also an extensive literature on stochastic binary choice. See Fishburn (1992) for a survey of this literature.

[4] McFadden (2003) provides extensions of the results in McFadden and Richter (1991) to the case with an infinite set of alternatives.

[5] Note that ARSP is a condition that is imposed jointly on the space of utility functions and on the random choice rule. Because we consider von Neumann-Morgenstern utility functions, ARSP implies linearity and extremeness in our setting.

an extension to a finitely additive probability measure on the smallest algebra containing those subsets. In the supplement (Gul and Pesendorfer (2004)) we apply coherency to our setting and show that it implies monotonicity, linearity, extremeness and mixture continuity.[6] Clark shows that coherency is necessary and hence our main theorem implies that monotonicity, mixture continuity, linearity, and extremeness are equivalent to coherency in our model.

Coherency can also be applied in settings where we only observe the choice behavior in a subset of the possible decision problems. In that case, coherency is necessary and sufficient for the implied random utility function to have an extension that is a probability measure. Thus the observed choice probabilities satisfy coherency if and only if one can construct a random utility function μ such that the observed behavior is consistent with $\mu-$maximization. This is in contrast to the conditions given in this paper. A finite data set may not violate any of our axioms but nevertheless be inconsistent with maximization of any random utility function.

A third strand of the literature related to our work is Falmagne (1978) and Barbera and Pattanaik (1986)[7]. Falmagne studies the case where choice problems are arbitrary subsets of a finite set of alternatives. His characterization of random choice identifies a finite number (depending on the number of available alternatives) of non-negativity conditions as necessary and sufficient for a random choice rule to maximize some regular random utility function. We can relate our results to those of Falmagne by considering a finite subset of decision problems. In particular, consider the decision problems consisting of degenerate lotteries that yield one of the prizes with probability 1. If a random choice rule on this restricted class of decision problems satisfies Falmagne's conditions then it can be extended to a random choice rule on all decision problems that satisfies monotonicity, mixture continuity, linearity, and extremeness. Our main result implies that the converse is true as well. Thus, Falmagne's conditions are necessary and sufficient for a random choice function over a finite set of prizes to have a mixture continuous, monotone, linear

[6] Like ARSP, coherency is a condition that is imposed jointly on the space of utility functions and on the random choice rule.

[7] Barbera and Pattanaik (1986) provide an exposition and refinement of the work of Falmagne (1978)

and extreme extension to the set of decision problems generated by lotteries over those prizes.[8]

ARSP, Coherency and Falmagne's conditions are complicated restrictions on arbitrary finite collections of decision problems. This makes them difficult to interpret. In contrast, monotonicity, linearity and extremeness are simple conditions in that each involves the comparison of pairs of decision problems. Moreover, each of these axioms is a straightforward extensions of axioms from the deterministic setting and therefore easy to interpret. The simplicity of the axioms facilitates the construction of experiments that attempt to falsify the theory.[9] In fact, linearity has been tested extensively (see for example, Kahnemann and Tversky (1979)). We could imagine similar tests for monotonicity and extremeness.

As in the case of linearity one would expect to find circumstances under which our other assumptions are violated in experimental settings. These violations can be the impetus for theories that generalize our model just like violations of linearity have stimulated theoretical research that generalizes expected utility (see Machina (1989)). To provide this impetus the axioms must be interpretable and must separate the key ingredients of the theory. Our axioms accomplish this for a stochastic version of expected utility theory. Violations of linearity or extremeness point to distinct violations of the expected utility hypothesis. Potentially, these violations could be addressed by generalizing the set of admissible utility functions. Violations of monotonicity point to a more basic failure of Chernoff's postulate 4 (or Sen's condition α) which requires an optimal choice to remain optimal when alternatives are removed from the choice set. Therefore, violations of monotonicity are inconsistent with maximization of a random utility function even if those utility functions are allowed to be non-linear.

---

[8] Assuming that Falmagne's conditions hold for any finite collection of decision problems, Cohen (1980) extends Falmagne's model to arbitrary infinite sets of alternatives (maintaining the assumption that decision problems are finite). In contrast, our model provides an extension of Falmagne's model to a specific infinite set - the set of lotteries over a finite set of prizes. Cohen does not place any restriction on the utility functions while we restrict attention to von Neumann-Morgenstern utility functions. Hence, monotonicity, mixture continuity, linearity, and extremeness imply that Cohen's conditions are satisfied in our setting but the converse is not true. Cohen's conditions imply monotonicity but not linearity or extremeness.

[9] It seems straightforward to construct experiments that test linearity, extremeness or monotonicity. However, like any continuity assumption, mixture continuity cannot be falsified by a finite data set. Hence, empirical tests based on our conditions must assume mixture continuity.

2. Random Choice and Random Utility

There is a finite set of prizes denoted $N = \{1, 2, \ldots, n+1\}$ for $n \geq 1$. The objects of choice are lotteries over the prizes N. Let $P := \{x \in I\!R^{n+1}_{+} | \sum_{i=1}^{n+1} x^i = 1\}$ be the unit simplex in $I\!R^{n+1}$. We denote with $x \in P$ a lottery over N.

A *decision problem* is a nonempty, finite set of lotteries $D \subset P$. Let $\mathcal{D}$ denote the set of all decision problems. We are concerned with a decision maker who makes stochastic choices from decision problems. The decision maker is characterized by a *random choice rule* that associates each decision problem D with a probability measure over choices.

Let $\mathcal{B}$ denote the Borel sets of P and Π be the set of all probability measures on the measurable space $(P, \mathcal{B})$.

Definition: *A random choice rule (RCR) is a function $\rho : \mathcal{D} \to \Pi$ with $\rho^D(D) = 1$.*

The probability measure ρ^D describes the agent's behavior when facing the decision problem D. We use $\rho^D(B)$ to denote the probability that the agent chooses a lottery in the set $B \in \mathcal{B}$ when faced with the decision problem D and write $\rho^D(x)$ instead of $\rho^D(\{x\})$. Note that ρ^D is defined to be a measure on $(P, \mathcal{B})$ rather than on the set of feasible choices D. Feasibility is ensured by the requirement that the support of ρ^D is D, i.e., $\rho^D(D) = 1$.

The purpose of this paper is to relate random choice rules and the behavior associated with maximizing a *random utility function.* We consider von Neumann-Morgenstern utility functions and therefore each utility function u can be identified with an element of $I\!R^{n+1}$. We write $u \cdot x$ rather than $u(x)$, where $u \cdot x = \sum_{i=1}^{n+1} u^i x^i$. Since $(u^1, \ldots, u^{n+1}) \cdot x \geq (u^1, \ldots, u^{n+1}) \cdot y$ if and only if $(u^1 - u^{n+1}, u^2 - u^{n+1}, \ldots, 0) \cdot x \geq (u^1 - u^{n+1}, u^2 - u^{n+1}, \ldots, 0) \cdot y$ for all $x, y \in P$, we can normalize the set of utility functions and work with $U := \{u \in I\!R^{n+1} \,|\, u^{n+1} = 0\}$.

Let $M(D, u)$ denote the maximizers of u in the choice problem D. That is,

$$M(D, u) = \{x \in D \,|\, u \cdot x \geq u \cdot y \, \forall y \in D\}$$

When the agent faces the decision problem D and the utility function u is realized the agent must choose an element in $M(D, u)$. Conversely, when the choice $x \in D$ is observed the agent's utility function must be in the set

$$N(D, x) := \{u \in U | \, u \cdot x \geq u \cdot y \, \forall y \in D\}$$

(For $x \notin D$, we set $N(D,x) = \emptyset$.)

Let $\mathcal{F}$ denote the smallest field (algebra) that contains $N(D,x)$ for all (D,x). A *random utility function* is a probability measure defined on $(U,\mathcal{F})$.

Definition: *A random utility function (RUF) is a function $\mu : \mathcal{F} \to [0,1]$ such that $\mu(U) = 1$ and $\mu(F \cup F') = \mu(F) + \mu(F')$ whenever $F \cap F' = \emptyset$ and $F, F' \in \mathcal{F}$. The RUF μ is countably additive if $\sum_{i=1}^{\infty} \mu(F_i) = \mu(\bigcup_{i=1}^{\infty} F_i)$ whenever $F_i, i = 1, \dots$ is a countable collection of pairwise disjoint sets in $\mathcal{F}$ such that $\bigcup_{i=1}^{\infty} F_i \in \mathcal{F}$.*

When we refer to a RUF μ, it is implied that μ is finitely additive but may not be countably additive. We refer to a countably additive μ as a countably additive RUF.

Example 1: Assume there are two prizes, i.e., $n+1 = 2$. The set U consists of all the linear combinations of the vectors $(1,0)$ and $(-1,0)$. There are three distinct (von Neumann-Morgenstern) utility functions, corresponding to the vectors $u = (0,0), u' = (1,0), u'' = (-1,0)$. The algebra $\mathcal{F}$ consists of all unions of the sets $\emptyset, F_0, F_1, F_{-1}$ where $F_0 = \{(0,0)\}$, $F_1 = \{\lambda(1,0) | \lambda > 0\}$ and $F_{-1} = \{\lambda(-1,0) | \lambda > 0\}$. Let $\mu(F_0) = 0, \mu(F_1) = \mu(F_{-1}) = 1/2$. The RUF μ describes an agent who is equally likely to have a strict preference for prize 1 and a strict preference for prize 2.

Example 2: Assume there are three prizes, i.e., $n+1 = 3$ and $U : \{(u,0) | u \in I\!R^2\}$. Let $F_{uv} := \{\alpha(u,0) + \beta(v,0) | \alpha, \beta > 0\}$. The set F_{uv} is the collection of von Neumann-Morgenstern utility functions that are positive linear transformations of a strict convex combinations of $(u,0)$ and $(v,0)$. The set F_{uu} is the collection of utility functions that are positive linear transformations of $(u,0)$. Let $\mathcal{H} := \{F_{uv} | u, v \in I\!R^2\} \cup \emptyset$. The algebra $\mathcal{F}$ is the collection of finite unions of sets in $\mathcal{H}$. The following random utility corresponds to the "uniform distribution" over von Neumann-Morgenstern utility functions. Let $\mu(F_{uv}) = 0$ if $u = \lambda v$ for some $\lambda \in I\!R$ and let $\mu(F_{uv}) = 1/2\pi \arccos\left(\frac{uv}{\|u\|\|v\|}\right)$ if $u \neq \lambda v$ for $\lambda \in I\!R$.[10] The RUF μ assigns each F_{uv} a measure proportional to the angle between the vectors u and v.

A *regular RUF* is one for which in every decision problem with probability 1 the realized utility function has a unique maximizer. For $x \in D$, let

$$N^+(D,x) := \{u \in U | \, u \cdot x > u \cdot y \;\forall y \in D, y \neq x\}$$

[10] $\|\cdot\|$ denotes the Euclidian norm.

be the set of utility functions that have x as the unique maximizer in D. (For $x \notin D$, we set $N^+(D,x) = \emptyset$.) The set $\bigcup_{x \in D} N^+(D,x)$ is the set of utility functions that have a unique maximizer in D. Proposition 6 shows that $\mathcal{F}$ contains $N^+(D,x)$ for all (x,D).

Definition: *The RUF μ is regular if $\mu(\bigcup_{x \in D} N^+(D,x)) = 1$ for all $D \in \mathcal{D}$.*

The definition of regularity can be re-stated as

$$\mu(N^+(D,x)) = \mu(N(D,x))$$

for all $D \in \mathcal{D}$ and $x \in D$.

The RUFs in Examples 1 and 2 are regular. In the one-dimensional case (illustrated in Example 1) the RUF μ is regular if and only if $\mu(F_0) = 0$, that is, the utility function that is indifferent between the two prizes ($u = (0,0)$) is chosen with probability zero. In the two-dimensional case (illustrated in Example 2) the RUF μ is regular if and only if $\mu(F_{uv}) > 0$ implies $u \neq \lambda v$, i.e., the utility functions $(v,0)$ is not a linear transformation of $(u,0)$. Hence, any set F_{uu} that consists of the positive linear transformations of a single utility function must have measure zero.[11]

The RCR ρ maximizes the regular RUF μ if for any $x \in D$, the probability of choosing x from D is equal to the probability of choosing a utility function that is maximized at x.

Definition: *The RCR ρ maximizes the regular RUF μ if $\rho^D(x) = \mu(N(D,x))$ for all $D \in \mathcal{D}$ and $x \in P$.*

Below, we describe the maximizing random choice rules for Examples 1 and 2.

Example 1 continued: The regular RUF μ in Example 1 is maximized by the RCR ρ that chooses each extreme point of every decision problem with equal probability. That is, for any decision problem $D = \{x_1, ..., x_k\}$ with $x_1^1 \leq ... \leq x_k^1$ the lotteries x_1 and x_k are each chosen with probability $1/2$. (Obviously, if $k = 1$ then the lottery x_1 is chosen with probability 1.)

[11] To see why $\mu(F_{uu}) = 0$ is necessary for regularity, let $\bar{u} = (u,0)$ and note that we can always find $x, y, x \neq y$ such that $\bar{u}x = \bar{u}y$. If $\mu(F_{uu}) > 0$ then there is a strictly positive probability of drawing a utility function that is indifferent between x and y in decision problem $D = \{x,y\}$. Hence, μ is not regular. Sufficiency of the stated condition is a consequence of Lemma 2 in the Appendix.

Example 2 continued: The regular RUF μ in Example 2 is maximized by the following RCR ρ. If all choices of the decision problem can be expressed as a convex combination of two extreme lotteries (and hence the decision problem is one-dimensional[12]) then each of the two extreme lotteries is chosen with equal probability. Hence, for one-dimensional decision problems, Example 2 reduces to Example 1 above. If D is two-dimensional, then $N(D,x)$ has the form $\{\alpha u + \beta v | \alpha \geq 0, \beta \geq 0\}$ for some $u, v \in I\!R^2$. Then, $\rho^D(x)$ is proportional to the angle between u and v. More precisely, $\rho^D(x) = \mu(F_{uv})$ where $\mu(F_{uv})$ is as defined in Example 2 above.

Not all RUF's are regular. In particular, as illustrated in Example 3 below, deterministic choice interpreted as a degenerate RUF is not regular whenever $n > 1$. In section 5, we extend the notion of RUF maximization to include non-regular RUFs.

Example 3: For the utility function $\bar{u} \in I\!R^{n+1}$ define $\mu_{\bar{u}}$ as follows: $\mu_{\bar{u}}(F) = 1$ if $\bar{u} \in F$ and $\mu_{\bar{u}}(F) = 0$ if $\bar{u} \notin F$. The RUF $\mu_{\bar{u}}$ corresponds to a deterministic utility function $\bar{u}$. In the case of two prizes $(n+1=2)$ the RUF $\mu_{\bar{u}}$ is regular if $\bar{u} \neq (0,0)$. When there are more than two prizes $(n+1>2)$ then $\mu_{\bar{u}}$ is not regular irrespective of the choice of $\bar{u}$. This follows because we can find distinct lotteries x, y such that $\bar{u}x = \bar{u}y$ and therefore $\bar{u}$ does not have a unique maximizer in the decision problem $D = \{x, y\}$.

We conclude this section by showing that there is a one-to-one correspondence between regular RUFs and their maximizers.

Theorem 1: *(i) Every regular μ has a unique maximizer. (ii) For every RCR ρ there is at most one regular RUF μ such that ρ maximizes μ.*

Proof: See Appendix 8.1 $\square$

Let $\rho : \mathcal{D} \to \Pi$ be defined by $\rho^D(B) := \sum_{x \in D \cap B} \mu(N(D,x))$ for all $D \in \mathcal{D}$, $B \in \mathcal{B}$. Since μ is regular, ρ is a well defined RCR. This is the only RCR that satisfies $\rho^D(x) = \mu(N(D,x))$ for all D, x and therefore part (i) follows. For part (ii) it suffices to show that if μ and μ' are two regular RUFs with $\mu(N(D,x)) = \mu'(N(D,x)) = \rho^D(x)$ for all D, x then $\mu(F) = \mu'(F)$ for all $F \in \mathcal{F}$. We prove this result in section 8.1.

[12] The dimension of a set is the dimension of its affine hull.

3. Properties of Random Choice Rules

This section describes the properties of random choice rules that identify random utility models.

We endow $\mathcal{D}$ with the Hausdorff topology. The Hausdorff distance between D and D' is given by

$$d_h(D, D') := \max\{\max_{x \in D} \min_{x' \in D'} \|x - x'\|, \max_{x' \in D'} \min_{x \in D} \|x - x'\|\}$$

This choice of topology implies that when lotteries are added to D that are close to some $x \in D$ then the choice problem remains close to D. We endow Π with the topology of weak convergence. For any $D, D' \subset \mathcal{D}$ and $\lambda \in [0,1]$, let $\lambda D + (1-\lambda)D' := \{\lambda x + (1-\lambda)y \mid x \in D, y \in D'\}$. Note that if $D, D' \in \mathcal{D}$ then $\lambda D + (1-\lambda)D' \in \mathcal{D}$.

We consider two notions of continuity for RCRs. The weaker notion (*mixture continuity*) is analogous to von Neumann-Morgenstern's notion of continuity for preferences over lotteries.

Definition: *The RCR ρ is mixture continuous if $\rho^{\alpha D + (1-\alpha)D'}$ is continuous in α for all $D, D' \in \mathcal{D}$.*

The stronger notion of continuity requires that the choice rule be a continuous function of the decision problem.

Definition: *The RCR ρ is continuous if $\rho : \mathcal{D} \to \Pi$ is a continuous function.*

Continuity implies mixture continuity since $\alpha D + (1-\alpha)D'$ and $\beta D + (1-\beta)D'$ are close (with respect to the Hausdorff metric) whenever α and β are close. To see that continuity is stronger than mixture continuity suppose that D' is obtained by rotating D. Mixture continuity permits the probability of choosing x in D to be very different even if the angle of rotation is very small.

The next property is *monotonicity.* Monotonicity says that the probability of choosing an alternative x cannot increase as more options are added to the decision problem.

Definition: *The RCR ρ is monotone if $x \in D \subset D'$ implies $\rho^{D'}(x) \leq \rho^{D}(x)$.*

Monotonicity is the stochastic analogue of Chernoff's Postulate 4 or equivalently, Sen's condition α, a well-known consistency condition on deterministic choice rules. This condition says that if x is chosen from D then it must also be chosen from every subset of D that contains x. Hence, Chernoff's Postulate 4 is monotonicity for deterministic choice rules.[13]

Our random utility model restricts attention to von Neumann-Morgenstern utility functions. As a consequence, the corresponding random choice rules must also be linear. Linearity requires that the choice probabilities remain unchanged when each element x of the choice problem D is replaced with the lottery $\lambda x + (1-\lambda)y$ for some fixed y.

Definition: *The RCR ρ is linear if $\rho^{\lambda D+(1-\lambda)\{y\}}(\lambda x + (1-\lambda)y) = \rho^D(x)$ for all $x \in D, \lambda \in (0,1)$.*

Linearity is analogous to the independence axiom of von Neumann-Morgenstern theory. Note that this "version" of the independence axiom corresponds exactly to the version used in experimental settings. In the experimental setting, a group of subjects is asked to make a choice from a binary choice problem $D = \{x, x'\}$. Then the same group must choose from a second choice problem that differs from the first by replacing the original lotteries x, x' with $\lambda x + (1-\lambda)y$ and $\lambda x' + (1-\lambda)y$. Linearity requires that the frequency with which the lottery x is chosen is the same as the frequency with which the lottery $\lambda x + (1-\lambda)y$ is chosen.

The final condition on random choice rules requires that from each decision problem only extreme points are chosen. The extreme points of D are denoted $\text{ext}\, D$. Note that the extreme points of D are those elements of D that are unique maximizers of some utility function. Hence, x is an extreme point of D if $N^+(D,x) \neq \emptyset$.

Definition: *The RCR ρ is extreme if $\rho^D(\text{ext}\, D) = 1$.*

A decision-maker who maximizes expected utility can without any loss, restrict himself to extreme points of the decision problem. Moreover, a decision maker who maximizes a *regular* RUF must choose an extreme point with probability 1. Hence, extremeness is a necessary condition for maximizing a regular RUF.

[13] Monotonicity rules out "complementarities" as illustrated in the following example of a choice rule given by Kalai et al. (2001). An economics department hires only in the field that has the highest number of applicants. The rationale is that a popular field is active and competitive and hence hiring in that field is a good idea. In other words, the composition of the choice set itself provides information for the decision-maker. Monotonicity rules this out.

4. Main Results

In this section we present our results on regular RUF maximization. Theorems 2 and 3 below are our main results. Theorem 2 establishes that mixture continuity, monotonicity, linearity, and extremeness are necessary and sufficient for ρ to maximize a regular RUF. Theorem 3 shows that replacing mixture continuity with continuity yields necessary and sufficient conditions for maximizing a regular and countably additive RUF.

Theorem 2: *The RCR ρ is mixture continuous, monotone, linear and extreme if and only if there exists a regular RUF μ such that ρ maximizes μ.*

Proof: See section 8.2.

It follows from Theorems 1 and 2 that there is a one-to-one correspondence between mixture continuous, monotone, linear and extreme RCRs and the RUFs they maximize.

To see the intuition for the "only if" part of Theorem 2, assume that ρ maximizes some regular μ. Hence, $\rho^D(x) = \mu(N(D,x))$ for all D, x. The choice rule ρ is monotone since $N(D \cup \{y\}, x) \subset N(D,x)$ whenever $x \in D$; it is linear since $N(D,x) = N(\lambda D + (1-\lambda)\{y\}, \lambda x + (1-\lambda)y)$. If x is not an extreme point of D then $N^+(D,x) = \emptyset$ and therefore regularity of μ implies that $\mu(N(D,x)) = \mu(N^+(D,x)) = 0$. Hence, ρ is extreme. For the proof of mixture continuity, see section 8.2 of the Appendix.

Next, we briefly sketch the proof of the "if" part of Theorem 2. Lemma 1 shows that monotonicity, linearity and extremeness of ρ imply $\rho^D(x) = \rho^{D'}(y)$ whenever $N(D,x) = N(D',y)$. To get intuition for the proof of Lemma 1, consider the choice problems D, D' illustrated in Figure 1.

Insert Figure 1 here

Note that $K := N(D,x) = N(D',y)$. By linearity we can translate and "shrink" D' without affecting the choice probabilities. In particular, as illustrated in Figure 1, we may translate D' so that the translation of y coincides with x and we may shrink D' so that it "fits into" D (as illustrated by the decision problem $\lambda D' + (1-\lambda)\{z\}$). Monotonicity together with the fact that only extreme points are chosen implies that the probability of choosing y from D' is at least as large as the probability of choosing x from D. Then, reversing the role of D and D' proves Lemma 1.

Finite additivity is proven in Lemma 4. To understand the argument for finite additivity consider the decision problems D, D', D'' as illustrated in Figure 2.

Insert Figure 2 here

Note that $N(D,x) = N(D',y) \cup N(D'',z)$. For a regular μ we have $\mu(N^+(D,x)) = \mu(N(D,x))$ for all (D,x) and hence we must show that $\mu(N(D,x)) = \mu(N(D',y)) + \mu(N(D'',z))$ which is equivalent to $\rho^D(x) = \rho^{D'}(y) + \rho^{D''}(z)$. Consider the decision problems $D_\lambda := (1-2\lambda)D + \lambda D' + \lambda D''$ as illustrated in Figure 2. By Lemma 1, we know that $\rho^{D_\lambda}(y_\lambda) = \rho^{D'}(y), \rho^{D_\lambda}(z_\lambda) = \rho^{D''}(z)$. Mixture continuity implies that $\rho^{D_\lambda}(B) \to \rho^D(x)$ for any Borel set B such that $B \cap D = \{x\} = \bar{B} \cap D$, where $\bar{B}$ denotes the closure of B. As $\lambda \to 0$ we have $y_\lambda \to x$ and $z_\lambda \to x$. This in turn implies that $\rho^{D_\lambda}(y_\lambda) + \rho^{D_\lambda}(z_\lambda) = \rho^{D'}(y) + \rho^{D''}(z) = \rho^D(x)$ as desired.

Next, we characterize the behavior associated with countably additive RUFs. Theorem 3 proves that continuity, monotonicity, linearity and extremeness are necessary and sufficient conditions for regular random utility maximization.

Theorem 3: *The RCR ρ is continuous, monotone, linear and extreme if and only if there exists a regular, countably additive RUF μ such that ρ maximizes μ.*

Proof: See section 8.3.

Theorem 3 implies that regular RUFs that are not countably additive lead to behavior that is not continuous. To see this, suppose ρ maximizes some regular but not countably additive μ. By Theorem 2, ρ must be monotone, linear and extreme. If ρ were continuous then (by Theorem 3) it would maximize some regular and countably additive μ'. But, by Theorem 1(ii), ρ cannot maximize two distinct regular RUFs.

For continuous ρ, extremeness can replaced with a weaker condition. Consider the choice problem D and a lottery x such that $x \in O$ for some open set O with $O \subset \operatorname{conv} D$. Clearly, the lottery x is not an optimal choice from D for *any* utility function $u \in U$, except

$u = (0, \dots, 0)$. Therefore x cannot be chosen from D with positive probability if the agent maximizes some regular RUF. Let $\operatorname{bd} X$ denote the boundary of the set $X \subset I\!R^{n+1}$.

Definition: *The RCR ρ is undominated if $\rho^D(\operatorname{bd}\operatorname{conv} D) = 1$ whenever $\dim D = n$.*

Undominated choice rules place zero probability on $x \in D$ such that any lottery in a neighborhood of x can be attained by a linear combination of lotteries in D. Such lotteries are never optimal for linear preferences unless the preference is indifferent among all options in P.

Theorem 4: *The RCR ρ is continuous, monotone, linear and undominated if and only if there exists a regular, countably additive RUF μ such that ρ maximizes μ.*

Proof: See section 8.4. □

To prove Theorem 4, we show that a continuous RCR is extreme if and only if it is undominated. Then the result follows from Theorem 3.

5. Counterexamples

In this section, we provide examples that show that none of the assumptions in Theorems 2, 3, and 4 are redundant. Example 4 provides a RCR that is continuous (hence mixture continuous), linear and extreme (hence undominated) but not monotone. This shows that monotonicity cannot be dispensed with in Theorems 2, 3 and 4.

Example 4: Let $n + 1 = 2$. Hence, P can be identified with the unit interval and $x \in P$ is the probability of getting prize 2. For $D \in \mathcal{D}$, let $\underline{m}(D)$ denote the smallest element in D, $\overline{m}(D)$ denote the largest element in D, and define

$$a(D) := \sup\{x - y \mid \underline{m}(D) \leq y \leq x \leq \overline{m}(D), (y, x) \cap D = \emptyset\}$$

Hence, $a(D)$ is the length of the largest open interval that does not intersect D, but is contained in the convex hull of D. If $D = \{x\}$ then $\rho^D(x) = 1$. If D is not a singleton, let

$$\rho^D(\underline{m}(D)) = \frac{a(D)}{\overline{m}(D) - \underline{m}(D)}$$
$$\rho^D(\overline{m}(D)) = 1 - \rho^D(\underline{m}(D))$$

and $\rho^D(x) = 0$ for $x \notin \{\underline{m}(D), \overline{m}(D)\}$. Then, ρ is continuous (hence mixture continuous), linear, extreme, (hence undominated) but not monotone.

Example 5 provides a RCR that is continuous (hence mixture continuous), monotone and linear but not undominated (and hence not extreme). This shows that the requirement that the choice rule is undominated cannot be dropped in Theorem 4 and the requirement that the choice rule is extreme cannot be dropped in Theorems 2 and Theorem 3.

Example 5: Let $n+1 = 2$ and let $x \in [0, 1]$ denote the probability of getting prize 2. For any $D = \{x_1, \dots, x_m\}$, where $x_1 < x_2 < \dots < x_m$, let

$$\rho^D(x_1) = \begin{cases} 1 & \text{if } m = 1 \\ 1/2 & \text{otherwise.} \end{cases}$$

For $k > 1$, let

$$\rho^D(x_k) = \frac{x_k - x_{k-1}}{2(x_m - x_1)}$$

Then, ρ is continuous, monotone and linear but not undominated (hence not extreme).

Example 6 provides a RCR that is continuous (hence mixture continuous), extreme (and hence undominated) and monotone but not linear. This shows that linearity cannot be dropped in Theorems 2, 3, and 4.

Example 6: Let $n + 1 = 2$ and let $x \in [0, 1]$ denote the probability of getting prize 2. As in Example 4, let $\underline{m}(D)$ and $\overline{m}(D)$ be the smallest and largest elements in D. Let $\rho^D(x) = 1$ for $D = \{x\}$. If D is not a singleton then

$$\rho^D(\overline{m}(D)) = \overline{m}(D)$$
$$\rho^D(\underline{m}(D)) = 1 - \overline{m}(D)$$

and $\rho^D(x) = 0$ for $x \notin \{\underline{m}(D), \overline{m}(D)\}$. Then, ρ is continuous, monotone and extreme but not linear.

Example 7 provides a RCR that is monotone, linear, and extreme (hence undominated) but not mixture continuous (and hence is not continuous). This shows that mixture continuity cannot be dispensed with in Theorem 2 and continuity cannot be dispensed with in Theorems 3 and 4.

Example 7: The RCR ρ takes on the values $0, 1/2$ and 1. If $N(D,x) = U$ and hence the decision problem is a singleton, then $\rho^D(x) = 1$. There are three cases in which ρ takes on the value $1/2$:

$$\rho^D(x) = 1/2 \text{ if } N(D,x) \text{ is a halfspace or}$$
$$\text{if there is } \epsilon > 0 \text{ such that } (1+\epsilon, -1, 0), (1, -1, 0) \in N(D,x) \text{ or}$$
$$\text{if there is } \epsilon > 0 \text{ such that } (-1, 1+\epsilon, 0), (-1, 1, 0) \in N(D,x)$$

In all other cases, $\rho^D(x) = 0$.

To see that this ρ constitutes a well defined RCR note that $N(D,x)$ is a halfspace if and only if D is one-dimensional and x is an extreme point of D. Clearly, a one-dimensional decision problem has two extreme points. If D is 2-dimensional then $\rho^D(x) = 1/2$ if x is the maximizer of $(1,-1,0)$ in D with the largest first coordinate or if x is the maximizer of $(-1,1,0)$ in D with the largest second coordinate.

This RCR is extreme by definition. It is linear because the probability of choosing x from D depends only on the set $N(D,x)$ which is invariant to linear translations of D. To see that the choice rule is monotone, note that the construction ensures that the probability of choosing x from D is monotone in $N(D,x)$. That is, $N(D,x) \subset N(D',y)$ implies $\rho^D(x) \leq \rho^{D'}(y)$. Since $N(D \cup \{y\}, x) \subset N(D,x)$ it follows that ρ is monotone. It remains to show that ρ is not mixture continuous.

Let $D = \{(1/4, 1/2, 1/4), (1/2, 1/4, 1/4)\}$ and let $D' = \{(3/8, 3/8, 1/4), (1/8, 1/8, 3/4)\}$. For $\lambda > 0$ the agent chooses from $\lambda D + (1-\lambda)D'$ either $\lambda(1/4, 1/2, 1/4) + (1-\lambda)(3/8, 3/8, 1/4)$ or $\lambda(1/2, 1/4, 1/4) + (1-\lambda)(3/8, 3/8, 1/4)$, each with probability $1/2$. For $\lambda = 0$ the agent chooses $(3/8, 3/8, 1/4)$ or $(1/8, 1/8, 3/4)$ each with probability $1/2$. Clearly, this violates mixture continuity at $\lambda = 0$.

6. Non-regular Random Utility

Not all RUFs are regular. For non-regular RUFs we cannot identify a unique maximizing RCR since there is a positive probability of a "tie" in some decision problems. More precisely, for some decision problem D there is a positive probability of choosing a utility function that does not have a unique maximizer in D.

To deal with non-regular RUFs we introduce tie-breakers. Suppose that the agent with RUF μ faces the decision problem D. Assume that in order to eliminate ties, the decision-maker chooses two utility functions (u, v) according to some measure η. If the set of maximizers of u in D (denoted $M(D, u)$) is a singleton, then the agent chooses the unique element of $M(D, u)$. Otherwise, the agent chooses an element of $M(D, u)$ that maximizes v; that is, an element of $M(M(D, u), v)$. If η is a product measure $\eta = \mu \times \hat{\mu}$ and $\hat{\mu}$ is regular then it is clear that this procedure will lead to a unique choice with probability one. In this case, the choice of v is independent of the choice of u and the regularity of $\hat{\mu}$ ensures that $M(M(D, u), v)$ is a singleton with probability 1. It turns out that independence is not necessary for a tie-breaker to generate a unique choice as long as the marginal on the second coordinate is a regular RUF. Therefore, our model does not restrict to product measures and allows for correlation.

In order to describe the lexicographic procedure above formally, we need to describe a measure on the set $U \times U$. Let $\mathcal{F}^2$ denote the smallest field that contains $\mathcal{F} \times \mathcal{F}$. The marginals η_i of η are defined by:

$$\eta_1(F) = \eta(F, U)$$

$$\eta_2(F) = \eta(U, F)$$

for all $F \in \mathcal{F}$.

Definition: *(i) The measure η on $\mathcal{F}^2$ is a tie-breaker if η_2 is regular. (ii) The measure η is a tie-breaker for μ if $\eta_1 = \mu$ and η_2 is regular.*

Let $N_l(D, x) = \{(u, v) \,|\, x \in M(M(D, u), v)\}$. Hence, $(u, v) \in N_l(D, x)$ if and only if x is a lexicographic maximizer of (u, v) in D. We show in Lemma 8 that $N_l(D, x) \in \mathcal{F}^2$ for all D, x. A random choice rule ρ maximizes the tie-breaker η if the probability of choosing

x in D is equal to the probability of choosing (u, v) in $N_l(D, x)$. The random choice rule maximizes the (not necessarily regular) RUF μ if ρ maximizes a tie-breaker for μ.

Definition: *(i) The RCR ρ maximizes the tie-breaker η if $\rho^D(x) = \eta(N_l(D, x))$ for all D, x. (ii) The RCR ρ maximizes the RUF μ if ρ maximizes a tie-breaker for μ.*

Part (ii) of the definition above applies to regular and non-regular RUFs. To see this note that

$$\mu(N^+(D, x)) \leq \eta(N_l(D, x)) \leq \mu(N(D, x))$$

for all D, x. The first inequality follows from the fact that if x is the unique maximizer of u in D then x is the lexicographic maximizer of (u, v) for all $v \in U$. The second inequality follows from the fact that any lexicographic maximizer of (u, v) is a maximizer of u. Hence, if η is tie-breaker for the regular RUF μ and $\rho^D(x) = \eta(N_l(D, x))$ for all D, x then $\rho^D(x) = \mu(N(D, x))$ for all D, x. Therefore, ρ maximizes μ.

Theorem 5 demonstrates that tie-breakers have a unique maximizing RCR. Moreover, this RCR is monotone, linear, mixture continuous and extreme.

Theorem 5: *Every tie-breaker is maximized by a unique RCR. If the RCR ρ maximizes a tie-breaker then ρ is monotone, linear, mixture continuous and extreme.*

Proof: See Appendix, section 8.5.

To prove part (i) of Theorem 5 we show that the function $\rho : \mathcal{D} \to \Pi$ defined as $\rho^D(B) := \sum_{x \in D \cap B} \eta(N_l(D, x))$ for all $D \in \mathcal{D}, B \in \mathcal{B}$ is a well defined random choice rule. To prove this we establish that $M(M(D, u), v)$ is a singleton with probability 1 when η_2 is regular. Clearly, this is the only RCR that satisfies $\rho^D(x) = \eta(N_l(D, x))$ and hence uniqueness follows. Part (ii) of Theorem 5 is analogous to the "only if" part of Theorem 2.

Example 3 above described the RUF that corresponds to a deterministic utility function $\bar{u}$. Below, we provide an example of a tie-breaker for this RUF.

Example 3 continued: There are three prizes ($n + 1 = 3$). Consider the RUF $\mu_{\bar{u}}$ which assigns probability 1 to the utility function $\bar{u} \neq (0, 0, 0)$. An example of a tie-breaker for

$\mu_{\bar{u}}$ is the measure $\eta = \mu_{\bar{u}} \times \mu$ where μ is the uniform RUF defined in Example 2. The tie-breaker η is maximized by the following RCR ρ. If $M(D, \bar{u}) = \{x\}$ and hence $\bar{u}$ has a unique maximizer in D then $\rho^D(x) = 1$. If $M(D, \bar{u})$ is not a singleton then the convex hull of $M(D, \bar{u})$ is a line segment. In that case, ρ^D assigns probability 1/2 to each end-point of this line segment.

Let $\hat{\mu}$ be any regular random utility.[14] Then, the product measure $\eta := \mu \times \hat{\mu}$ is a tie-breaker for μ. By Theorem 5, every tie-breaker has a maximizer and therefore it follows that every non-regular RUF has a maximizer. For a non-regular RUF the choice of a tie-breaker affects behavior and therefore there are multiple maximizing random choice rules. In contrast, regular random utilities have a unique maximizer. Theorem 6 summarizes these facts.

Theorem 6: *(i) Every RUF μ has a maximizer. (ii) A RUF has a unique maximizer if and only if it is regular.*

Proof: See Appendix, section 8.6.

Theorem 5 shows that the generalization of RUF maximization to non-regular RUFs preserves the properties identified in section 4. If ρ is a maximizer of some (not necessarily regular) RUF then it satisfies monotonicity, linearity, mixture continuity and extremeness. Therefore we can apply Theorem 2 to conclude that ρ must also maximize some *regular* RUF μ'.

Theorem 7: *If the RCR ρ maximizes some RUF then ρ maximizes a regular RUF.*

Proof: Follows from Theorem 5 and Theorem 2. □

Consider a non-regular RUF μ. Let η be a tie-breaker for μ and let ρ be the maximizer of η. By Theorem 7 the RCR ρ also maximizes a regular RUF μ'. Hence,

$$\mu'(N(D, x)) = \eta(N_l(D, x)) = \rho^D(x)$$

for all $D \in \mathcal{D}$ and $x \in D$. We call this μ' a dilation of μ. A dilation μ' of μ satisfies

$$\mu(N^+(D, x)) \leq \mu'(N(D, x)) \leq \mu(N(D, x))$$

[14] Lemma 3 proves the existence of a regular RUF.

Intuitively, a dilation of μ takes probability mass from lower dimensional subsets of U and (with the aid of the tie-breaker) spreads it over adjacent $n-$dimensional sets. Below, we illustrate a dilation of the RUF in Example 3.

Example 3 continued: There are three prizes ($n+1=3$). Consider the RUF $\mu_{\bar{u}}$ which assigns probability 1 to the utility function $\bar{u} \neq (0,0,0)$. The following regular random utility μ' is a dilation of $\mu_{\bar{u}}$. Recall that for any $u, v \in U$, $F_{uv} := \{\alpha u + \beta v \,|\, \alpha, \beta > 0\}$. Let $\mu'(F_{uv}) = 1$ if $u \neq \lambda v$ (and hence F_{uv} is two-dimensional) and $\bar{u}$ is in the relative interior of F_{uv}. Let $\mu'(F_{uv}) = 1/2$ if $u \neq \lambda v$ for $\lambda \in I\!R$ and $\bar{u}$ is on the boundary of F_{uv}. That is, $\bar{u} = \lambda u$ or $\bar{u} = \lambda v$ for some $\lambda > 0$. In all other cases, $\mu'(F_{uv}) = 0$. In particular, every one-dimensional subset of U has $\mu'-$measure 0 and therefore, μ' is regular. The RUF μ' is maximized by same RCR as the uniform tie-breaker described above: If $M(D, \bar{u}) = \{x\}$ then $\bar{u}$ is in the interior of $N(D, x)$. Therefore, $\rho^D(x) = \mu'(N(D, x)) = 1$ in this case. If $M(D, x)$ is not a singleton then ρ^D assigns probability 1/2 to each extreme point of $M(D, \bar{u})$. (Note that $M(D, \bar{u})$ has at most two extreme points).

Theorem 8 shows that except for the case of complete indifference a dilation of a non-regular random utility is not countably additive. In other words, ties cannot be broken in a manner that preserves countable additivity. Recall that $o = (0, \ldots, 0)$ denotes the utility function that is indifferent between all prizes.

Theorem 8: *If μ' is a dilation of some non-regular μ such that $\mu(o) = 0$ then μ' is not countably additive.*

Proof: See section 8.7. □

Theorem 8 is closely related to Theorem 3 above. Theorem 3 implies that a maximizer of a regular, countably additive RUF is continuous. In the proof of Theorem 8 we show that a maximizer of a non-regular RUF μ with $\mu(o) = 0$ must fail continuity and therefore, Theorem 3 implies Theorem 8. We illustrate Theorem 8 by demonstrating that the dilation in Example 3 above is not countably additive.

Example 3 continued: In Example 3 above, we define a dilation μ' of the random utility $\mu_{\bar{u}}$. To see that μ' is not countably additive, let $v \neq \lambda \bar{u}$ and let v_n be on the relative interior of the line segment connecting $\bar{u}$ and v. Choose the sequence v_n so that it converges to $\bar{u}$. Note that $\mu'(F_{vv_n}) = 0$ for all n yet $\mu'(\bigcup_n F_{vv_n}) = 1/2$. Hence, the dilation μ' is not countably additive. Note, that the original random utility $\mu_{\bar{u}}$ is countably additive.

We can interpret the results in this section as a justification for restricting attention to regular RUFs. When tie-breakers are used to resolve the ambiguity associated with non-regular RUFs, the resulting behavior maximizes some regular random utility. In this sense, the restriction to regular RUFs is without loss of generality. However, applying a tie-breaker to a non-regular μ typically results in a regular RUF (i.e., dilation of μ) that fails countable additivity.

7. Appendix A: Preliminaries

In this section, we define the concepts and state results from convex analysis that are used in the proofs. Throughout this section, all points and all sets are in $n-$dimensional Euclidian space $I\!R^n$. For any $x \in I\!R^n$ we use x^i to denote the i'th coordinate of x and o to denote the origin. If $x = \sum_i \lambda_i x_i$ with $\lambda_i \in I\!R$ for all $i = 1, \ldots, k$ then x is a (linear) combination of the $x_1, \ldots, x_k$. If $\lambda_i \geq 0$, then x is a positive combination, if $\sum_i \lambda_i = 1$ then x is an affine combination and if $\lambda_i \geq 0, \sum_i \lambda_i = 1$ then x is a convex combination of $x_1, \ldots, x_k$. For any set A, we let $\operatorname{aff} A$ ($\operatorname{pos} A$, $\operatorname{conv} A$) denote the set of all affine (positive, convex) combinations of points in A. The set A is affine (a cone, convex) if $A = \operatorname{aff} A$ ($A = \operatorname{pos} A$, $A = \operatorname{conv} A$). The interior of a set A is denoted $\operatorname{int} A$. The relative interior of A, denoted $\operatorname{ri} A$, is the interior of A in the relative topology of $\operatorname{aff} A$. The dimension of the affine set A is the dimension of the subspace $A - x$ for $x \in A$. The dimension of any set A, denoted $\dim A$ is the dimension of the affine hull of A.

The open ball with radius ϵ and center x is denoted $B_\epsilon(x)$. The unit sphere is denoted $S = \{u \in I\!R^n \,|\, \|u\| = 1\}$, and the n-dimensional cube is denoted $E^* := \{u \in I\!R^n \,|\, |u^i| = 1 \text{ for some } i \text{ and } u_j = 0 \forall j \neq i\}$. We use e to denote the vector of 1's in $I\!R^n$.

A set of the form $K(u, \alpha) := \{z \in I\!R^n \,|\, u \cdot z \leq \alpha\}$ for $u \neq o$, is called a halfspace. For $x \neq o$, the set $H(x, \alpha) := K(x, \alpha) \cap K(-x, -\alpha)$ is called a hyperplane. A set A is polyhedral (or is a polyhedron) if it can be expressed as the intersection of a finite collection of halfspaces. Obviously, polyhedral sets are closed and convex. The set A is a polytope if

$A = \operatorname{conv} B$ for some finite set B. Every polytope is a polyhedron and a polyhedron is a polytope if and only if it is bounded. A cone is polyhedral if and only if it can be expressed as $\operatorname{pos} C$ for some finite C. Let $\mathcal{K}^*$ denote the set of all polyhedral cones and $\mathcal{K}$ denote the set of all *pointed* polyhedral cones; that is, the elements of $\mathcal{K}$ are those elements of $\mathcal{K}^*$ that have o as an extreme point.

For the polyhedron A and $x \in A$, the set $N(A,x) = \{u \in I\!R^n \,|\, u \cdot y \leq u \cdot x \,\forall y \in A\}$ is called the normal cone to A at x. When D is a finite set, we write $N(D,x)$ rather than $N(\operatorname{conv} D, x)$. The set $N(A,x)$ is polyhedral whenever A is polyhedral. If K is a polyhedral cone then $L = N(K,o)$ is called the polar cone of K and satisfies $K = N(L,o)$.

A face A' of a polyhedron A is a nonempty convex subset of A such that if $\alpha x + (1-\alpha)y \in A'$ for some $x, y \in A, \alpha \in (0,1)$ then $\{x,y\} \subset A'$. Let $F(A)$ denote the set of all nonempty faces of the nonempty polyhedron A and let $F^0(A) := \{\operatorname{ri} F \,|\, F \in F(A)\}$. Let $F(A,u) = \{x \in A \,|\, u \cdot x \geq u \cdot y \,\forall y \in A\}$. For $A \neq \emptyset$, the set $F(A,u)$ is called an exposed face of A. Clearly every exposed face of A is a face of A. A singleton set is a face of A if and only if it is an extreme point of A. For any polyhedron A, A itself is a face of A and it is the only face $F \in F(A)$ such that $\dim(F) = \dim(A)$. Every face of a polyhedron is a polyhedron; A'' is a face of A' and A' is a face of the polyhedron A implies A'' is a face of A and finally, every face of a polyhedron is an exposed face (hence $F(A) = \bigcup_{u \in I\!R^n} F(A,u)$).

Proposition 1: *Let A, A' be two polyhedra and $x, y \in A$. Then:* (i) $\dim A = n$ *if and only if* $o \in \operatorname{ext} N(A,x)$. (ii) $L = N(A,x)$ *implies* $N(L,o) = \operatorname{pos}(A - \{x\})$ (iii) $x \in \operatorname{ext} A$ *if and only if* $\dim N(A,x) = n$. (iv) $\operatorname{ri} N(A,x) \cap \operatorname{ri} N(A,y) \neq \emptyset$ *implies* $N(A,x) = N(A,y)$. (v) $\operatorname{ri} A \cap \operatorname{ri} A' \neq \emptyset$ *implies* $\operatorname{ri} A \cap \operatorname{ri} A' = \operatorname{ri}(A \cap A')$.

Proof: (i) If $o \notin \operatorname{ext} N(A,x)$, then $\{o\}$ is not a face of $N(A,x)$ and therefore there exists $u \neq o$ such that $u, -u \in N(A,x)$. Hence, $A \subset \{z \,|\, u \cdot z \leq u \cdot x\} \cap \{z \,|\, -u \cdot z \leq -u \cdot x\}$. But $\{z \,|\, u \cdot z \leq u \cdot x\} \cap \{z \,|\, -u \cdot z \leq -u \cdot x\}$ has dimension $n-1$ and therefore, $\dim A < n$. The argument can be reversed. (ii) Let $L = N(A,x)$ and $K = \operatorname{pos}(A - \{x\})$. Clearly, K is a polyhedral cone and $L = N(K,o)$ is its polar cone. Hence, $N(L,o) = K$ as desired. (iii) Let $L = N(\operatorname{pos}(A - \{x\}), o)$. Then, $N(A,x) = N(A - \{x\}, o) = L$ and $N(L,o) = \operatorname{pos}(A - \{x\})$. Therefore, $x \in \operatorname{ext} A$ iff $o \in \operatorname{ext} N(L,o)$. By part (i), $o \in \operatorname{ext} N(L,o)$ iff $\dim L = n$ and therefore $x \in \operatorname{ext} A$ iff $\dim N(A,x) = n$. (iv) Schneider (1993) notes this after stating Lemma 2.2.3. (v) Theorem 6.5 of Rockafeller (1970) proves the same result for all convex sets. □

Proposition 2: *(i) Let A be a polytope or polyhedral cone. Then, $x, y \in \operatorname{ri} F$ for some $F \in F(A)$ implies $N(A,x) = N(A,y)$. (ii) Let A be a polytope with $\dim A = n$ and $u \neq o$. Then, $x \in \operatorname{ri} F(A,u)$ implies $u \in \operatorname{ri} N(A,x)$.*

Proof: Suppose $x, y \in \operatorname{ri} F$ for some $F \in F(A)$. If $u \in N(A,x)$ then $x \in F(A,u)$. Since $y \in \operatorname{ri} F$ and $x \in F$, there exists $\lambda > 1$ such that $z := \lambda x + (1-\lambda)y \in A$. Hence, $x = \alpha y + (1-\alpha)z$ for some $\alpha \in (0,1)$. Since $F(A,u)$ is a face of A, we conclude that $y \in F(A,u)$ and therefore $u \in N(A,y)$. By symmetry, we have $N(A,x) = N(A,y)$. In Schneider (1993) page 99, (ii) is stated as (2.4.3), a consequence of Theorem 2.4.9. □

Proposition 3: *If $D_i \in \mathcal{D}$ for $i = 1, \ldots, m$ then*

$$N(D_1 + \cdots + D_m, \sum_i x_i) = \bigcap_{i=1}^m N(D_i, x_i)$$
$$N_l(D_1 + \cdots + D_m, \sum_i x_i) = \bigcap_{i=1}^m N_l(D_i, x_i)$$

Proof: Follows from elementary arguments. □

Proposition 4: *If $K \in \mathcal{K}^*$ then $K = N(D,o)$ for some $D \in \mathcal{D}$ with $o \in D$.*

Proof: Let $A = N(K,o) \cap \operatorname{conv} E^*$. Clearly, A is bounded and polyhedral. That $N(A,o) = N(N(K,o),o)$ is obvious. Since $N(N(K,o),o) = K$, $\operatorname{ext} A \cup \{o\}$ is the desired set. □

Let $\mathcal{N}(A) := \{N(A,x) \,|\, x \in A\}$ and let $\mathcal{N}^0(A) := \{\operatorname{ri} K \,|\, K \in \mathcal{N}(A)\}$. A finite collection of subsets $\mathcal{P}$ of X is called a *partition* (of X) if $\emptyset \notin \mathcal{P}$, $A, B \in \mathcal{P}$, $A \cap B \neq \emptyset$ implies $A = B$, and $\bigcup_{A \in \mathcal{P}} A = X$. If $\mathcal{P}$ is partition of X and $\emptyset \neq Y \subset X$ then we say that $\mathcal{P}$ *measures* Y if there exists $A_i \in \mathcal{P}$ for $i = 1, \ldots, m$ such that $\bigcup_{i=1}^m A_i = Y$. Note that the partition $\mathcal{P}$ measures Y if and only if $A \in \mathcal{P}$, $A \cap Y \neq \emptyset$ implies $A \subset Y$. We say that the partition $\mathcal{P}$ *refines* $\mathcal{P}'$, if $\mathcal{P}$ measures each element of $\mathcal{P}'$.

Proposition 5: *(i) For any nonempty polyhedron A, $F^0(A)$ is a partition of A and measures each element of $F(A)$. (ii) For any polytope A such that* $\dim(A) = n$, *$\mathcal{N}^0(A)$ is a partition of $I\!R^n$.*

Proof: (i) That $F^0(A)$ is a partition of A follows from the fact that the set of relative interiors of faces of any closed, convex set is a pairwise disjoint cover i.e., a decomposition of A, (Theorem 2.1.2 of Schneider (1993)) and the fact that a polyhedron has a finite number of faces. Then, suppose $B \in F(A)$, $H \in F^0(A)$ and $B \cap H \neq \emptyset$. Since any face of $B \in F(A)$ is also a face of A and $F^0(B)$ is a partition of B, we can express B as $\bigcup_{i=1}^m H_i$ for $H_1, \ldots, H_m \in F^0(A)$. But since $F^0(A)$ is a partition, it follows that $H_i \cap H \neq \emptyset$ implies $H = H_i$. Hence, $F^0(A)$ measures each element of $F(A)$.

(ii) For any u the face $F(A, u)$ is a non-empty convex set. Therefore, $\mathrm{ri}\, F(A, u)$ is non-empty (Theorem 1.1.12 of Schneider (1993)). By Proposition 2(ii), $x \in \mathrm{ri}\, F(A, u)$ implies $u \in \mathrm{ri}\, N(A, x)$ and hence $u \in \bigcup_{K \in \mathcal{N}^0(A)} K$. It follows that $I\!R^n \subset \bigcup_{K \in \mathcal{N}^0(A)} K$. To complete the proof we must show that $K, K' \in \mathcal{N}^0$ and $K \cap K' \neq \emptyset$ implies $K = K'$. Suppose, $\mathrm{ri}\, N(A, x) \cap \mathrm{ri}\, N(A, y) \neq \emptyset$. Then, for $u \in \mathrm{ri}\, N(A, x) \cap \mathrm{ri}\, N(A, y)$ Proposition 2(ii) yields $x, y \in \mathrm{ri}\, F(A, u)$. But then Proposition 2$(i)$ establishes $N(A, x) = N(A, y)$ and therefore $\mathrm{ri}\, N(A, x) = \mathrm{ri}\, N(A, y)$. Hence, $\mathcal{N}^0$ is a partition. $\square$

Let $\mathcal{F}$ be the smallest field that contains $\mathcal{K}^*$ and let $\mathcal{H} := \{\mathrm{ri}\, K \mid K \in \mathcal{K}\} \cup \emptyset$. A collection of subsets $\mathcal{P}$ of X is called a *semiring* if $\emptyset \in \mathcal{P}$, $A, B \in \mathcal{P}$ implies $A \cap B \in \mathcal{P}$, and $A, B \in \mathcal{P}$ and $B \subset A$ implies there exists disjoint sets $A_1, \ldots, A_m \in \mathcal{P}$ such that $\bigcup_i A_i = A \backslash B$.

Proposition 6: *(i) $\mathcal{H}$ is a semiring. (ii) $\mathcal{F} = \{\cup_{i=1}^m H_i \mid H_i \in \mathcal{H}$ for $i = 1, \ldots, m\}$.*

Proof: (i) First, we show that $H, H' \in \mathcal{H}$ implies $H \cap H'$ in $\mathcal{H}$. Let $H = \mathrm{ri}\, K$ and $H' = \mathrm{ri}\, K'$ for $K, K' \in \mathcal{K}$. Note that $o \in \mathrm{ext}\, K \cap \mathrm{ext}\, K'$ and hence $o \in \mathrm{ext}(K \cap K')$. If $H \cap H' = \emptyset$, we are done. Otherwise, by Proposition 1(v), $H \cap H' = \mathrm{ri}(K \cap K') \in \mathcal{H}$ as desired.

Next, we show that for all polytopes A, A' such that $\dim(A + A') = n$, $\mathcal{N}^0(A + A')$ is a partition that measures each element of $\mathcal{N}^0(A)$ and by symmetry of $\mathcal{N}^0(A')$. Proposition 5(ii), $\mathcal{N}^0(A + A')$ is a partition of $I\!R^n$. Recall that $\mathcal{N}^0(A + A')$ refines $\mathcal{N}^0(A)$ if for each $H \in \mathcal{N}^0(A)$ and $H'' \in \mathcal{N}^0(A+A')$, $H \cap H'' \neq \emptyset$ implies $H'' \subset H$. Hence, assume $H = \operatorname{ri} N(A,x)$ for some $x \in A$, $H'' = \operatorname{ri} N(A+A', y+x')$ for some $y \in A, x' \in A'$ and $\operatorname{ri} N(A,x) \cap \operatorname{ri} N(A+A', y+x') \neq \emptyset$. Then, by Propositions 1(v) and 3,

$$\begin{aligned}\emptyset \neq H'' \cap H &= \operatorname{ri} N(A+A', y+x') \cap \operatorname{ri} N(A,x)\\ &= \operatorname{ri}[N(A+A', y+x') \cap N(A,x)]\\ &= \operatorname{ri} N(A+A+A', x+y+x')\end{aligned}$$

Since A is a convex set, $N(A+A+A', x+y+x') = N(A+A', \frac{x+y}{2}+x') \in \mathcal{N}(A+A')$. It follows that $\operatorname{ri} N(A+A', \frac{x+y}{2}+x') \cap H'' \neq \emptyset$ and therefore, by Proposition 1(iv), $\operatorname{ri} N(A+A', \frac{x+y}{2}+x') = H''$, establishing $H'' \cap H = H''$ (i.e., $H'' \subset H$) as desired.

Assume that $H, H' \in \mathcal{H}$ such that $H' \subset H$. Hence, by Proposition 4, $H \in \mathcal{N}^0(A)$ and $H' \in \mathcal{N}^0(A')$ for some polytopes A, A'. By Proposition 1(i) each of these polytopes and hence $A+A'$ has dimension n. Hence, $\mathcal{N}^0(A+A')$ refines both $\mathcal{N}^0(A)$ and $\mathcal{N}^0(A')$ and therefore measures $H \backslash H'$ proving that $\mathcal{H}$ is semiring.

(ii) We first show that $\mathcal{F} \subset \{\cup_{i=1}^m H_i \,|\, H_i \in \mathcal{H} \text{ for } i = 1, \ldots, m\}$. Clearly, the set of all finite unions of elements of a semiring is a field. Hence, $\{\cup_{i=1}^m H_i \,|\, H_i \in \mathcal{H} \text{ for } i = 1, \ldots, m\}$ is a field. Let $K \in \mathcal{K}$, then $F(K) \subset \mathcal{K}$ and hence $F^0(K) \subset \mathcal{H}$. By Proposition 5($i$), $\bigcup_{H \in F^0(K)} H = K$ and hence $\{\cup_{i=1}^m H_i \,|\, H_i \in \mathcal{H} \text{ for } i = 1, \ldots, m\}$ contains $\mathcal{K}$. Let $K \in \mathcal{K}^*$. Then, by Proposition 4, there exists A, x such that $N(A,x) = K$. Since $\bigcup_{\text{ext } B} N(B,x) = I\!R^n$, Proposition 3 implies $\bigcup_{y \in \text{ext } E^*} N(A+E^*, x+y) = N(A,x)$. Since $\dim(A+E^*) = n$, by Proposition 1(i), each $N(A+E^*, x+y) \in \mathcal{K}$. Since, $\{\cup_{i=1}^m H_i \,|\, H_i \in \mathcal{H} \text{ for } i = 1, \ldots, m\}$ is a field, we conclude $K \in \{\cup_{i=1}^m H_i \,|\, H_i \in \mathcal{H} \text{ for } i = 1, \ldots, m\}$ and hence $\mathcal{F} \subset \{\cup_{i=1}^m H_i \,|\, H_i \in \mathcal{H} \text{ for } i = 1, \ldots, m\}$.

Since $\mathcal{F}$ is a field, to show that $\{\cup_{i=1}^m H_i \,|\, H_i \in \mathcal{H} \text{ for } i = 1, \ldots, m\} \subset \mathcal{F}$, it is enough to show that $H \in \mathcal{F}$ for all $H \in \mathcal{H}$. Let $H = \operatorname{ri} K$ for some $K \in \mathcal{K}$. Since $\mathcal{K}^* \subset \mathcal{F}$, $K \in \mathcal{F}$. By Proposition 5(i), $F^0(K)$ is a partition of K that measures each face of K. Hence,

$$\begin{aligned}K &= \operatorname{ri} K \cup (\bigcup_{F \in F(K), F \neq K} F)\\ \emptyset &= \operatorname{ri} K \cap (\bigcup_{F \in F(K), F \neq K} F)\end{aligned}$$

Since $\mathcal{F}$ is a field that contains $F(K)$, it follows that $\operatorname{ri} K = K \cap (\bigcup_{F \in F(K), F \neq K} F)^c \in \mathcal{F}$ as desired. $\square$

Proposition 7: *Let $D_i \in \mathcal{D}$ converge to $D \in \mathcal{D}$ and let $K = N(D, x) \in \mathcal{K}$ for some $x \in D$. There exist $K_j \in \mathcal{K}$, k_j and $\epsilon_j > 0$ for $j = 1, 2, \dots$ such that (i) $K_{j+1} \subset K_j$ for all j, (ii) $\bigcap_j K_j = K$, and (iii) $\bigcup_{x_i \in D_i \cap B_{\epsilon_j}(x)} N(D_i, x_i) \subset K_j$ for $i > k_j$.*

Proof: Since $K \in \mathcal{K}$ Proposition 1(i) implies $\dim \operatorname{conv} D = n$. Let $y^* \in \operatorname{int} \operatorname{conv} D$ and let $\tilde{D}_j = \{x\} \cup (\frac{j}{j+1} D + \frac{1}{j+1}\{y^*\})$. Note that $y^* \in \operatorname{int} \operatorname{conv} \tilde{D}_j$. Define $K_j := N(\tilde{D}_j, x)$.

To prove (i) let $u \in K_{j+1}$ and hence $u \cdot x \geq u \cdot (\frac{j+1}{j+2} y + \frac{1}{j+2} y^*)$ for all $y \in D$. Note that $u \cdot x \geq u \cdot y^*$ since $y^* \in \operatorname{int} \operatorname{conv} \tilde{D}_{j+1}$. It follows that $u \cdot x \geq u \cdot (\frac{j}{j+1} y + \frac{1}{j+1} y^*)$ for all $y \in D$ and hence $u \in K_j$.

If $u \in K$ then $u \cdot x \geq u \cdot y$ for all $y \in D$ and hence $u \cdot x \geq u \cdot (\frac{j}{j+1} y + \frac{1}{j+1} y^*)$ for all $y \in D$ and therefore $u \in K_j$ for all j. Let $u \in \bigcap_j K_j$ then $u \cdot x \geq u \cdot (\frac{j}{j+1} y + \frac{1}{j+1} y^*)$ for all j and all $y \in D$. It follows that $u \cdot x \geq u \cdot y$ for all $y \in D$ and hence $u \in K$. This proves (ii).

To prove (iii), first, we observe that $u \cdot y > u \cdot x$ for all $u \in N(\tilde{D}_j, \frac{j}{j+1} y + \frac{1}{j+1} y^*), u \neq o$. To see this, note that for $u \in N(\tilde{D}_j, \frac{j}{j+1} y + \frac{1}{j+1} y^*), u \neq o$ there is z with $u \cdot z > 0$. Since $y^* + \epsilon' z \in \operatorname{int} \operatorname{conv} \tilde{D}_j$ for some $\epsilon' > 0$ and since $u \cdot (\frac{j}{j+1} y + \frac{1}{j+1} y^*) \geq u \cdot (y^* + \epsilon' z)$, we conclude that $u \cdot y > u \cdot y^*$. But $u \cdot (\frac{j}{j+1} y + \frac{1}{j+1} y^*) \geq u \cdot x$ and therefore, $u \cdot y > u \cdot x$.

Recall that S denotes the unit sphere. For $y \in D$, let $R_j(y) := N(\tilde{D}_j, \frac{j}{j+1} y + \frac{1}{j+1} y^*) \cap S$. Clearly, $R_j(y)$ is compact. By the argument above, $u \cdot y > u \cdot x$ for $u \in R_j(y)$. Since $R_j(y)$ is compact and D is finite, there is an $\alpha > 0$ such that $\max_{y \in D} u \cdot (y - x) \geq \alpha$ for all $u \in R_j := \bigcup_{y \in D, y \neq x} R_j(y)$. Note that if $u \notin K_j$ then $\lambda u \in R_j$ for some $\lambda > 0$.

Choose $\epsilon_j > 0$ so that $|u \cdot z| < \alpha/4$ for all $u \in R_j$ and $z \in B_{\epsilon_j}(o)$. Choose k_j so that $B_{\epsilon_j}(y) \cap D_i \neq \emptyset$ for all $y \in D$ and $i > k_j$. Then, for all $u \in R_j(y), x_i \in D_i \cap B_{\epsilon_j}(x), y_i \in D_i \cap B_{\epsilon_j}(y)$ we have $u \cdot (x_i - y_i) \leq u \cdot (x - y) + \max_{x_i \in B_{\epsilon_j}(x)} u \cdot (x_i - x) - \min_{y_i \in B_{\epsilon_j}(y)} u \cdot (y_i - y) < u \cdot (x - y) + \alpha/2 < 0$ and hence $u \notin N(D_i, x_i) \cap S$ for $x_i \in D_i \cap B_{\epsilon_j}(x)$. We conclude that $\bigcup_{x_i \in D_i \cap B_{\epsilon_j}(x)} N(D_i, x_i) \subset K_j$ for all $i > k_j$. $\square$

Proposition 8: *Let $K \in \mathcal{K}$ and $\epsilon > 0$. There exist $D, D' \in \mathcal{D}$, $K' \in \mathcal{K}$ and an open set O such that $D \cap B_1(o) = D' \cap B_1(o) = \{o\}$, $K = N(D, o), K' = N(D', o)$, $d_h(D, D') < \epsilon$ and $K \cap S \subset O \subset K'$.*

Proof: By Proposition 4, there is $D \in \mathcal{D}$ such that $o \in D$ and $K = N(D, o)$. Since $N(D, o) = N(\lambda D, o)$ for $\lambda > 0$, we may choose D so that $D \cap B_2(o) = \{o\}$. By Proposition 1(i), $\dim D = n$. Choose $y \in \operatorname{int}\operatorname{conv} D$ and $\lambda \in (0, 1)$ so that $D' := \{o\} \cup ((1-\lambda)D + \lambda\{y\})$ satisfies $d_h(D, D') < \epsilon$ and $D' \cap B_1(o) = \{o\}$. Clearly, $\dim \operatorname{conv} D' = n$ and hence $K' := N(D', o) \in \mathcal{K}$. If $K = \{o\}$ then $K' = K$ and $O = \emptyset$ have the desired property and we are done. Therefore, assume $K \neq \{o\}$. Obviously, $0 > u \cdot y$ for $u \in K, u \neq o$ and $y \in \operatorname{int}\operatorname{conv} D$. Hence, $0 > u \cdot x \; \forall x \in D' \backslash \{o\}, \forall u \in K, u \neq o$. Since $K \cap S$ is compact there is $\epsilon' > 0$ such that $-\epsilon' > u \cdot x \; \forall x \in D' \backslash \{o\}, \forall u \in K \cap S$. Let $\epsilon'' = \min_{x \in D' \backslash \{o\}} \epsilon'/(2\|x\|)$. Then $0 > -\epsilon'/2 \geq u \cdot x + (u' - u) \cdot x = u' \cdot x \; \forall x \in D' \backslash \{o\}, \forall u' \in \bigcup_{u \in K \cap S} B_{\epsilon''}(u)$. Let $O := \bigcup_{u \in K \cap S} B_{\epsilon''}(u)$. Clearly, $K \cap S \subset O \subset K'$ as desired. $\square$

8. Appendix B: Proofs

It is convenient to view random choice rules as maps from nonempty finite subsets of the $n-$dimensional Euclidean space $\mathbb{R}^n$ (rather than P) to probability measures on the Borel subsets of $\mathbb{R}^n$. To see how this can be done, let $\hat{P} = \{x \in \mathbb{R}^n_+ \mid \sum_{i=1}^n x^i \leq 1\}$. Hence, $\hat{P}$ is the $n-$dimensional "Machina-Marschak Triangle". There is an obvious way to interpret ρ as a RCR on finite subsets of $\hat{P}$ and a RUF as a probability measure on the algebra generated by polyhedral cones in $\mathbb{R}^n$. This is done with the aid of the following two bijections. Define, $T_0 : \mathbb{R}^n \to U$ and $T_1 : \hat{P} \to P$ as follows:

$$T_0(u^1, \ldots, u^n) = (u^1, \ldots, u^n, 0) \text{ and}$$
$$T_1(x^1, \ldots, x^n) = (x^1, \ldots, x^n, 1 - \textstyle\sum_{i=1}^n x^i)$$

Note that $\hat{P}$ is convex and both T_0, T_1 are homeomorphisms satisfying the following properties:

$$T_0(\gamma u + \beta v) = \alpha T_0(u) + \beta T_0(v)$$
$$T_1(\gamma x + (1-\gamma)y) = \gamma T_1(x) + (1-\gamma)T_1(y)$$
$$T_0(u) \cdot T_1(v) = u \cdot v$$

for all $u, v \in \mathbb{R}^n$, $x, y \in \hat{P}$, $\alpha, \beta \in \mathbb{R}$, and $\gamma \in (0, 1)$.

Let $\hat{\rho}^{\hat{D}}(x) = \rho^{T_1(\hat{D})}(T_1(x))$. We extend the RCR $\hat{\rho}$ to all finite non-empty subsets of $\mathbb{R}^n$ in the following manner: Choose $z \in \operatorname{int}\hat{P}$. For $D \subset \mathbb{R}^n$ let $\gamma_D = \max\{\gamma \in$

$(0,1] \mid \gamma D + (1-\gamma)\{z\} \subset \hat{P}\}$. Note that γ_D is well-defined since $\hat{P}$ is closed and $z \in \operatorname{int} \hat{P}$. Also, if $D \subset \hat{P}$, then $\gamma_D = 1$. Extend $\hat{\rho}$ to all finite, nonempty $D \subset I\!R^n$ by letting $\hat{\rho}^D(x) = \hat{\rho}^{\gamma D + (1-\gamma)\{z\}}(\gamma x + (1-\gamma) z)$ for all x, D.

For the extended RCR, the following definitions of linearity and mixture continuity will be used.

Definition: *The RCR is linear if $\rho^D(x) = \rho^{tD+\{y\}}(tx+y)$ for all $t > 0, y \in I\!R^n$ and $x \in D$.*

Definition: *The RCR is mixture continuous if $\rho^{tD+t'D'}$ is continuous in t, t' for all $t, t' \geq 0$.*

Continuity, monotonicity, extremeness and undominatedness of $\hat{\rho}$ are defined the same way as the corresponding properties for ρ. It follows from the properties of T_1 stated above that $\hat{\rho}$ is mixture continuous (continuous, monotone, linear, extreme, undominated) if and only if ρ is mixture continuous (continuous, monotone, linear, extreme, undominated). Furthermore, $\hat{\rho}$ maximizes $\mu \circ T_0$ if and only if μ maximizes ρ. Hence, in the proofs we work in $I\!R^n$ so that ρ refers to the corresponding $\hat{\rho}$ and μ to $\mu \circ T_0$.

Lemma 1: *If ρ is monotone, linear and extreme then $x \in D$, $x \in D'$ and $N(D,x) = N(D',x')$ implies $\rho^D(x) = \rho^{D'}(x')$.*

Proof: By linearity, $\rho^{D-\{x\}}(o) = \rho^D(x)$. Therefore, it suffices to show that $N(D,o) = N(D',o)$, $o \in D, D'$ implies $\rho^D(o) = \rho^{D'}(o)$.

We first show that if $N(D,o) = N(D',o)$ there exists $\lambda \in (0,1)$ such that $D'' := \lambda D' \subset \operatorname{conv} D$. By Proposition 1($ii$), $\operatorname{pos} D = N(L,o)$ for $L = N(D,o)$. Let $y \in D'$. Since $D' \subset N(L,o)$ it follows that $y = \sum \alpha_i x_i$, $x_i \in D, \alpha_i \geq 0$. Since $o \in D$, $\lambda y \in \operatorname{conv} D$ for λ sufficiently small proving the assertion.

By linearity $\rho^{D''}(o) = \rho^{D'}(o)$. Then, monotonicity and extremeness imply that $\rho^{D''}(o) \geq \rho^{D'' \cup D}(o) = \rho^D(o)$. Hence, $\rho^{D'}(o) \geq \rho^D(o)$. A symmetric argument ensures $\rho^D(o) \geq \rho^{D'}(o)$ and hence $\rho^D(o) = \rho^{D'}(o)$ as desired. □

A RUF is *full-dimensional* if algebra elements that have dimension smaller than n have measure zero.

Definition: *The RUF μ is full-dimensional if $\mu(F) = 0$ whenever $\dim F < n$.*

Lemma 2: *A RUF μ is full-dimensional if and only if it is regular.*

Proof: Assume μ is full-dimensional. We first establish that $I\!R^n = \bigcup_{x \in D} N^+(D, x) \cup F$ where F is a finite union of polyhedral cones of dimension less than n. It is easy to see that $I\!R^n = \bigcup_{x \in D} N(D, x)$. By Proposition 5($i$), each $N(D, x)$ can be expressed as the disjoint union of sets $\operatorname{ri} A$ for $A \in F(N(D, x))$. Recall that each face of a polyhedral cone is a polyhedral cone. Note that $A \in F(K)$ and $A \neq K$ implies $A = H(u, \alpha) \cap K$ for some $u \neq o$ and some $\alpha \in I\!R$. Hence, $A \neq K$ implies $\dim A < n$. If $\dim N(D, x) = n$ then $\operatorname{ri} N(D, x) = \operatorname{int} N(D, x) = N^+(D, x)$. Therefore, $N(D, x) = N^+(D, x) \cup F'$ where F' is a finite union of polyhedral cones with dimension less than n and hence $I\!R^n = \bigcup_{x \in D} N(D, x) = \bigcup_{x \in D} N^+(D, x) \cup F$ where F is a finite union of polyhedral cones of dimension less than n. If μ is full-dimensional then $\mu(F) = 0$. Therefore, $1 = \mu(I\!R^n) = \mu\left(\bigcup_{x \in D} N^+(D, x)\right)$ which proves the "only if" part of the lemma.

If μ is not full-dimensional then there exists a set $F \in \mathcal{F}$ such that $\dim F < n$ and $\mu(F) > 0$. By Proposition 6, $\mathcal{H} := \{\operatorname{ri} K | K \in \mathcal{K}\}$ is a semiring and every element of $\mathcal{F}$ can be written as a finite union of elements in $\mathcal{H}$. Therefore, $\mu(K) > 0$ for some $K \in \mathcal{K}^*$ with $\dim K < n$. By Proposition 1(i), $\dim K < n$ implies there is $x \neq 0$ such that $x, -x \in N(K, o)$. Let $D = \{x, -x\}$ and note that $K \subset N(D, x) \cap N(D, -x)$. Hence, $\mu(N^+(D, x) \cup N^+(D, -x)) \leq 1 - \mu(K) < 1$ and μ is not regular. $\square$

Lemma 3: *The set of regular RUFs is nonempty.*

Proof: Let V be the usual notion of volume in $I\!R^n$. For any $K \in \mathcal{K}^*$, let $\mu_V(\operatorname{int} K) = \frac{V(B_1(o) \cap K)}{V(B_1(o))}$. Obviously, $\dim K < n$ implies $\operatorname{int} K = \emptyset$ and hence $V(B_1(o) \cap K) = 0$. By Proposition 5(i), $K \backslash \operatorname{int} K$ can be written as a finite union of set of dimension less than n. Hence, $\mu_V(K) = \mu_V(\operatorname{int} K)$ and therefore μ_V is a RUF. Since μ_V assigns probability 0 to every set of dimension less than n, by Lemma 2, μ_V is a regular RUF. $\square$

8.1 Proof of Theorem 1

Let μ, μ' be regular RUFs with $\mu(N(D, x)) = \mu'(N(D, x)$ for all (D, x). We must show that $\mu(F) = \mu'(F)$ for all $F \in \mathcal{F}$.

Let $F \in \mathcal{F}$. By Proposition 6, we can write F as a finite union of elements in $\mathcal{H}$. In fact, it is easy to see that we can write F as a finite union of disjoint elements of $\mathcal{H}$. Recall that $\mathcal{H} := \{\operatorname{ri} K | K \in \mathcal{K}\}$. To prove Theorem 1(ii) it therefore suffices to show that $\mu(\operatorname{ri} K) = \mu'(\operatorname{ri} K)$ for all $K \in \mathcal{K}$.

By Proposition 4, for every $K \in \mathcal{K}$ there is D, x such that $K = N(D, x)$. Note that $N^+(D, x) \subset \operatorname{ri} N(D, x) \subset N(D, x)$. If μ is regular then $\mu(N^+(D, x)) = \mu(N(D, x))$. Therefore, $\mu(\operatorname{ri} K) = \mu(K)$ for all $K \in \mathcal{K}$. It follows that $\mu(\operatorname{ri} K) = \mu'(\operatorname{ri} K)$ for all $K \in \mathcal{K}$. $\square$

8.2 Proof of Theorem 2

Let μ be regular RUF. By Theorem 1 there exists a unique ρ that maximizes μ. Hence, ρ and μ satisfy

$$\rho^D(x) = \mu(N(D, x)) \tag{1}$$

for all $D \in \mathcal{D}$ and $x \in D$. We argue in the text that ρ must be monotone, linear and extreme. To prove mixture continuity, we must show that $\rho^{tD+t'D'}$ is continuous in t, t'. By equation (1) and Proposition 3, $\rho^{tD+t'D'}(tx + t'x') = \mu(N(D, x) \cap N(D', x'))$ which implies that $\rho^{tD+t'D'}$ is continuous in (t, t') for $t, t' > 0$. Hence, it remains to show that $\rho^{tD+D'} \to \rho^{D'}$ as $t \to 0$. Choose $\epsilon > 0$ small enough so that $B_\epsilon(x') \cap D' = \{x'\}$ and choose t small enough so that $x' + tx \in B_\epsilon(x')$ for all $x \in D$. Proposition 3 and the fact that $\bigcup_{x \in D} N(D, x) = I\!R^n$ imply that

$$\rho^{tD+D'}(B_\epsilon(x')) = \mu(\bigcup_{x \in D} (N(D, x) \cap N(D', x'))) = \mu(N(D', x')) = \rho^{D'}(x')$$

which completes the proof of mixture continuity. This proves the only if part of the Theorem.

Let ρ be a mixture continuous, monotone linear and extreme RCR. By Proposition 4, for any $K \in \mathcal{K}^*$ there exists (D, x) such that $K = N(D, x)$. We define $\mu : \mathcal{H} \to I\!R$ as follows:

$$\mu(\operatorname{ri} K) = \rho^D(x) \tag{2}$$

for D, x such that $K = N(D, x), K \in \mathcal{K}$. Lemma 1 ensures that μ is well-defined. If $\dim K < n$ then Proposition 1(iii) implies that x is not an extreme point of D. Since ρ is extreme this in turn implies $\rho^D(x) = 0$, so $\mu(\operatorname{ri} K) = 0$ for any $K \in \mathcal{K}^*$ such that $\dim(K) < n$. Note that $F \in F^0(K)$ and $\dim F = \dim K$ implies $F = \operatorname{ri} K$. It follows from Proposition 5(i) that

$$\mu(\operatorname{int} K) = \mu(K) \tag{3}$$

for $K \in \mathcal{K}$.

In Lemmas 4 and 5 it is understood that the function μ (defined above) and the RCR ρ satisfy (2).

Lemma 4: *If ρ is mixture continuous, monotone, linear and extreme then μ is finitely additive.*

Proof: Assume $\operatorname{ri} K_0 = \bigcup_{i=1}^m \operatorname{ri} K_i$ and $K_i \in \mathcal{K}$ for all $i = 1, \ldots, m$ with $\operatorname{ri} K_i, i = 1, \ldots, n$ pairwise disjoint. By Proposition 4, there exist $D_i \in \mathcal{D}$ and $x_i \in D_i$ such that $N(D_i, x_i) = K_i$ for all $i = 0, \ldots, m$. Let $D = D_0 + \cdots + D_m$ and without loss of generality, assume that the D_i's are "generic" that is, for each $y \in D$, there exists a unique collection of y_j's such that $y = \sum_j y_j$ and for each $y' \in D_0 + \cdots + D_{i-1} + D_{i+1} + \cdots + D_m$ there exist a unique collection of y_j's for $j \neq i$ such that $y = \sum_{j \neq i} y_j$. Let $\beta^i > 0$ for all i and let $D(\beta) = \beta^0 D_0 + \cdots + \beta^m D_m$. Note that $N(\beta^i D_i, \beta^i y_i) = N(D_i, y_i)$ for $\beta^i > 0$ and hence Proposition 3 implies

$$N(D(\beta), \sum_i \beta^i y_i) = \bigcap_{i=1}^m N(D_i, y_i) \tag{4}$$

whenever $\beta^i > 0$ and $y_i \in D_i$ for all i.

Fix $i \in \{0, \ldots, m\}$ and let $\beta_k = (\beta_k^0, \ldots, \beta_k^m)$ be such that $\beta_k^j = \frac{1}{k}$ for $j \neq i$ and $\beta_k^i = 1$. For $y \in \bigcup_{j=0}^m D_j$, let

$$Z(y) = \{z = (z^0, \ldots z^m) \in \times_{j=0}^m D_j \mid z^j \in D_j \text{ for all } j, \, z^j = y \text{ for some } j\}$$
$$G_\beta(y) = \{y' \in D(\beta) \mid y' = \sum_{j=0}^m \beta^j z^j \text{ for } z \in Z(y)\}$$

Let $G(y) = G_{(1,\ldots,1)}(y)$. By our genericity assumption, for each $y \in \bigcup_{j=0}^m D_j$ there exists a unique j such that $y \in D_j$. Hence, the function $\phi : G(y) \to G_{\beta_k}(y)$ such that

$\phi(y_0 + \cdots + y_m) = \beta_k^0 y_0 + \cdots + \beta_k^m y_m$ is well-defined. Again, by our genericity assumption ϕ is a bijection for k sufficiently large. But since $N(D(\beta), \sum_i \beta^i y_i) = N(D, \sum_i y_i)$, we have $\rho^{D(\beta_k)}(G_{\beta_k}(y)) = \rho^D(G(y))$ for all $y \in \bigcup_{j=0}^m D_j$ and for sufficiently large k. Choose open sets O, O' such that $\{y\} = O \cap D_i$, $D_i \backslash \{y\} = O' \cap D_i$. By mixture continuity, $\lim_{k\to\infty} \rho^{D(\beta_k)}(G_{\beta_k}(y)) = \lim_{k\to\infty} \rho^{D(\beta_k)}(O) \geq \rho^{D_i}(O) = \rho^{D_i}(y)$ and similarly, $\lim_{k\to\infty} \rho^{D(\beta_k)}(D(\beta_k)\backslash G_{\beta_k}(y)) = \lim_{k\to\infty} \rho^{D(\beta_k)}(O') \geq \rho^{D_i}(O') = \rho^{D_i}(D_i \backslash \{y\})$. That is, $\rho^{D(\beta_k)}(G_{\beta_k}(y)) \to \rho^{D_i}(y)$ and hence we conclude for all $i = 0, \ldots, m$ and $y \in D_i$

$$\rho^D(G(y)) = \rho^{D_i}(y) \tag{5}$$

By the definition of μ, (4) implies that for $z^j \in D$, $j = 0, \ldots, m$ and $y = \sum_{j=0}^m z^j$,

$$\rho^D(y) = \mu[\operatorname{int} N(D, y)] = \mu[\bigcap_{j=0}^m \operatorname{int} N(D_j, z^j)] \tag{6}$$

Since $\operatorname{int} N(D, x_i) \cap \operatorname{int} N(D, x_j) = \emptyset$ and $\operatorname{int} N(D, x_i) \subset \operatorname{int} N(D, x_0)$ for $i, j \geq 1$, $i \neq j$, (6) implies

$$\rho^D(G(x_i) \cap G(x_j)) = 0 \text{ and } \rho^D(G(x_i)\backslash G(x_0)) = 0$$

for $i, j \geq 1$, $i \neq j$. Thus,

$$\begin{aligned}\rho^D(G(x_0)) &= \rho^D(\bigcup_{i=1}^m (G(x_0) \cap G(x_i)) \\ &= \rho^D(\bigcup_{i=1}^m G(x_i)) = \sum_{i=1}^m \sum_{y \in G(x_i)} \rho^D(y) = \sum_{i=1}^m \rho^D(G(x_i))\end{aligned} \tag{7}$$

Again, by the definition of μ, (6) and (7) imply that

$$\mu[\operatorname{int} N(D_0, x_0)] = \rho^{D_0}(x_0) = \sum_{i=1}^m \rho^{D_i}(x_i) = \sum_{i=1}^m \mu[\operatorname{int} N(D_i, x_i)]$$

as desired. $\square$

Next, we extend μ to $\mathcal{F}$. Equation (2) defines μ for every element of $\mathcal{H}$. By Proposition 6, $\mathcal{F}$ consists of all finite unions of elements in $\mathcal{H}$. In fact, it is easy to see that $\mathcal{F}$ consists of all finite unions of disjoint sets in $\mathcal{H}$. To extend μ to $\mathcal{F}$, set $\mu(\emptyset) = 0$ and define

$\mu(F) = \sum_{i=1}^m \mu(H_i)$ where $H_1, \ldots, H_m$ is some disjoint collection of sets in $\mathcal{H}$ such that $\bigcup_{i=1}^m H_i = F$. To prove that μ is well-defined and additive on $\mathcal{F}$, note that if $H'_j, j = 1, \ldots, k$ is some other disjoint collection such that $\bigcup_{j=1}^k H'_j = F$, then $\sum_{i=1}^m \mu_i(H_i) = \sum_{i=1}^m \sum_{j=1}^k \mu(H_i \cap H'_j) = \sum_{j=1}^k \mu_i(H'_j)$.

Next, we show that $\mu(\mathbb{R}^n) = 1$. It is easy to see that $\bigcup_{x \in E^*} N(E^*, x) = \mathbb{R}^n$. Note also that $N(E^*, x) \in \mathcal{K}$ for all x and $\mu(N(E^*, x)) = \mu(\operatorname{int} N(E^*(x))$ by Equation (3). Since interiors of normal cones at distinct points are disjoint, we have $\bigcup_{x \in E^*} \operatorname{int} N(E^*, x)) \subset \mathbb{R}^n$. Therefore, we have

$$\mu(\mathbb{R}^n) \leq \sum_{x \in E^*} N(E^*, x) = \sum_{x \in E^*} \mu(\operatorname{int} N(E^*, x)) \leq \mu(\mathbb{R}^n)$$

Since $\sum_{x \in E^*} \mu(\operatorname{int} N(E^*, x)) = \rho^{E^*}(E^*) = 1$ it follows that $\mu(\mathbb{R}^n) = 1$.

We have established that μ is a finitely additive probability measure and therefore a RUF.

Lemma 5: *The RCR ρ maximizes the RUF μ.*

Proof: Since ρ^D is a discrete measure, it suffices to show that $\rho^D(x) = \mu(N(D, x))$ for all $x \in D$. By the construction, this holds for all D, x such that D has dimension n (i.e., whenever $N(D, x) \in \mathcal{K}$). It remains to show that $\rho^D(x) = \mu(N(D, x))$ for lower dimensional decision problems.

Let $\alpha > 0$. Since $\dim(D + \alpha E^*) = n$, $\rho^{D+\alpha E^*}(x + \alpha y) = \mu(\operatorname{int} N(D + \alpha E^*, x + \alpha y)$. Then, Proposition 3 and the fact that the interiors of normal cones at distinct points are disjoint implies

$$\begin{aligned}
\rho^{D+\alpha E^*}(\{x\} + \alpha E^*) &= \sum_{y \in E^*} \rho^{D+\alpha E^*}(x + \alpha y) = \sum_{y \in E^*} \mu(\operatorname{int} N(D + \alpha E^*, x + \alpha y) \\
&= \mu(\bigcup_{y \in E^*} \operatorname{int} N(D + \alpha E^*, x + \alpha y) \\
&= \mu(\bigcup_{y \in E^*} N(D + \alpha E^*, x + \alpha y) = \mu(N(D, x))
\end{aligned}$$

The last equality follows from the fact that $\bigcup_{y \in E^*} N(E^*, y) = \mathbb{R}^n$. Choose open sets O, O' such that $\{x\} = O \cap D$, $D \backslash \{x\} = O' \cap D$. By mixture continuity,

$$\lim_{\alpha \to 0} \rho^{D+\alpha E^*}(\{x\} + \alpha E^*) = \lim_{\alpha \to 0} \rho^{D+\alpha E^*}(O) \geq \rho^D(O) = \rho^D(x)$$

and similarly,

$$\lim_{\alpha \to 0} \rho^{D+\alpha E^*}([D+\alpha E^*]\backslash[\{x\}+\alpha E^*]) = \lim_{\alpha \to 0} \rho^{D+\alpha E^*}(O') \geq \rho^D(O') = \rho^D(D\backslash\{x\})$$

That is,

$$\lim_{\alpha \to 0} \rho^{D+\alpha E^*}(\{x\}+\alpha E^*) = \rho^D(x)$$

Hence

$$\rho^D(x) = \mu(N(D,x))$$

for all $D \in \mathcal{D}$, $x \in I\!R^n$ and therefore ρ maximizes μ. $\square$

8.3 Proof of Theorem 3

Theorem 2 and Lemmas 6 proves the "only if" part of Theorem 3, while Theorem 2 and Lemma 7 proves the "if" part of Theorem 3.

Lemma 6: *Let ρ maximize the regular RUF μ. If ρ is continuous then μ is countably additive.*

Proof: By Theorem 11.3 of Billingsley (1986) any finitely additive and countably subadditive real-valued function on a semiring extends to a countably additive measure on $\sigma(\mathcal{H})$, the $\sigma-$field generated by $\mathcal{H}$. Since $I\!R^n \in \mathcal{H}$ and $\mu(I\!R^n) = 1$, the extension must be a (countably additive) probability measure. Hence, to prove that μ is countably additive it suffices to show that μ is countably subadditive on $\mathcal{H}$.

Let $\bigcup_{i=1}^m H_i = H_0$. Since $\mathcal{H}$ is a semiring we can construct a partition of H_0 that measures each H_i. Then, the finite additivity of μ implies the finite subadditivity of μ. To prove countable subadditivity, consider a countable collection of set $K_i, i = 0, \ldots$ such that $K_i \in \mathcal{K}$ and $\operatorname{ri} K_0 = \bigcup_{i=1}^\infty \operatorname{ri} K_i$. We must show that $\mu(\bigcup_{i=1}^\infty \operatorname{int} K_i) \leq \sum_{i=1}^\infty \mu(\operatorname{int} K_i)$.

By Proposition 5(i), each $K \in \mathcal{K}$ can be expressed as the disjoint union of sets $\operatorname{ri} A$ for $A \in F(K)$. Recall that each face of a polyhedral cone is a polyhedral cone. Note that $A \in F(K)$ and $A \neq K$ implies $A = H(u,\alpha) \cap K$ for some $u \neq o$ and some $\alpha \in I\!R$. Hence, $A \neq K$ implies $\dim A < n$. If $\dim K = n$ then $\operatorname{ri} K = \operatorname{int} K$. Therefore, $K = \operatorname{int} K \cup F$ where F is a finite union of polyhedral cones with dimension less than n. Since μ is full-dimensional (by Lemma 2) this implies that $\mu(\operatorname{int} K) = \mu(K)$. Since $\operatorname{ri} K_0 = \bigcup_{i=1}^{\infty} \operatorname{ri} K_i$, we have $K_0 = \bigcup_{i=1}^{\infty} K_i$ and it suffices to show that $\mu(\bigcup_{i=1}^{\infty} K_i) \leq \sum_{i=1}^{\infty} \mu(K_i)$

Proposition 8 implies that for each K_i, ϵ' there are $D_i, \tilde{D}_i \in \mathcal{D}$, $\tilde{K}_i \in \mathcal{K}$ and an open set O such that (i) $K_i \cap S \subset O_i \subset \tilde{K}_i$ with $K_i = N(D_i, o), \tilde{K}_i = N(\tilde{D}_i, o)$; (ii) $d_h(D_i, \tilde{D}_i) < \epsilon'$ and (iii) $\rho^{D_i}(B_1(o)) = \mu(K_i), \rho^{\tilde{D}_i}(B_1(o)) = \mu(\tilde{K}_i)$.

Note that (i) implies that $K_0 \cap S \subset \bigcup_{i=1}^{\infty} O_i$. Since $K_0 \cap S$ is compact, there exists a finite collection $O_i, i \in I$, $0 \notin I$, that covers $K_0 \cap S$. Hence $\tilde{K}_i, i \in I$ covers K_0. Since ρ is continuous (ii) and (iii) imply that we may choose ϵ' small enough so that $\mu(K_i) \geq \mu(\tilde{K}_i) - 2^i\epsilon$. Then, finite subadditivity implies $\mu(K_0) \leq \sum_{i \in I} \mu(\tilde{K}_i) - \epsilon \leq \sum_{i=1}^{\infty} \mu(K_i) - \epsilon$. Since ϵ was arbitrary the result follows. $\square$

Lemma 7: *If ρ maximizes the regular, countably additive, RUF μ then ρ is continuous.*

Proof: Assume that D_i converges to D. It suffices to show that $\limsup \rho^{D_i}(G) \leq \rho^D(G)$ for any closed $G \subset I\!R^n$ (Billingsley (1999), Theorem 2.1). Without loss of generality, assume $D \cap G = \{x\}$ for some $x \in D$.

Case 1: $\dim \operatorname{conv} D = n$**.** Then, Proposition 1(i) implies $N(D, x) \in \mathcal{K}$. By Proposition 7 there are $\epsilon_j > 0$, k_j, and $K_j, j = 1, 2, \ldots$ such that $K_{j+1} \subset K_j, \bigcap_j K_j = N(D, x)$ and

$$\bigcup_{y \in D_i \cap B_{\epsilon_j}(x)} N(D_i, y) \subset K_j \tag{8}$$

for all $i > k_j$.

Since D_i converges to D and $D \cap G = \{x\}$, for all $\epsilon_j > 0$, there exists m_j such that $i \geq m_j$ implies

$$D_i \cap G \subset B_{\epsilon_j}(x) \tag{9}$$

Let $F_j = K_j \backslash N(D, x)$. Since μ is countably additive and $F_j \downarrow \emptyset$ we conclude that $\mu(F_j) \to 0$. Hence, for all $\epsilon > 0$ there exist m such that $j \geq m$ implies

$$\mu(K_j) \leq \mu(N(D, x)) + \epsilon \tag{10}$$

For a given ϵ choose j so that (10) is satisfied. Then, choose $i > \max\{m_j, k_j\}$ so that both (8) and (9) are satisfied. By Proposition 1(iv), the interiors of normal cones at distinct points of D_i are disjoint. Since μ is full-dimensional, we have $\mu(N(D_i,x)) = \mu(\operatorname{int} N(D_i,x))$. Therefore,

$$\rho^{D_i}(G) = \sum_{y\in D_i\cap G} \mu(N(D_i,y)) = \bigcup_{y\in D_i\cap G} \mu(N(D_i,y)) \le \mu(K_j) \le \rho^D(G)+\epsilon$$

Since, ϵ is arbitrary, $\rho^D(G) \ge \limsup \rho^{D_i}(G)$ as desired.

Case 2: $\dim \operatorname{conv} D < n$**.** Note that $x \in M(D_i,u)$ implies $M(\lambda D_i + (1-\lambda)E^*, u) \subset \lambda x + (1-\lambda)E^*$. Hence, we conclude

$$\rho^{D_i}(x) \le \rho^{\lambda D_i+(1-\lambda)E^*}(\lambda\{x\} + (1-\lambda)E^*)$$

Since $\dim \operatorname{conv}[\lambda D_i + (1-\lambda)E^*] = n$, the argument above establishes

$$\limsup \rho^{\lambda D_i+(1-\lambda)E^*}(\lambda\{x\} + (1-\lambda)E^*) \le \rho^{\lambda D+(1-\lambda)E^*}(\lambda\{x\} + (1-\lambda)E^*)$$

Choose $\lambda \in (0,1)$ such that $\|x-y\| < \frac{1-\lambda}{\lambda}\|x'-y'\|$ for all $x,y \in D$ and $x',y' \in E^*$, $x' \ne y'$. Note that $M(\lambda D + (1-\lambda)E^*, u) = \lambda M(D,u) + (1-\lambda)M(E^*,u)$. Hence, for all $w \in M(\lambda D+(1-\lambda)E^*,u)\cap[\lambda\{x\}+(1-\lambda)E^*]$ there exists $x_D \in M(D,u)$ and $x_{E^*}, y_{E^*} \in E^*$ such that $w = \lambda x_D + (1-\lambda)x_{E^*} = \lambda x + (1-\lambda)y_{E^*}$. Hence $\lambda(x-x_D) = (1-\lambda)(x_{E^*} - y_{E^*})$. From our choice of λ, we conclude that $x = x_D$. Therefore

$$\rho^{\lambda D+(1-\lambda)E^*}(\lambda\{x\} + (1-\lambda)E^*) \le \rho^D(x)$$

The last three display equations yield $\limsup \rho^{D_i}(x) \le \rho^D(x)$ as desired. □

8.4 Proof of Theorem 4

Theorem 4 follows from Theorem 3 and Lemma 9 below.

Lemma 9: *A continuous RCR is undominated if and only if it is extreme.*

Proof: Note that $\operatorname{ext} D \subset \operatorname{bd}\operatorname{conv} D$. Hence, every extreme RCR is undominated. For the converse, consider a D such that $\dim D = n$. Let $D_k = \operatorname{ext} D \cup (\frac{k-1}{k}D + \frac{1}{k}\{y\})$ for $y \in \operatorname{int}\operatorname{conv} D$. Note that D_k converges to D and $D_k \cap \operatorname{bd}\operatorname{conv} D_k = \operatorname{ext} D$. Therefore, ρ is undominated implies $\rho^{D_k}(\operatorname{ext} D) = 1$ for all k. By continuity, $\rho^D(\operatorname{ext} D) = 1$ as desired. Let m be any number such that $1 < m \le n$. To conclude the proof, we show that if

$\rho^D(\text{ext}\, D) = 1$ for all $D \in \mathcal{D}$ such that $\dim D = m$ then $\rho^D(\text{ext}\, D) = 1$ for all $D \in \mathcal{D}$ such that $\dim D = m-1$. Let $\dim D = m-1$ and $x \in D \backslash \text{ext}\, D$. Choose $y \in \text{ext}\, D$ and $z \notin \text{aff}\, D$. Define $D_k = D \cup \{\frac{k-1}{k} y + \frac{1}{k} z\}$ and note that $\dim D_k = m$, D_k converges to D and $\text{ext}\, D_k = (\text{ext}\, D) \cup \{\frac{k-1}{k} y + \frac{1}{k} z\}$ for all k. Hence, there exists an open set O such that $x \in O$ and $O \cap \text{ext}\, D_k = \emptyset$ for all k. By assumption, $\rho^{D_k}(O) = 0$ for all k. Then, by continuity $\rho^D(x) \leq \rho^D(O) = 0$. □

8.5 Proof of Theorem 5

For $D \in \mathcal{D}$ and $x \in D$ let $\mathcal{P}_x(D)$ denote the collection subset of D that contain x. For $X \subset I\!R^n$, let $\neg X = I\!R^n \backslash X$. Note that

$$N_l(D, x) = \bigcup_{B \in \mathcal{P}_x(D)} \Big(\big(\bigcap_{y \in B} N(D, y) \cap \bigcap_{y \in \neg B \cap D} \neg N(D, y) \big) \times N(B, x) \Big) \tag{11}$$

where we let the intersection over an empty index set (i.e., for $B = D$) equal $I\!R^n$. Define

$$N_l^+(D, x) := \bigcup_{B \in \mathcal{P}_x(D)} \Big(\big(\bigcap_{y \in B} N(D, y) \cap \bigcap_{y \in \neg B \cap D} \neg N(D, y) \big) \times N^+(B, x) \Big) \tag{12}$$

Lemma 8: $N_l(D, x) \in \mathcal{F}^2$.

Proof: $\mathcal{F}$ is a field that contains $N(D', y)$ for all (D', y). Since $\mathcal{F}^2$ contains $\mathcal{F} \times \mathcal{F}$ equation (11) implies that $\mathcal{F}^2$ contains $N_l(D, x)$. □

Let η be a tie-breaker and let $\rho : \mathcal{B} \to \Pi$ be defined as

$$\rho^D(B) = \sum_{x \in D \cap B} \eta(N_l(D, x)) \tag{13}$$

for all $D \in \mathcal{D}$, $B \in \mathcal{B}$. Clearly, this ρ is the only candidate for a maximizer of the tie-breaker η. Lemma 9 below shows that the ρ defined in (13) is a well defined RCR. This proves that every tie-breaker has a unique maximizing RCR.

Lemma 9: *Then function ρ defined in (13) is a RCR.*

Proof: To prove that ρ is a RCR we suffices to show that $\sum_{x \in D} \rho^D(x) = 1$ for all D, x. First, we show that $\eta(N_l(D, x)) = \eta(N_l^+(D, x))$ for all D, x. Clearly, $\eta(N_l(D, x)) \geq$

$\eta(N_l^+(D,x))$. If $\eta(N_l(D,x)) > \eta(N_l^+(D,x))$ then by (11) and (12) there is $F \in \mathcal{F}$ such that $\eta(F \times N(B,x)) > \eta(F \times N^+(B,x))$. Since $\eta(U \backslash F \times N(B,x)) \geq \eta(U \backslash F \times N^+(B,x))$, this implies that $\eta(U \times N(B,x)) > \eta(U \times N^+(B,x))$ contradicting the regularity of η_2.

For $x \neq y$, $N_l^+(D,x) \cap N_l^+(D,y) = \emptyset$. Also, $\bigcup_{x \in D} N_l(D,x) = I\!R^n \times I\!R^n$. Therefore,

$$\rho^D(x) = \sum_{x \in D} \eta(N_l(D,x)) \geq \eta(\bigcup_{x \in D} N_l(D,x)) = \eta(I\!R^n \times I\!R^n) = 1$$
$$\rho^D(x) = \sum_{x \in D} \eta(N_l^+(D,x)) = \eta(\bigcup_{x \in D} N_l^+(D,x)) \leq \eta(I\!R^n \times I\!R^n) = 1$$

Hence ρ^D is a RCR. $\square$

Lemma 10: *Let the RCR ρ maximize the tie-breaker η. Then ρ is monotone, linear, mixture continuous and extreme.*

Proof: Note that for all D, $x \in D$, y, and $\lambda \in (0,1)$, $N_l(D \cup \{y\}, x) \subset N_l(D,x)$ and $N_l(\lambda D + (1-\lambda)\{y\}, x + (1-\lambda)y) = N_l(D,x)$. Hence, monotonicity and linearity of ρ follows immediately from its definition.

Next, we prove that ρ is extreme. For any $B \subset D$, let $F_B(D)$ denote the intersection of all faces of $F(\operatorname{conv} D)$ that contain B. Obviously, $B \subset F_B(D) \cap D$. Suppose there exists $z \in F_B(D) \cap D, z \notin B$. Then, $u \in \bigcap_{y \in B} N(D,y)$ implies $u \in N(D,z)$ and therefore $\bigcap_{y \in B} N(D,y) \cap \bigcap_{y \in \neg B \cap D} \neg N(D,y) = \emptyset$. Hence, in (11) it suffices to consider B such that $B = F \cap D$ for some face $F \in F(\operatorname{conv} D)$. But if $B = F \cap D$ for some $F \in F(\operatorname{conv} D)$ and $x \in B$ is not extreme point of D then it is not an extreme point of B. But then, the regularity of η_2 ensure $\eta(I\!R^n, N^+(B,x)) = 0$, proving the extremeness of ρ.

To prove mixture continuity, it suffices to show that $\rho^{tD+t'D'}$ is continuous in t, t'. By Proposition 3, $\rho^{tD+t'D'}(tx + t'x') = \mu(N_l(D,x) \cap N_l(D',x'))$ which implies that $\rho^{tD+t'D'}$ is continuous in (t,t') for $t, t' > 0$. Hence, it remains to show that $\rho^{tD+D'} \to \rho^{D'}$ as $t \to 0$. Choose $\epsilon > 0$ small enough so that $B_\epsilon(x') \cap D' = \{x'\}$ and choose t small enough so that $x' + tx \in B_\epsilon(x')$ for all $x \in D$. Proposition 3 and the fact that $\bigcup_{x \in D} N_l(D,x) = I\!R^n$ imply that

$$\rho^{tD+D'}(B_\epsilon(x')) = \mu(\bigcup_{x \in D} (N_l(D,x) \cap N_l(D',x'))) = \mu(N_l(D',x')) = \rho^{D'}(x')$$

which establishes mixture continuity and completes the proof of the Lemma. $\square$

8.6 Proof of Theorem 6

In Lemma 3 we construct a regular RUF μ_V. Obviously, $\mu \times \mu_V$ is a tie-breaker for μ. Then Theorem 5 proves part (i) of the theorem.

Let ρ be such that $\rho^D(x) = \eta(N_l(D, x))$ for all D, x and $\eta = \mu \times \mu_V$ where μ_V is the regular RUF constructed in Lemma 3. By Theorem 5, this identifies a unique RCR ρ that is a maximizer of μ. To construct a second maximizer, note that since μ is not full-dimensional there exists some polyhedral cone K_* such that $\dim K_* < n$ and $\mu(K_*) > 0$. By the argument given in the proof of Lemma 2, there is $x_* \neq 0$ such that $K_* \subset N(D_*, x_*) \cap N(D_*, -x_*)$ for $D_* = \{-x_*, x_*\}$. Define μ_* as follows:

$$\mu_*(K) = \frac{V(B_1(o) \cap K \cap N(D_*, x_*))}{V(B_1(o) \cap N(D_*, x_*))}$$

Repeating the arguments made for μ_V establishes that μ_* is a regular RUF.[15] Then, let ρ_* be defined by $\rho_*^D(x) = \eta_*(N(D, x))$ where $\eta_* = \mu \times \mu_*$. Again by Theorem 5, ρ_* is a maximizer of μ. Note that $1 = \rho_*^{D_*}(x_*) \neq \rho^{D_*}(x_*) = .5$. Hence, $\rho_* \neq \rho$ and we have shown that there are multiple maximizers of μ. $\square$

8.7 Proof of Theorem 8

Lemma 11: *Let ρ maximize some RUF μ such that $\mu(o) = 0$. If ρ is continuous then μ is regular.*

Proof: If ρ maximizes some μ then

$$\mu(N^+(D, x)) \leq \rho^D(x) \leq \mu(N(D, x)) \tag{14}$$

Suppose μ is not regular. By Lemma 2, this implies that μ is not full-dimensional. By Proposition 6, $\mathcal{H} := \{\mathrm{ri}\, K | K \in \mathcal{K}\}$ is a semiring and every element of $\mathcal{F}$ can be written as a finite union of elements in $\mathcal{H}$. Therefore, $\mu(K) > 0$ for some $K \in \mathcal{K}$ with $\dim K < n$. By Proposition 1(i), $\dim K < n$ implies there is $x \neq 0$ such that $x, -x \in N(K, o)$. Since K is a pointed cone, o is an extreme point of K and therefore Proposition 1(iii) implies

[15] A similar construction is used in Regenwetter and Marley (2001), p. 880.

that $N^+(K,o)$ is non-empty. Hence there is z such that $uz < 0$ for all $u \in K$, $u \neq o$. Let $D_k := \{x, {}^1\!/_k(-z), -x\}$ and let $D = \{-x, o, x\}$. Let O be an open ball that contains o but does not contain $x, -x$. Since $\mu(K) = \mu(K \backslash \{o\})$, for all k sufficiently large, (14) implies $\rho^{D_k}(O) \geq \mu(K \backslash \{o\}) = \mu(K) > 0$. But $\rho^D(O) = 0$ since ρ is extreme. $\square$

To prove Theorem 8, let μ' be a dilation of μ for some non-regular μ such that $\mu(o) = 0$. Let ρ maximize μ' (and hence maximize μ). Since ρ maximizes μ, Lemma 11 implies that ρ is not continuous. Since μ' is regular, Theorem 3 implies μ' is not countably additive. $\square$

References

1. Barbera, S. and P.K. Pattanaik, "Falmagne and the Rationalizability of Stochastic Choices in Terms of Random Orderings", *Econometrica*, 1986, Vol. 54, pp. 707–715.

2. Billingsley, P., *Probability and Measure*, 1986, John Wiley & Sons, New York.

3. Chernoff, H., "Rational Selection of Decision Functions," *Econometrica* 1954, 22, 422–443.

4. Clark, S. A., "The Random Utility Model with an Infinite Choice Space," *Economic Theory*, 1995, 7, 179–189.

5. Cohen, M. A., "Random Utility Systems - the Infinite Case", *Journal of Mathematical Psychology*, 1980, 22, 1–23.

6. Dekel, E., B. Lipman and A. Rustichini, "A Unique Subjective State Space for Unforeseen Contingencies," *Econometrica*, 2001, 69, 891-934.

7. Falmagne, J.-C., "A Random Utility Model for a Belief Function", Synthese, 1983, 57, 35-48.

8. Falmagne, J.-Cl., "A Representation Theorem for Finite Random Scale Systems", *Journal of Mathematical Psychology*, 1978, 18, pp. 52–72.

9. Fishburn, P. C., "Induced Binary Probabilities and the Linear Ordering Polytope: A Status Report," *Mathematical Social Sciences*, 1992, 23, 67–80.

10. Gul F. and W. Pesendorfer, "Supplement to Random Expected Utility," mimeo, 2004.

11. Kalai, G, A. Rubinstein and R. Spiegler, "Rationalizing Choice Functions by Multiple Rationales", *Econometrica*, 2002, 70, 2481-2488.

12. Kahneman, D and A. Tversky, "Prospect Theory: An Analysis of Decision under Risk", *Econometrica*, 1979, 47, 263–292.

13. Kreps, D., "A Preference for Flexibility," *Econometrica*, 1979, 47, 565–576.

14. Luce, R. D., and Suppes, P, "Preference, Utility, and Subjective Probability". In: Luce, R.d., Bush, R. R., and Galanter, E., editors, Handbook of Mathematical Psychology, vol. III, 1965, 249-410. Wiley, New York.

15. McFadden, D. L. "Revealed Stochastic Preference: A Synthesis. Mimeo, Department of Economics, Berkeley, 2003.

16. McFadden, D. and M. Richter, "Revealed Stochastic Preference", in: J. S. Chipman, D. McFadden and M. K. Richter, editors, *Preferences, Uncertainty and Optimality.* Westview Press, Boulder, CO, 1991.

17. Regenwetter, M. and A. A. J. Marley, “Random Relations, Random Utilities and Random Functions”, *Journal of Mathematical Psychology*, 2001, 45, 864-912.

18. Rockafellar, T., *Convex Analysis*, 1970, Princeton University Press, Princeton, New Jersey.

19. Schneider, R., *Convex Bodies: The Brunn-Minkowski Theory*, 1993, Cambridge University Press, Victoria, Australia.

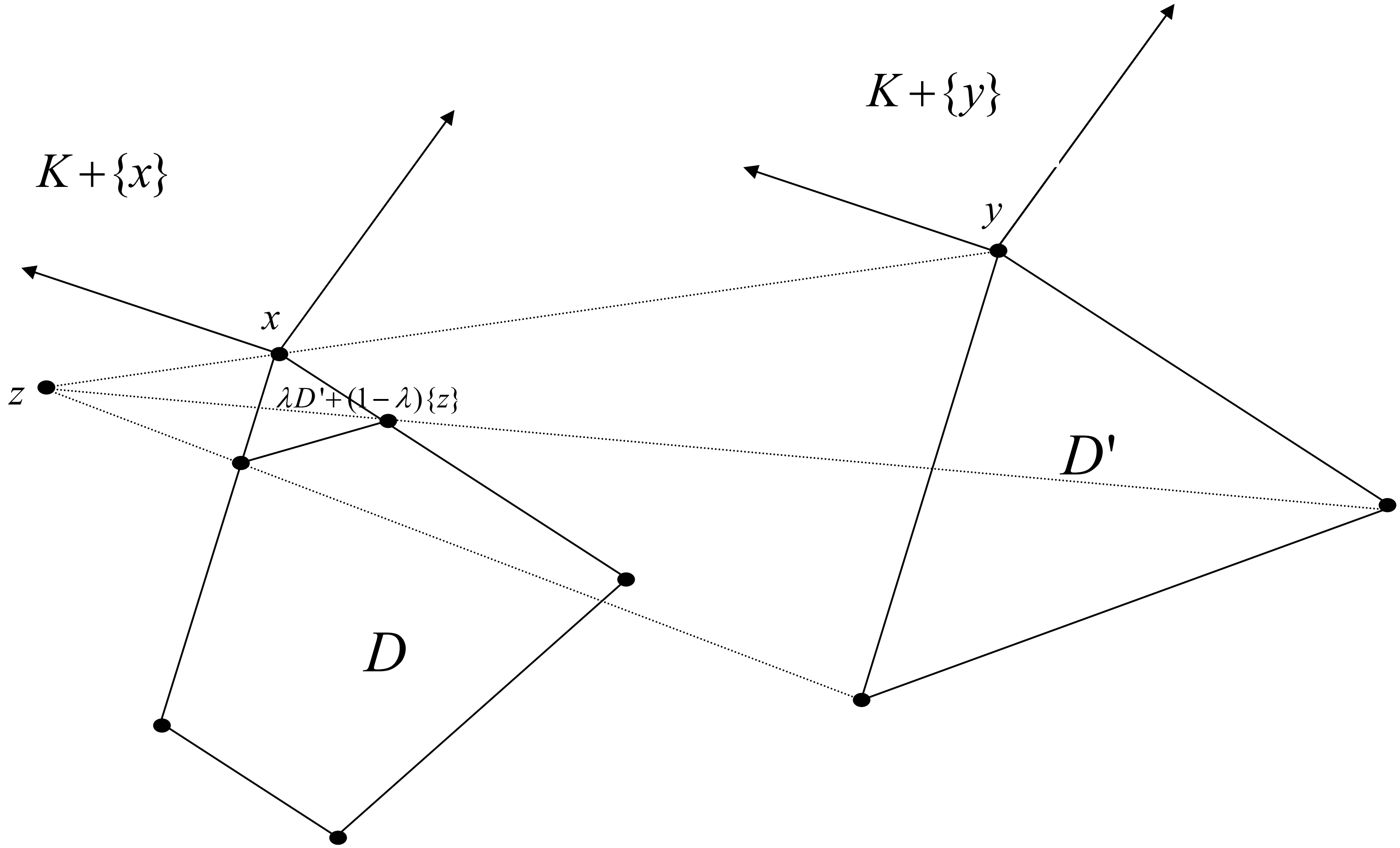
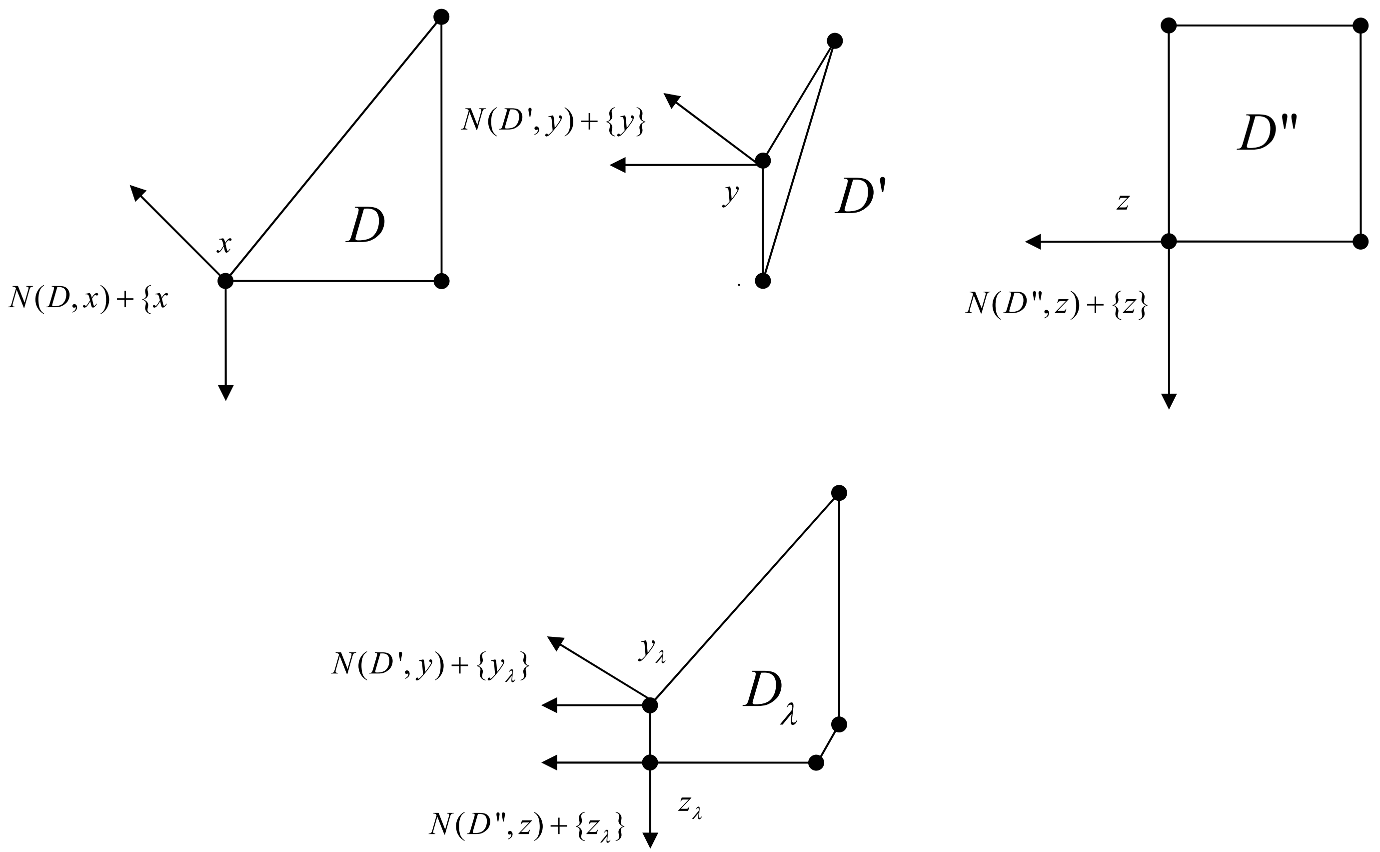


Figure 1



$N(D', y) + \{y\}$
D'
y
D''
z
x
D
$N(D, x) + \{x$
$N(D'', z) + \{z\}$
y_λ
$N(D', y) + \{y_\lambda\}$
D_λ
z_λ
$N(D'', z) + \{z_\lambda\}$

Figure 2

Supplement to Random Expected Utility[†]

Faruk Gul

and

Wolfgang Pesendorfer

Princeton University

August 2004

Abstract

This supplement provides a detailed discussion of the related literature in Gul and Pesendorfer (2004). In particular, we relate our results to the work of McFadden and Richter (1971), Clark (1995) and Falmagne (1978)

McFadden and Richter (1991) and Clark (1995) provide conditions that are necessary and sufficient for a random choice rule (RCR) to maximize a random utility function (RUF). The models of Clark and McFadden and Richter (1991) do not restrict to the linear structure of lotteries and von Neumann-Morgenstern utility functions. To facilitate a comparison, we adapt their conditions to the model analyzed Gul and Pesendorfer (2004).

[†] This research was supported by grants SES0236882 and SES0214050 from the National Science Foundation.

This adaptation is possible because both Clark's condition and McFadden and Richter's condition can be thought of as a joint restriction on random choice rules and the space of utility functions. The notation below is taken from section 2 of Gul and Pesendorfer (2004).

Clark (1995) introduces an axiom termed "coherency". Coherency is closely related to a theorem of De Finetti's which provides a necessary and sufficient condition for a function defined on a collection of subsets to have an extension to a finitely additive probability measure on the smallest algebra containing those subsets. Clark (1995) shows that a random choice rule is coherent if and only if it maximizes some regular random utility function.

The definition below adapts Clark's axiom to the setting of Gul and Pesendorfer (2004). For $A \subset U$ let I_A denote the indicator function on the set A. Hence $I_A(u) = 1$ if $u \in A$ and $I_A(u) = 0$ otherwise. We write $I_A \geq 0$ as a shorthand for $I_A(u) \geq 0 \, \forall u \in U$.

Definition: *The RCR ρ is coherent if for every finite sequence $\{D_i, x_i\}_{i=1}^m$, with $D_i \in \mathcal{D}, x_i \in D_i$ and every finite sequence $\{\lambda_i\}_{i=1}^m$ in $I\!R^m$*

$$\sum_{i=1}^{n} \lambda_i I_{N(D_i,x_i)} \geq 0 \textit{ implies } \sum_{i=1}^{n} \lambda_i \rho^{D_i}(x_i) \geq 0$$

Clark (1995) shows that coherency is necessary and sufficient for the existence of a regular RUF μ such that for all D and $x \in D$

$$\rho^D(x) = \mu(N(D, x))$$

Fact 1: *A coherent RCR ρ is mixture continuous, monotone, linear and extreme.*

Proof: To show extremeness, let $y \in D$ with $y \notin \operatorname{ext} D'$ and let $D = \operatorname{ext} D'$. Then $I_{N(D,x)} = I_{N(D',x)}$ for all $x \in D$ and therefore coherency implies

$$\sum_{x \in D} \rho^{D'}(x) = \sum_{x \in D} \rho^D(x) = 1$$

which in turn implies that $\rho^D(y) = 0$ and establishes extremeness.

To show monotonicity, let $D' = D \cup \{y\}$. Then $N(D, x) \supset N(D', x)$ for all $x \in D$ and therefore $I_{N(D,x)} - I_{N(D',x)} \geq 0$ which implies $\rho^D(x) \geq \rho^{D'}(x)$.

To show linearity and mixture continuity note that for any coherent RCR ρ

$$N(D, x) = N(D', x') \text{ implies } \rho^D(x) = \rho^{D'}(x') \tag{i}$$

and

$$N(D, x) = N(D', x') \cup N(D'', x'') \text{ implies } \rho^D(x) \leq \rho^{D'}(x') + \rho^{D''}(x'') \tag{ii}$$

Since $N(D, x) = N(\lambda D + (1 - \lambda)\{y\}, \lambda x + (1 - \lambda y)$ linearity follows from (i). Using (i) and (ii) it is straightforward to adapt the argument given in the proof of Theorem 2 in Gul and Pesendorfer (2004) to demonstrate that ρ is mixture continuous. $\square$

Clark's theorem implies that coherency is necessary for a random choice rule to maximize a regular random utility function. By Theorem 2 in Gul and Pesendorfer (2004) the maximizer of a regular random utility function is monotone, linear, mixture continuous and extreme. Together with Fact 1 this implies that a RCR is monotone, linear, mixture continuous and extreme if and only if it is coherent.

Coherency can also be applied in settings where we only observe the choice behavior in a subset of the possible decision problems. In that case, coherency is necessary and sufficient for the implied RUF to have an extension that is a probability measure. Thus whenever the observed choice probabilities satisfy coherency, one can construct a RUF μ such that the observed behavior is consistent with $\mu-$maximization.

Coherency is difficult to interpret behaviorally. Moreover, it seems difficult to construct experiments that "test" for coherency. By contrast, it seems quite straightforward to construct tests of extremeness, linearity and monotonicity. In fact, the experimental literature on expected utility has focused on the linearity axiom to point out violations of the expected utility framework and develop alternatives. This process of searching for violations of a theory and generalizing the theory to incorporate the documented violations requires interpretable axioms.

McFadden and Richter (1991) introduce a stochastic version of the strong axiom of revealed preference, an axiom they term *Axiom of Revealed Stochastic Preference (ARSP).* McFadden and Richter (1991) study a case where each utility function under consideration has a unique maximizer and show that ARSP is necessary and sufficient for (regular) random utility maximization.

In the definition below, we adapt ARSP to the framework in Gul and Pesendorfer (2004).

Definition: *The RCR satisfies ARSP iff for all* $(D_i, x_i)_{i=1}^m$ *with* $D_i \in \mathcal{D}, x_i \in D_i$

$$\sum_{i=1}^{m} \rho^{D_i}(x_i) \leq \max_{u \in U^*} \sum_{i=1}^{m} I_{N^+(D_i,x_i)}(u) \tag{$*$}$$

To see that ARSP is necessary for regular random utility maximization, note that if ρ maximizes a regular RUF μ, then

$$\sum_{i=1}^{m} \rho^{D_i}(x_i) = \int_U \sum_{i=1}^{m} I_{N^+(D_i,x_i)}(u)\mu(du)$$

Obviously, the r.h.s. of the equation above is less than or equal to the r.h.s. of ($*$).

Fact 2 below shows that ARSP implies monotonicity, linearity, extremeness and mixture continuity. Hence, Theorem 2 in Gul and Pesendorfer (2004) implies that a random choice rule satisfies ARSP if and only if it is monotone, linear, mixture continuous and extreme.

Fact 2: *If the RCR ρ satisfies ARSP then it is mixture continuous, monotone, linear and extreme.*

Proof: Extremeness is trivial because $N^+(D, x)$ is empty unless x is an extreme point of D.

For monotonicity, apply ARSP to $\{(D, x), (D\backslash\{y\}, z)_{z \neq x,y}\}$. This yields $\rho^D(x) \leq \rho^{D\backslash\{y\}}(x)$ and hence monotonicity.

Next, we show that

$$\rho^D(x) = \rho^{D'}(x') \text{ if } N^+(D, x) = N^+(D', x') \tag{i'}$$

Apply ARSP to $\{(D, x), (D', y)_{y \neq x'}\}$ to get $\rho^D(x) \leq \rho^{D'}(x')$. Reversing the roles of D and D' yields the reverse inequality and hence the result. Linearity now follows because $N^+(D, x) = N^+(\lambda D + (1-\lambda)\{y\}, \lambda x + (1-\lambda)y)$. To prove mixture continuity, we proceed as above. First, we show that (i) and (ii) in the proof of Fact 1 above hold. To prove (ii) apply ARSP to $\{(D, x), (D', y)_{y \neq x'}, (D'', y)_{y \neq x''}\}$. Since $N(D, x) = N(D', x')$ implies $N^+(D, x) = N^+(D', x')$ (i) follows from (i'). Using (i) and (ii) it is again straightforward to adapt the argument given in the proof of Theorem 2 to demonstrate that ρ is mixture continuous. $\square$

Falmagne (1978) studies a model with finitely many alternatives. Let Y be a finite set. A decision problem is a non-empty subset D of Y. Let $\mathcal{D}^*$ be the corresponding collection of decision problems. Let U^* be the set of all one-to-one utility functions on Y, and let $\mathcal{F}^*$ be the algebra generated by the equivalence relation that identifies all ordinally equivalent utility functions (i.e. $u \in F$ implies $v \in F$ if and only if $[v(x) \geq v(y)$ iff $u(x) \geq u(y)]$ for all $x, y \in Y$). Let Π^* denote the set of all probability measures on $\mathcal{F}^*$. Falmagne identifies a finite number (depending on the number of available alternatives) of non-negativity conditions as necessary and sufficient for random utility maximization.

Definition: *For any RCR ρ, define the difference function Δ of ρ inductively as follows: $\Delta_x(\emptyset, D) = \rho^D(x)$ for all $x \in D$ and $D \subset Y^*$. Let $\Delta_x(A \cup \{y\}, D) = \Delta_x(A, D) - \Delta_x(A, D \cup \{y\})$ for any $A, D \subset Y^*$ such that $x \in D$, $A \cap D = \emptyset$ and $y \in Y^* \backslash (A \cup D)$.*

Falmagne (1978) shows that the RCR ρ maximizes some $\mu \in \Pi^*$ if and only if $\Delta_x(A, Y \backslash A) \geq 0$ for all A and $x \in Y \backslash A$. This condition turns out to be equivalent to $\Delta_x(A, D) \geq 0$ for all x, A, D such that $A \cap D = \emptyset$ and $x \in D$.

Note that for $A = \{y\}$, the condition $\Delta_x(A, D) \geq 0$ for all $x \in D$, $y \notin D$ corresponds to our monotonicity assumption and says that the probability of choosing x from D is at least as high as the probability of choosing x from $D \cup \{y\}$. These conditions also require that the difference in the probabilities between choosing x from D and $D \cup \{y\}$ is decreasing as alternative z is added to D and that analogous higher order differences be decreasing. While monotonicity is a straightforward (necessary) condition, the higher order conditions are more difficult to interpret.

We can relate our theorem to Falmagne's if we interpret Y to be the set of extreme points of our simplex of lotteries P. Suppose Falmagne's conditions are satisfied and hence a RCR (on $\mathcal{D}^*$) maximizes some RUF μ. We can extend this RUF μ to a RUF $\hat{\mu}$ on our algebra $\mathcal{F}$ (i.e., the algebra generated by the normal cones $N(D,x)$) by choosing a single u from each $[u]$ and setting $\hat{\mu}(\{\lambda u \,|\, \lambda \geq 0\}) = \mu([u])$ where $[u]$ is the (equivalence) class of utility functions ordinally equivalent to u. Hence, $\hat{\mu}$ is a RUF on $\mathcal{F}$ that assigns positive probability to a finite number of rays and zero probability to all cones that do not contain one of those rays. By utilizing our Theorem 1, we can construct some mixture continuous, monotone, linear and extreme $\hat{\rho}$ that maximizes $\hat{\mu}$. This $\hat{\rho}$ must agree with ρ whenever $D \subset P$ consists of degenerate lotteries. Hence, any random choice functions that satisfies Falmagne's conditions can be extended to a random choice function over lotteries that satisfies our conditions. Conversely, if a Falmagne random choice function can be extended to a random choice function (on $\mathcal{F}$) satisfying our conditions, then by Theorem 2, this function maximizes a RUF. This implies that the restriction of this function to sets of degenerate lotteries maximizes a Falmagne RUF and satisfies the conditions above. Thus, Falmagne's conditions are necessary and sufficient for a random choice function over a finite set to have a mixture continuous, monotone, linear and extreme extension to the set of all lotteries over that set.

References

1. Barbera, S. and P.K. Pattanaik, "Falmagne and the Rationalizability of Stochastic Choices in Terms of Random Orderings", *Econometrica*, 1986, Vol. 54, pp. 707–715.

2. Clark, S. A., "The Random Utility Model with an Infinite Choice Space," *Economic Theory*, 1995, 7, 179–189.

3. Falmagne, J.-Cl., "A Representation Theorem for Finite Random Scale Systems", *Journal of Mathematical Psychology*, 1978, 18, pp. 52–72.

4. Gul, F and W. Pesendorfer, "Random Expected Utility", mimeo 2004.

5. McFadden, D. and M. Richter, "Revealed Stochastic Preference", in: J. S. Chipman, D. McFadden and M. K. Richter, editors, *Preferences, Uncertainty and Optimality.* Westview Press, Boulder, CO, 1991.