Random Expected Utility†

Faruk Gul

and

Wolfgang Pesendorfer

Princeton University

June 2003

Abstract

We analyze decision-makers who make stochastic choices from sets of lotteries. A random choice rule associates with each decision problem a probability measure over the feasible choices. A random utility is a probability measure over von Neumann-Morgenstern utility functions. We show that a random choice rule maximizes some random utility if and only if it is *mixture continuous*, *monotone* (the probability that x is chosen from a choice problem is non-increasing as alternatives are added to the choice problem), *extreme* (chooses an extreme point with probability one), and *linear* (satisfies the independence axiom).

† This research was supported by grants from the National Science Foundation.

1. Introduction

In this paper, we develop and analyze a model of random choice and random expected utility. Modelling 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, the 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 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 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.

Modelling random choice as a consequence of random utility maximization 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 model 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. Therefore, identifying the behavioral implications of random choice that results from random utility maximization has been a central concern of the random choice literature. This amounts to answering the following question: what conditions on ρ are necessary and sufficient for there to exist a random utility μ that is maximized by ρ?

We study behavior that results from 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, consider the demand analysis in a portfolio choice problem. Understanding random choice in this context 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 linear structure of the set of risky prospects facilitates the simpler conditions that we identify as necessary and sufficient.

One (trivial) example of a random utility is a measure that places probability 1 on the utility function that is indifferent between all choices. Clearly, this random utility is consistent with *any* behavior. A *regular* random utility is one where in any decision problem, with probability 1, the realized utility function has a unique maximizer. Hence, for a regular random utility ties are 0-probability events.

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) mixture continuity, (ii) monotonicity, (iii) linearity, and (iv) extremeness.

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 *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 *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.

[1] Sattath and Tversky (1976) use the same axiom and refer to it as regularity.

A random choice rule is *extreme* if extreme points of the choice set are chosen with probability 1. Extreme points are those elements of the choice problem that are unique optima for *some* von Neumann-Morgenstern utility function. Hence, if a random utility is regular, then the corresponding random choice rule must be extreme.

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

A deterministic utility function is a special case of a random utility. Clearly, it is not regular since there are choice problems for which ties occur with positive probability. However, we can use a tie-breaking rule that turns this non-regular random utility into a regular random utility. Using this tie-breaking rule, we establish that for *any* random utility μ there is a *regular* random utility μ' such that a maximizer of μ' is also a maximizer of μ.

When the random utility corresponds to a deterministic utility function, then the corresponding random choice rules will typically fail continuity (but satisfy mixture continuity). We show that this failure of continuity corresponds to a failure of countable additivity of the random utility. Put differently, suppose that a random choice rule maximizes a random utility. Then the random choice rule is continuous if and only if the random utility is countably additive. Our second main result is follows from this observation and our first result discussed above: *a random choice rule maximizes some regular, countably additive, random utility if and only if the random choice rule is continuous, monotone, linear and extreme.*

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. Of course, the independence axiom is not the only implication of expected utility theory. Our theorems identify all of implications of random expected utility maximization that are relevant for the typical experimental setting.

The majority of the work on random choice and random utility studies binary choice; that is, the case where $\mathcal{D}$ consists of all two-element subsets of some finite set Y. In order to avoid the ambiguities that arise from indifference, it is assumed that U consists of one-to-one functions. Since there is no way to distinguish ordinally equivalent utility functions, a class of such functions is viewed as a realization of the random utility. Fishburn (1992) offers an extensive survey of this part of the literature.

There are three strands of literature that have investigated the implications of random utility maximization in situations where the choice sets may not be binary.

McFadden and Richter (1970) provide a condition that is analogous to the strong axiom of revealed preference of demand theory and show that this condition is necessary and sufficient for maximizing a randomly drawn utility from the set of strictly concave and increasing functions. Applying this theory to a portfolio choice problem would require additional restrictions on the admissible utility functions. These restrictions in turn imply restrictions on observable behavior beyond those identified by McFadden and Richter. The contribution of this paper is to identify the additional restrictions that result from expected utility maximization.

Clark (1995) provides a test for verifying (or falsifying) if any (finite or infinite) data set is consistent with expected utility maximization. Falmagne (1978), Barbera and Pattanaik (1986) study the case where choice problems are arbitrary subsets of a finite set of alternatives. Their characterization of random choice identifies a finite number (depending on the number of available alternatives) of non-negativity conditions as necessary and sufficient for random utility maximization.

In section 5, we provide a detailed discussion of the relationship between our results and those provided by McFadden and Richter (1970), Clark (1995), and Falmagne (1978).

2. Random Choice and Random Utility

There is a finite set of prizes denoted $N = \{1, 2, \dots, n+1\}$ for $n \geq 1$. Let P be the unit simplex in $\mathbb{R}^{n+1}$ and $x \in P$ denote 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. The agent makes random choices when confronted with a decision problem. Let $\mathcal{B}$ denote the Borel sets of P and Π be set of all probability measures on the measurable space $\{P, \mathcal{B}\}$.

A *random choice rule* is a function $\rho : \mathcal{D} \to \Pi$ with $\rho^D(D) = 1$. The probability measure ρ^D with support 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 when faced with the decision problem D and write $\rho^D(x)$ instead of $\rho(D)(\{x\})$.

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

A *random utility* is a probability measure defined on an appropriate algebra of U. 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 then the agent must choose an element in $M(D, u)$. Conversely, when the choice $x \in D$ is observed then 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 $\mathcal{F}$ be the smallest field (algebra) that contains $N(D, x)$ for all (D, x). A random utility is a finitely additive probability measure on $\mathcal{F}$.

Definition: *A random utility 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 $F, F' \in \mathcal{F}$. A random utility μ*

is countably additive if $\sum_{i=1}^{\infty} \mu(F_i) = \mu(\bigcup_{i=1}^{\infty} F_i)$ *whenever* $F_i, i = 1, \ldots$ *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 random utility 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 random utility.

Next, we define what it means for a random choice rule to maximize a random utility. For $x \in D$, let

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

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

If $u \in U$ does not have a unique maximizer in D then the resulting choice from D is ambiguous. Since $N^+(D, x)$ contains all the utility functions that have x as the unique maximizer, the set $\bigcup_{x \in D} N^+(D, x)$ is the set of utility functions that have a unique maximizer in D. If $\mu(\bigcup_{x \in D} N^+(D, x)) < 1$ there is a positive probability of drawing a utility function for which the resulting choice is ambiguous. For such μ, it is not possible to identify a unique random choice rule as the maximizer of ρ. Conversely, if random utility functions such that $\mu(\bigcup_{x \in D} N^+(D, x)) < 1$ are allowed, the hypothesis of random utility maximization loses its force. For example, let $u^o = (\frac{1}{n}, \ldots, \frac{1}{n}, 0) \in U$ denote the utility function that is indifferent between all prizes. Consider the random utility μ_{u^o} such that $\mu(F) = 1$ if and only if $u^o \in F$. The random utility μ_{u^o} is the degenerate measure that assigns probability 1 to every set that contains u^o. An agent whose random utility is μ_{u^o} will be indifferent with probability 1 among all $x \in D$ for all $D \in \mathcal{D}$. To avoid this difficulty, the literature on random utility maximization restricts attention to random utilities that generate ties with probability 0. We call such random utilities *regular*.

Definition: *The random utility* 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 $x \in D$.

When there are two prizes ($n+1=2$) the set U consists of all the linear combinations of the vectors $(1,0)$ and $(-1,0)$. In this case, there are three distinct (von Neumann-Morgenstern) utility functions, corresponding to the vectors $u=(0,0), u'=(1,0), u''=(-1,0)$. The algebra $\mathcal{F}$ in this case consists of all unions of the sets $\emptyset, F_0, F_1, F_2$ where $F_0=\{(0,0)\}$, $F_1=\{\lambda(1,0)|\lambda>0\}$ and $F_2=\{\lambda(-1,0)|\lambda>0\}$.

With two prizes the random utility 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. Note that F_0 has dimension 0 whereas the other non-empty algebra elements have dimension 1. Hence, regularity is equivalent to assigning a zero probability to the lower dimensional element in the algebra $\mathcal{F}$. Lemma 1 shows that this characterization of regularity holds for all n. A random utility μ is regular if and only if μ is *full-dimensional*, i.e., $\mu(F)=0$ for every $F\in\mathcal{F}$ that has dimension $k<n$.[2]

A random choice rule ρ maximizes the regular random utility 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. Thus, the random choice rule ρ maximizes the regular random utility μ if

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

for all D.

As note above, a single expected utility function $\bar{u}$ can be viewed as a special random utility $\mu_{\bar{u}}$, where $\mu_{\bar{u}}(F)=1$ if $\bar{u}\in F$ and $\mu_{\bar{u}}(F)=0$ if $\bar{u}\notin F$. In the case with two prizes the random utility $\mu_{\bar{u}}$ 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 $\bar{u}$. To see this, note that the set $F=\{u=\lambda\bar{u}$ for $\lambda>0\}$ is an element of $\mathcal{F}$ with $\mu_{\bar{u}}(F)>0$ but F has dimension $1<n$.

Thus we can view deterministic utility functions as random utility but typically not as regular random utilities. To extend the concept of random expected utility maximization to all (not necessarily regular) random utilities, we introduce to notion of a tie-breaker. Let μ be any random utility function. Suppose that the agent with random utility μ draws the utility function u when facing the choice problem D. If the set of maximizers of u in D

[2] The dimension of F is the dimension of the affine hull of F.

(denoted $M(D,u)$) is a singleton, then the agent chooses the unique element of $M(D,u)$. If the set $M(D,u)$ is not a singleton then the agent draws another $\hat{u}$ according to some random utility $\hat{\mu}$ to decide which element of $M(D,u)$ to choose. If $\hat{\mu}$ chooses a unique maximizer from each $M(D,u)$ with probability 1, this procedure will lead to the following random choice rule:

$$\rho^D(x) = \int \hat{\mu}(N(M(D,u),x)\mu(du) \tag{2}$$

The integral in (2) is the Lebesgue integral. Lemma 2 shows that the integral in (2) is well-defined for all μ and $\hat{\mu}$. Thus to ensure that ρ defined by (2) is indeed a random choice rule, we need only verify that $\sum_{x \in D} \rho^D(x) = 1$ for all $D \in \mathcal{D}$. Lemma 3 ensures that this is the case whenever $\hat{\mu}$ is regular.

Definition: *The random choice rule ρ maximizes the random utility μ if there exists some random utility $\hat{\mu}$ (a tie-breaker) such that (2) is satisfied.*

The definition above extends the definition of random utility maximization to all random utilities. Note that we require the tie-breaking rule not to vary with the decision problem. Hence, we do not consider cases where the agent uses one tie-breaking rule for the decision problem D and a different one for the decision problem D'. A non-regular random utility together with this type of a tie-breaker can be interpreted as a regular random utility with a lexicographically less important dimension.

Note that for a regular random utility μ this definition reduces to the definition in equation (1). In particular, if $\mu, \hat{\mu}$ are random utilities and μ is regular, then

$$\int \hat{\mu}(N(M(D,u),x)\mu(du) = \mu(N(D,x))$$

To see this, first note that $\int \hat{\mu}(N(M(D,u),x)\mu(du) = \int_{N^+(D,x)} \hat{\mu}(N(M(D,u),x)\mu(du)$ since μ is regular. If $N^+(D,x) = \emptyset$ then obviously $\int_{N^+(D,x)} \hat{\mu}(N(M(D,u),x)\mu(du) = 0 = \mu(N(D,x)$. If $N^+(D,x) \neq \emptyset$ then

$$\begin{aligned}
\int_{N^+(D,x)} \hat{\mu}(N(M(D,u),x)\mu(du) &= \int_{N^+(D,x)} \hat{\mu}(N(\{x\},x))\mu(du) \\
&= \int_{N^+(D,x)} \hat{\mu}(U)\mu(du) \\
&= \mu(N^+(D,x)) \\
&= \mu(N(D,x))
\end{aligned}$$

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

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

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.

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

Definition: *The random choice rule 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 random choice rule 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 from the probability of choosing x from D' no matter how close D and D' are with respect to the Hausdorff metric.

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: *A random choice rule 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. 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.

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.

For any $D, D' \subset \mathcal{D}$ and $\lambda \in [0,1]$, let $\lambda D + (1-\lambda)D' := \{\lambda x + (1-\lambda)y \,|\, x \in D, y \in D'\}$. Note that if $D, D' \in \mathcal{D}$ then $\lambda D + (1-\lambda)D' \in \mathcal{D}$.

Definition: *A random choice rule 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 in familiar contexts of choice under uncertainty. 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 $\operatorname{ext} D$. 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: *A random choice rule ρ is extreme if $\rho^D(\operatorname{ext} D) = 1$.*

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* random utility must choose an extreme point with probability 1. Hence, extremeness is a necessary condition of maximization of a regular random utility.

4. Results

In Theorem 1, we establish that the notion of random utility maximization presented in section 2 can be applied the all random utilities. That is, every random utility can be maximized. Theorem 1 also establishes that regularity of the random utility is necessary and sufficient for the existence of a unique maximizer.

Theorem 1: *(i) Let μ be a random utility. Then, there exists random choice rule ρ that maximizes μ. (ii) The random utility μ has a unique maximizer if and only if it is regular.*

Proof: See section 7.

To provide intuition for Theorem 1, consider first a regular random utility μ. Let ρ be defined by equation (1) in section 2. That is, $\rho^D(x) = \mu(N(D,x))$ for all x, D. Then, the regularity of μ implies

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

for all D, x. Recall that x is an extreme point of D if $N^+(D,x) \neq \emptyset$. Hence, $\sum_{x\in D} \rho^D(x) = \sum_{x\in D} \mu(N^+(D,x)) = 1$ establishing that ρ is a random utility. Note that equation (1) uniquely identifies the maximizer of a regular μ.

When μ is not regular, we can choose any regular random utility $\hat{\mu}$ and define ρ by (2). Since $\hat{\mu}$ is regular, Lemma 3 implies that ρ is a random choice rule. By changing the tie-breaker $\hat{\mu}$, we can generate a different random choice rule that also maximizes μ. Hence, regularity is necessary for μ to have a unique maximizer.

Theorem 2 below is our main result. It establishes that monotonicity, mixture continuity, extremeness and linearity are necessary and sufficient for ρ to maximize a random utility.

Theorem 2: *Let ρ a random choice rule. There exists a regular random utility μ such that ρ maximizes μ if and only if ρ is mixture continuous, monotone, linear and extreme.*

Proof: See section 7.

We briefly sketch the proof of Theorem 2. First assume that ρ maximizes μ and, for simplicity, assume that μ is regular. Hence, μ and ρ satisfy (1). The choice rule ρ is monotone since $N(D \cup \{y\}, x) \subset N(D, x)$ whenever $x \in D$; it is linear since $N(D, x) = N(\lambda D + (1-\lambda)\{y\}, \lambda x + (1-\lambda)y)$. Since $N^+(D, x) = \emptyset$ whenever $x \notin \text{ext}\, D$, extremeness follows immediately from the fact that μ is regular and therefore $\mu(N^+(D, x)) = \mu(N(D, x))$. For the proof of mixture continuity see Lemma 7.

To prove that mixture continuity, monotonicity, linearity and extremeness are sufficient for random utility maximization, we first show that monotonicity, linearity and extremeness of ρ imply $\rho^D(x) = \rho^{D'}(y)$ whenever $N(D, x) = N(D, y)$ (Lemma 6). To get intuition for the proof of Lemma 6, 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 6.

Finite additivity is proven in Lemma 8. 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 6, 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 $B \cap D = \{x\}$. 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.

In the proof Theorem 1 we show that every mixture continuous, monotone, linear and extreme random choice rule maximizes some random utility μ by constructing a random utility μ such that $\rho^D(x) = \mu(N(D,x))$ for all D, x. Since ρ is extreme μ must be regular. Then, it follows from the converse implication of Theorem 2 that a random choice rule maximizes some random utility if and only if it maximizes a regular random utility.

Corollary 1: *Let ρ be a random choice rule. Then, ρ maximizes some random utility μ if and only if it maximizes some regular random utility.*

Proof: In the proof Theorem 2 we have shown that if ρ maximizes some random utility then it is mixture continuous, monotone, linear, and extreme. We have also shown that if ρ is mixture continuous, monotone, linear and extreme then, there exists a random utility μ such that $\rho^D(x) = \mu(N(D,x))$ for all D, x. To conclude the proof, we observe that this μ is regular: From Proposition 1(*iii*) and extremeness we infer that μ is full-dimensional. Lemma 1 then implies that μ is regular. □

Example 1 below considers a random utility $\mu_{\bar{u}}$ that corresponds to a single (deterministic) utility function. It shows that maximizers of $\mu_{\bar{u}}$ are not continuous. Moreover, if μ is a regular random utility that has the property that the maximizer of μ is also a maximizer of $\mu_{\bar{u}}$ then μ is not countably additive.

Example 1: Consider the case of three prizes ($n+1=3$) and the (non-regular) random utility $\mu_{\bar{u}}$ for $\bar{u} = (-2, -1, 0)$. That is, $\mu_{\bar{u}}$ is the random utility associated with deterministic utility function $\bar{u}$. Let ρ be a random choice rule that maximizes $\mu_{\bar{u}}$. First, we observe that the ρ is not continuous. To see this, let $x = (0,1,0), y = (.5, 0, .5)$ and $z = .5(x+z)$. For $k > 4$, let $z_k = z + \frac{1}{k}\bar{u}$ and let $z_{-k} = z - \frac{1}{k}\bar{u}$. Let $D_k = \{x, y, z_k\}, D_{-k} = \{x, y, z_{-k}\}$ and $D = \{x, y, z\}$. Note that $\bar{u} \cdot z_{-k} < \bar{u} \cdot x = u \cdot y < u \cdot z_k$ for all $k = 4, 5, \ldots$. Hence, $\rho^{D_k}(z_k) = 1$ and $\rho^{D_{-k}}(z_{-k}) = 0$. Since z_k and z_{-k} converge to z, we have $\rho^{D_k}(O) = 1$ and $\rho^{D_{-k}}(O) = 0$

for k sufficiently large and any open ball that contains z but does not contain x and y. Since D_k and D_{-k} converge to D this establishes that ρ is not continuous. Next, we show that the failure of continuity of ρ implies that the corresponding regular random utility μ is not countably additive. Clearly we have $\mu(N^+(D_k, x)) = 0$ and $\mu(N^+(D_k, y)) = 0$ for all $k > 4$ since x and y are not chosen from D_k for any k. However, $\bigcup_{k>4} N^+(D_k, x) = N^+(D, x)$ and $\bigcup_{k>4} N^+(D_k, y) = N^+(D, y)$ and $\mu(N^+(D, x)) + \mu(N^+(D, y)) = 1$ because μ is regular and x, y are the only extreme points of D. Therefore μ is not countably additive.

Note that the failure of continuity shown for Example 1 will typically result if the random utility corresponds to a deterministic utility function. More precisely, assume that $n + 1 \geq 3$ and consider a random utility $\mu_{\bar{u}}$ such that $\bar{u} \neq (0, ..., 0)$. Then any ρ that maximizes $\mu_{\bar{u}}$ is not continuous.[3] Moreover, if ρ is mixture continuous, monotone, linear and extreme then the corresponding regular random utility will fail countable additivity. Theorem 3 below shows this relation between countable additivity of a regular random utility and continuity of its maximizer holds generally.

Theorem 3: *Let ρ maximize the regular random utility μ. Then, ρ is continuous if and only if μ is countably additive.*

Proof: See appendix.

Corollary 1 and Theorem 3 yield the following characterization of the countably additive random utility model.

Corollary 2: *Let ρ be a random choice rule. Then, there exists a regular, countably additive random utility μ such that the random choice rule ρ maximizes μ if and only if ρ is continuous, monotone, linear and extreme.*

Proof: Suppose ρ is continuous, monotone, linear and extreme. By Theorem 1 and Corollary 1 there exists a regular μ such that ρ maximizes μ. Since ρ is continuous, Theorem 3 implies μ is countably additive. For the converse, assume that ρ maximizes the regular,

[3] The argument given for the failure of continuity in Example 1 can easily be modified to deal with the more general case.

countably additive random utility μ. Theorem 2 establishes that ρ is monotone, linear and extreme. Theorem 3 implies that ρ is continuous. □

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

Definition: *A random choice rule 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: *Let ρ be a random choice rule. Then, there exists a regular, countably additive random utility μ such that the random choice rule ρ maximizes μ if and only if ρ is continuous, monotone, linear and undominated.*

Proof: see Section 8. □

To prove Theorem 4, we show that a continuous random choice rule is extreme if and only if it is undominated. Then the result follows from Corollary 2.

Note that as an alternative to the finite choice problems analyzed in this paper, we could have identified each choice problem D with its convex hull and chosen the collection of polytopes as the domain of choice problems. With the exception of Theorem 4 all our results hold for this alternative domain. However, an undominated and continuous choice rule may not be extreme if the choice problems are polytopes.[4] Theorem 4 is true, however, if the domain is the union of all finite choice problems and all polytopes.

[4] This alternative domain was suggested by a referee. The editor provided a counter-example to Theorem 4 for this case: consider a uniform distribution over the boundary points of the polytope. This random choice rule satisfies the four properties of Theorem 4 but is obviously not extreme and hence does not maximize a random utility.

5. Counterexamples

In this section, we provide examples that show that none of the assumptions in Theorem 2 and 4 and in Corollaries 1 and 2 are redundant. Example 2 provides a random choice rule that is continuous (hence mixture continuous), linear and extreme (hence undominated) but not monotone. This shows that monotonicity cannot be dispensed with.

Example 2: 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

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

Hence, $a(D)$ is the largest open interval that does not intersect D, but is contained in the convex hull of D. Let $\rho^D(x) = 0$ for $x \notin \{\underline{m}(D), \overline{m}(D)\}$. If D is a singleton, the ρ^D is defined in the obvious way. Otherwise, let

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

and

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

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

The next example provides a random choice rule 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 Theorem 2 and the Corollaries.

Example 3: Let $n+1=2$ and let $x \in [0,1]$ denote the probability of getting prize 2. For any $D = \{x_1, \ldots, x_m\}$, where $x_1 < x_2 <, \ldots, < x_m$, let

$$\rho^D(x_1) = \begin{cases} 1 & \text{if } m = 1 \\ 0 & \text{otherwise.} \end{cases}$$

For $k > 1$, let

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

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

Example 4 provides a random choice rule 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 and 4 and the Corollaries.

Example 4: Let $n+1=2$ and let $x \in [0,1]$ denote the probability of getting prize 2. As in Example 2, let $\underline{m}(D)$ and $\overline{m}(D)$ be the smallest and largest elements in D. Let

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

and

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

Then, ρ is continuous, monotone and extreme but not linear.

The final example constructs a random choice rule 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 Corollary 1 and continuity cannot be dispensed with in Corollary 2 and Theorem 4.

Example 5: Let $n+1=3$ and assume that ρ is defined as follows. Each extreme point is chosen either with probability 0 or with probability 1/2. For any choice problem D that has two extreme points (and therefore has dimension 1) each extreme point is chosen with probability $1/2$. For any choice problem that has more than 2 extreme points (and therefore has dimension 2) the extreme point $x \in D$ is chosen probability $1/2$ if (and only if) $N(D,x)$ contains $(1,-1,0)$ or $(-1,1,0)$ and $N^+(D,x)$ contains some $u=(u_1,u_2,u_3)$ such that $u_1+u_2>0$.

Note that this random choice rule corresponds to a random utility model with a tie-breaking rule that depends on the dimension of the decision problem. The agent draws either the utility function $(1,-1,0)$ or the utility function $(-1,1,0)$ each with probability

$^1\!/_2$. If he faces a choice problem with 2 extreme points (and hence $N(D,x)$ is a halfspace) then he breaks ties by randomizing uniformly. Therefore, each extreme point of a 1-dimensional decision problem is chosen with probability 1/2. If the agent faces a choice problem with more than two extreme points (and hence $N(D,x)$ is "smaller" than a halfspace) then he breaks ties by choosing points that maximize the utility function $(^1\!/_2, {}^1\!/_2, 0)$.

This random choice rule is extreme by definition. It is linear because the probability of choosing x from D depends only on $N(D,x)$. As we argued in the intuition for Theorem 1, the set $N(D,x)$ is invariant to linear translations of D and therefore the choice rule is linear. 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, if $N(D,x) \subset N(D',y)$ then the probability of choosing y from D' is at least as large as the probability of choosing x from D. Since $N(D \cup \{y\}, x) \subset N(D,x)$, monotonicity follows. It remains to show that the choice rule is not mixture continuous. Let $D = \left\{ \left(^1\!/_4, {}^1\!/_2, {}^1\!/_4\right), \left(^1\!/_2, {}^1\!/_4, {}^1\!/_4\right) \right\}$ and let $D' = \left\{ \left(^3\!/_8, {}^3\!/_8, {}^1\!/_4\right), \left(^1\!/_8, {}^1\!/_8, {}^3\!/_4\right) \right\}$. For $\lambda > 0$ the agent chooses from $\lambda D + (1-\lambda)D'$ either $\lambda \left(^1\!/_4, {}^1\!/_2, {}^1\!/_4\right) + (1-\lambda)\left(^3\!/_8, {}^3\!/_8, {}^1\!/_4\right)$ or $\lambda \left(^1\!/_2, {}^1\!/_4, {}^1\!/_4\right) + (1-\lambda)\left(^3\!/_8, {}^3\!/_8, {}^1\!/_4\right)$, each with probability $^1\!/_2$. For $\lambda = 0$ the agent chooses $\left(^3\!/_8, {}^3\!/_8, {}^1\!/_4\right)$ or $\left(^1\!/_8, {}^1\!/_8, {}^3\!/_4\right)$ each with probability $^1\!/_2$. Clearly, this violates mixture continuity at $\lambda = 0$.

6. Related Literature

In order to compare results from McFadden and Richter (1970), Falmagne (1978), and Clark (1995) to our own, we present a framework general enough to include all the models presented in the four papers. This framework consists of a random choice structure $\mathcal{C} = \{Y^*, \mathcal{D}^*, \mathcal{B}^*\}$ and a random utility space $\mathcal{U} = \{U^*, \mathcal{F}^*, \Pi^*\}$, where Y^* is the set of choice objects $\mathcal{D}^*$ is the set of decision problems (i.e., a collection of subsets of Y^*), $\mathcal{B}^*$ is an algebra on Y^* such that $\mathcal{D}^* \subset \mathcal{B}^*$, U^* is a set of utility functions, $\mathcal{F}^*$ is an algebra on U^* and Π^* is a set of probability measures on $\mathcal{F}^*$.

For any $(\mathcal{C}, \mathcal{U})$, $D \in \mathcal{D}^*$, $u \in U^*$, $x \in D$, define

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

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

$$N^+(D,x) = \{v \in U \,|\, v(x) > v(y) \forall y \in D \backslash \{x\}\}$$

A model $(\mathcal{C},\mathcal{U})$ is a random choice structure $\mathcal{C}$ and a random utility space $\mathcal{U}$ such that $\mathcal{F}^*$ is the smallest algebra that contains all sets of the form $N(D,x)$ for all $D \in \mathcal{D}^*$ and $x \in D$. Given a model $(\mathcal{C},\mathcal{U})$, a random choice rule is a function ρ that associates a probability measure ρ^D on the algebra $\mathcal{B}$ such that $\rho^D(D) = 1$. A random utility is a finitely additive probability measure on $\mathcal{F}^*$. The random choice rule ρ maximizes the random utility μ if and only if

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

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

For any $u \in U$ and $B \in \mathcal{B}$, let $I_{M(D,u)}(B) = 1$ if $M(D,u) \subset B$ and $I_{M(D,u)}(B) = 0$ otherwise. McFadden and Richter (1970) study a case where ties occur with probability zero, that is, $\mu(N^+(D,x)) = \mu(N(D,x))$. McFadden and Richter prove the following result: There exists $\mu \in \Pi^*$ such that ρ maximizes μ if and only if for all $(D_i, B_i)_{i=1}^m$ such that $D_i \in \mathcal{D}^*$, $B_i \in \mathcal{B}^*$ for $i = 1, \ldots, m$

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

To see that the McFadden-Richter condition is necessary for random utility maximization, note that if ρ maximizes $\mu \in \Pi^*$, then

$$\sum_{i=1}^m \rho^{D_i}(B_i) = \int_{u \in U^*} \sum_{i=1}^m I_{M(D_i,u)}(B_i)\mu(du)$$

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

To relate the McFadden-Richter conditions to our Theorem 2, we apply them to our framework and show that they imply monotonicity, linearity, extremeness and mixture continuity. Thus, we can use Theorem 2 to prove a version of the McFadden-Richter theorem in our setting. Let $\mathcal{C} = \{P, \mathcal{D}, \mathcal{B}\}$ and $\mathcal{U} = \{U, \mathcal{F}, \Pi^*\}$ and $\Pi^* \subset \Pi$ be the set of all regular random utilities on $\mathcal{F}$.

Monotonicity: Applying the McFadden-Richter conditions condition to

$$(D, \{x\}), (D\backslash\{y\}, D\backslash\{x, y\})$$

yields $\rho^D(x) \leq \rho^{D\setminus\{y\}}(x)$ and hence monotonicity.

Linearity: Applying the McFadden-Richter conditions to

$$(D, B_1),\ (\lambda D + (1-\lambda)\{y\}, B_2)$$

with $B_1 = \{x\}$ and $B_2 = \lambda(D\setminus\{x\}) + (1-\lambda)\{y\}$ yields

$$\rho^D(x) \leq \rho^{\lambda D+(1-\lambda)\{y\}}(\lambda x + (1-\lambda)y)$$

A symmetric argument for $B_1 = D\setminus\{x\}, B_2 = \{\lambda x+(1-\lambda)y\}$ yields the opposite inequality and establishes linearity.

Extremeness: To see that the McFadden-Richter conditions yield extremeness note that $I_{M(D,u)}(B) = 0$ unless B contains an extreme point of D.

Mixture Continuity: Using Proposition 3 of the next section, it can be shown that the McFadden-Richter conditions also imply mixture continuity.

Clark (1995) studies the case where Y^* is arbitrary and $\mathcal{D}^*$ is any (finite or infinite) collection of choice sets. He assumes that each $D \in \mathcal{D}^*$ is finite and each $u \in U^*$ has a unique maximizer in each D. Then, the collection of choice probabilities $\rho^D(x)$ such that $D \in \mathcal{D}^*$ and $x \in D$ induce a function $\mu : \mathcal{N} \to [0,1]$ where $\mathcal{N} := \{N(D,x) \,|\, D \in \mathcal{D}^*, x \in D\}$. He provides a condition on the choice probabilities $\rho^D(x)$ that is necessary and sufficient for μ to have an extension to $\mathcal{F}^*$ that is a probability measure. Thus whenever the observed choice probabilities satisfy his condition, one can construct a random utility μ such that the observed behavior is consistent with $\mu-$maximization. Clark's condition on observed choice probabilities is 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.

If a finite data set satisfies Clark's condition then there is a random utility that could have generated the data. Conversely, if a finite data set is inconsistent with random utility maximization then Clark's conditions will detect this inconsistency. Hence, Clark's condition provides the most powerful test of random utility maximization. This is in contrast to the conditions given in McFadden and Richter (1970) and the axioms in this paper. A finite data set may not violate any of our axioms but nevertheless be inconsistent with random utility maximization. However, Clark's condition is difficult to interpret behaviorally. By contrast, our conditions have a straightforward economic interpretation.

As we have done in the case of the McFadden-Richter theorem, we can relate Clark's theorem to our Theorem 2 by letting $\mathcal{C} = \{P, \mathcal{D}, \mathcal{B}\}$ and $\mathcal{U} = \{U, \mathcal{F}, \Pi^*\}$ as above and using his condition to establish monotonicity, linearity, extremeness and mixture continuity. Given Proposition 3 (to be used for verifying mixture continuity), deriving these properties from Clark's property is not difficult. Hence, we can prove a version of Clark's theorem (one that applies only when all choice problems are observable) by utilizing Theorem 2.

Falmagne (1978) studies the case where Y^* is any finite set, $\mathcal{B}$ is the algebra of all subsets of Y^*, U^* is the set of all one-to-one utility functions on Y^*, $\mathcal{F}^*$ is 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 $x, y \in Y^*$), and Π^* is the set of all probability measures on $\mathcal{F}^*$. 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 random utility maximization. Formally,

Definition: *For any random choice rule 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 random choice rule 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 by considering Y^* as the set of extreme points of our simplex of lotteries P. Suppose, Falmagne's conditions are satisfied and hence ρ maximizes some random utility μ. We can extend this μ to a random utility $\hat{\mu}$ 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 $\hat{\mu}(\{\lambda u \mid \lambda \geq 0\}) = \mu([u])$ where $[u]$ is the (equivalence) class of utility functions ordinally equivalent to u. Hence, $\hat{\mu}$ is a random utility 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 random utility. This implies that the restriction of this function to sets of degenerate lotteries maximizes a Falmagne random utility 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.

7. Preliminaries

In this section, we define the concepts and state results from convex analysis that are used in the proofs. The proofs of the Propositions can be found in the appendix. 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 $x_1, \ldots, x_k$. We let $\text{aff}\, A$ ($\text{pos}\, A$, $\text{conv}\, A$) 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 relative interior A, denoted $\operatorname{ri} A$, is the interior of A in the relative topology of $\operatorname{aff} A$.

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 $E^* := \{u \in I\!R^n \,|\, |u^i| = 1$ for some i 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 *pointed* polyhedral cones, that is, cones 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 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 $\operatorname{ri} A \cap \operatorname{ri} A' = \operatorname{ri}(A \cap A')$.*

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 $u \in \operatorname{ri} N(A,x)$.*

Proposition 3: *Let A_i be polytopes, for $i = 1, \dots, m$. Then,*

$$N(A_1 + \cdots + A_m, \sum_i x_i) = \bigcap_{i=1}^{m} N(A, x_i)$$

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

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 $i = 1, \dots, m$ 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 $\mathbb{R}^n$.*

Let $\mathcal{F}$ be the smallest field that contains all polyhedral cones and let $\mathcal{H} := \{\operatorname{ri} K \,|\, K \in \mathcal{K}\} \cup \emptyset$. 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 $A_1, \dots, A_m \in \mathcal{P}$ such that $\bigcup_i A_i = A \backslash B$.

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

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) $\bigcup_{y \in D_i \cap B_\epsilon(x)} N(D_i, y) \subset K_j$ for $i > k_j$.*

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 $o \in D \cap D'$, $K = N(D, o)$, $K' = N(D', o)$, $d_h(D, D') < \epsilon$ and $K \cap S \subset O \subset K'$.*

8. Proofs

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

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

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

Let $\hat{\rho}^{\hat{D}}(x) = \rho^{T_1(\hat{D})}(T_1(x))$. We extend the random choice rule $\hat{\rho}$ to all finite nonempty subsets of $I\!R^n$ in the following manner: Choose $z \in \operatorname{int} \hat{P}$. For $D \subset I\!R^n$ 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 random choice rule, the following definitions of linearity and mixture continuity will be used.

Definition: *A random choice rule 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: *A random choice rule 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 $\mathbb{R}^n$ so that ρ refers to the corresponding $\hat{\rho}$ and μ to $\mu \circ T_0$.

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

Lemma 1: *A random utility* μ *is full-dimensional if and only if it is regular.*

Proof: Suppose μ is full-dimensional. Clearly $\mathbb{R}^n = \bigcup_{x \in \text{ext}\, D} N(D,x)$ and by Proposition 5(i), $\bigcup_{x \in \text{ext}\, D} N(D,x) = \bigcup_{x \in \text{ext}\, D} \bigcup_{B \in F^0(N(D,x))} B$. By Proposition 1($i$) $\text{int}\, N(D,x) = \text{ri}\, N(D,x) \subset N^+(D,x)$. Therefore, $\mathbb{R}^n = \bigcup_{x \in \text{ext}\, D} N^+(D,x) \cup F$ where F is a finite union of polyhedral cones of dimension less than n. Since μ is full-dimensional $\mu(F) = 0$ and $\mu\left(\bigcup_{x \in D} N^+(D,x)\right) = 1$.

If μ is not full-dimensional then there exists a set $F \in \mathcal{F}$ such that $\dim F < n$ and $\mu(F) > 0$. Since $\mathcal{H}$ is a semiring, every element of $\mathcal{F}$ can be written as a finite union of elements in $\mathcal{H}$. Therefore, $\mu(K) > 0$ for some polyhedral cone 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. $\square$

8.1 Proof of Theorem 1:

Lemma 2: (i) *The set of regular random utilities is nonempty.* (ii) *For any random utilities* μ, $\hat{\mu}$, *the integral* $\int \hat{\mu}(N(M(D,u),x))\mu(du)$ *is well-defined and satisfies*

$$\sum_{x \in D} \int \hat{\mu}(N(M(D,u),x))\mu(du) = \int \sum_{x \in D} \hat{\mu}(N(M(D,u),x))\mu(du)$$

Proof: (i) Let V be the usual notion of volume in $I\!R^n$. For any polyhedral cone K, let $\mu_V(\operatorname{int} K) = \frac{V(B_1(o)\cap K)}{V(B_1(o))}$. Obviously, $\dim K < n$ implies $\mu_V = 0$. 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 random utility. Since μ_V assigns probability 0 to all set of dimension less than n, by Lemma 1, μ_V is a regular random utility.

(ii) Let $f : I\!R \to$ be any simple function (i.e., the cardinality of $f(I\!R^n)$ is finite). Such a function f is $\mathcal{F}$-measurable if $f^{-1}(r) \in \mathcal{F}$ for all $r \in I\!R$. Countable additivity plays no role in the definition of the Lebesgue integral. Hence, $\int f\mu(du)$ exists whenever the simple function f is $\mathcal{F}$-measurable. That $\int(f+g)\mu(du) = \int f\mu(du) + \int g\mu(du)$ for all simple, $\mathcal{F}$-measurable functions f, g is obvious. Hence, to complete the proof, we need only verify that for all $x \in I\!R^n, D \in \mathcal{D}$, the function $f := \hat{\mu}(N(M(D,\cdot),x))$ is a simple, $\mathcal{F}$-measurable function.

Fix x, D and let $M := \{M(D,u) \,|\, u \in I\!R^n\}$. Clearly, M is nonempty. Since each element of M is a subset of the finite set D, the set M is also finite. Let $M_r = \{D' \in M \,|\, \hat{\mu}(N(D',x)) = r\}$. Note that the function f takes on values r such that $M_r \neq \emptyset$. Since nonempty M_r's form a partition of the finite set M, there are at most a finite set of r's for which $M_r \neq \emptyset$. Hence, f is a simple function. Note that $f(u) = r$ if and only if $M(D,u) \in M_r$. Hence $f^{-1}(r) = \bigcup_{D' \in M_r} N(D',x)$ and therefore f is measurable. □

Lemma 3: *Let μ be a random utility and $\hat{\mu}$ be a regular random utility. Define ρ by*

$$\rho^D(x) = \int \hat{\mu}(N(M(D,u),x))\mu(du)$$

Then, ρ is a random choice rule.

Proof: Obviously, $\rho^D(x) \geq 0$ for all D, x. Hence, we need only verify that $\sum_{x\in D} \rho^D(x) = 1$ for all $D \in \mathcal{D}$. Since $\hat{\mu}$ is regular, by Lemma 2, we have

$$\begin{aligned}\sum_{x\in D} \int \hat{\mu}(N(M(D,u),x))\mu(du) &= \int \sum_{x\in D} \hat{\mu}(N(M(D,u),x))\mu(du) \\ &= \int \sum_{x\in D} \hat{\mu}(N^+(M(D,u),x))\mu(du) = 1\end{aligned}$$

□

To prove Theorem 1, let μ be a regular random utility. Then ρ maximizes μ if and only if $\rho^D(x) = \mu(N(D,x))$ for all (D,x). This defines ρ uniquely and therefore μ has a unique maximizer.

To prove the converse, suppose μ is not full-dimensional. Then, by Lemma 1, μ is not full-dimensional. We will construct two distinct maximizers of μ. The first maximizer is the ρ defined by $\rho^D(x) = \int \mu_V(N(M(D,u),x))\mu(du)$ for the regular random utility μ_V constructed in the proof of Lemma 2. By Lemma 3 this ρ is a random choice rule.

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 1, 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 random utility. Then, let ρ_* be defined by $\rho_*^D(x) = \int \mu_*(N(M(D,u),x))\mu(du)$. By Lemma 3, ρ_* is a random choice rule. 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. $\square$

8.2 Proof of Theorem 2

We first show that ρ defined as

$$\rho^D(x) = \int \hat{\mu}(N(M(D,u),x))\mu(du)$$

is monotone, linear, extreme and mixture continuous.

Lemma 4: *ρ is monotone and linear.*

Proof: In the proof of Lemma 2(ii) we established that the function $f = \hat{\mu}(N(M(D,u),\cdot))$ is a simple, $\mathcal{F}$-measurable function. Obviously, if g is another simple, $\mathcal{F}$-measurable function such that $g(u) \leq f(u)$ for all $u \in I\!R^n$, then $\int g(u)\mu(du) \leq \int f(u)\mu(du)$. Therefore, to prove monotonicity, we need to show that $\hat{\mu}(N(M(D',u),x)) \leq \hat{\mu}(N(M(D,u),x))$ for all x, D, D' such that $x \in D$ and $D' = D \cup \{y\}$ for some y. For any such x, D, D', note that if $x \notin M(D', u)$ then we are done. If $x \in M(D', u)$ then $M(D, u) \subset M(D', u)$ and hence $N(M(D', u), x)) \subset N(M(D, u), x))$ and we are done.

To prove linearity, note that $\lambda M(D, u) + \{y\} = M(\lambda D + \{y\}, u)$ and $N(\lambda D' + \{y\}, \lambda x + \{y\}) = N(D', x)$. Hence, $N(M(D, u), x)) = N(\lambda M(D, u) + \{y\}, \lambda x + y) = N(M(\lambda D + \{y\}, u), \lambda x + y)$ as desired. □

Lemma 5: *ρ is extreme.*

Proof: *Claim 1:* Let $A = \text{conv}\, D$. Then, $\rho^D(x) = \int \hat{\mu}(N(F(A, u), x))\mu(du)$ for all $x \in D$.
Proof: Obviously, $\text{conv}\, M(D, u) = F(A, u)$. Then $N(F(A, u), x) = N(\text{conv}\, M(D, u), x) = N(M(D, u), x)$ for all $x \in D$. This proves claim 1.

Claim 2: $x, y \in \text{ri}\, F$ implies $[x \in F$ iff $y \in F$ for all $F \in F(A)]$.
Proof: It is enough to show that $x, y \in \text{ri}\, F$ and $x \in F'$ implies $y \in F'$ for all $F \in F(A)$. By Proposition 2(i), $x, y \in \text{ri}\, F$ implies $N(A, x) = N(A, y)$. Since every face F' of A is an exposed face, claim 2 follows.

Claim 3: $x, y \in \text{ri}\, F$ implies $\rho^D(x) = \rho^D(y)$.
Proof: Let $F_1 := F(A, u)$ and $F_2 := F(F_1, u')$. Then, $u' \in N(F(A, u), x)$ if and only if $x \in F_2$. Since F_2 is a face of F_1 which is a face of A, it follows that F_2 is face of A. Hence, Claim 2 yields $N(F(A, u), x) = N(F(A, u), y)$. By Claim 1, $\rho^D(x) = \rho^D(y)$ which proves claim 3.

To prove the Lemma, assume $x \in D \setminus \text{ext}\, D$. By Proposition 5($i$) $x \in \text{ri}\, F$ for some face F of $\text{conv}\, D$. Since x is not an extreme point of F it follows that $\text{ri}\, F$ is not a singleton. Therefore, there exists $y \in \text{ri}\, F \setminus D$. Let $D' = D \cup \{y\}$. By Claim 3, $\rho^{D'}(x) = \rho^{D'}(y)$. By Claim 1, $\rho^D(z) = \rho^{D'}(z)$ for all $z \in D$. Therefore, $1 = \sum_{z \in D} \rho^D(z) = \sum_{x \in D} \rho^{D'(z)} = \sum_{z \in D'} \rho^{D'(z)} - \rho^{D'}(y) = 1 - \rho^D(x)$. It follows that $\rho^D(x) = 0$, establishing extremeness.□

Lemma 6: *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. □

Lemma 7: *ρ is mixture continuous.*

Proof: Fix $D, D' \in \mathcal{D}$ and assume that $t_m \geq 0, t'_m \geq 0$ for all m, t_m converges to t and t'_m converges to t'.

Case 1: $t, t' > 0$**:** Let $z = tx + t'x'$ for some $x \in D, x' \in D'$. Choose an open ball O such that $tD + t'D' \cap O = \{z\}$. Choose m^* large enough so that for all $m \geq m^*$, $O \cap t_m D + t'_m D' = B_m$ where $B_m := \{t_m x + t'_m x \,|\, tx + t'x' = z\}$.

Claim 4: For any polytope A and $x \in \operatorname{ext} A$, $y \neq x$ implies $N(A,x) \neq N(A,y)$.
Proof: Note that $x \in \operatorname{ext} A$ implies $N^+(A,x) \neq \emptyset$ and obviously, $u \in N^+(A,x)$ implies $u \notin N(A,y)$ for any $y \neq x$. This proves claim 4.

By Proposition 3, $N(t_m D + t'_m D', z_m) = N(tD + t'D', z)$ for all $z_m \in B_m$. Then, it follows from Claim 4 that if B_m is not a singleton then no element of B_m is an extreme point of $t_m D + t'_m D', z_m$. Since we have already shown that ρ is extreme, $\rho^{t_m D + t'_m D'}(B_m) = 0$ whenever B_m is not a singleton. Recall that in addition to extremeness, monotonicity and linearity of ρ have also been established. Therefore we can apply Lemma 6 to establish $\rho^{t_m D + t'_m D'}(z) = \rho^{tD + t'D'}(z_m)$ for all $z_m \in B_m$ and we have $\rho^{t_m D + t'_m D'}(O) = \rho^{tD + t'D'}(O)$ for all $m \geq m^*$, establishing mixture continuity at (t, t') in case 1.

Case 2: $t' = 0$**:** It is easy to verify that $M(t_m D + t'_m D', u) = M(t_m D, u) + M(t'_m D', u)$. Proposition 3 implies $N(M(t_m D, u) + M(t'_m D', u), t_m x + t'_m x') = N(M(t_m D, u), t_m x) \cap N(M(t'_m D', u), t'_m x')$ for $x \in D, x' \in D'$. Since $\bigcup_{x' \in D'} N(M(t'_m D', u), t'_m x') = I\!R^n$, we have

$$\bigcup_{x' \in D'} N(M(t_m D + t'_m D', u), t_m x + t'_m x') = N(M(t_m D, u), t_m x) \text{ and therefore}$$

$$\hat{\mu}(\bigcup_{x' \in D'} N(M(t_m D + t'_m D', u), t_m x + t'_m x')) = \hat{\mu}(N(M(t_m D, u), t_m x))$$

Since

$$\hat{\mu}(\bigcup_{x'\in D'} N(M(t_m D + t'_m D', u), t_m x + t'_m x')) \leq \sum_{x'\in D'} \hat{\mu}(N(M(t_m D + t'_m D', u), t_m x + t'_m x'))$$

we conclude

$$\begin{aligned}
1 = \sum_{x\in D} \rho^{t_m D}(tx) &= \sum_{x\in D} \hat{\mu}(N(M(t_m D, u), t_m x)) \\
&= \sum_{x\in D} \hat{\mu}(\bigcup_{x'\in D'} N(M(t_m D + t'_m D', u), t_m x + t'_m x')) \\
&\leq \sum_{x\in D} \sum_{x'\in D'} \hat{\mu}(N(M(t_m D + t'_m D', u), t_m x + t'_m x')) \\
&= \sum_{x\in D} \sum_{x'\in D'} \rho^{t_m D + t'_m D'}(t_m x + t'_m x) = 1
\end{aligned}$$

The display equations above imply

$$\hat{\mu}(N(M(t_m D, u), t_m x)) = \sum_{x'\in D'} \hat{\mu}(N(M(t_m D + t'_m D', u), t_m x + t'_m x'))$$

Since $\hat{\mu}(N(M(tD, u), tx)) = \hat{\mu}(N(M(t_m D, u), t_m x))$ we obtain

$$\begin{aligned}
\rho^{tD}(x) = \hat{\mu}(N(M(tD, u), tx)) &= \sum_{x'\in D'} \hat{\mu}(N(M(t_m D + t'_m D', u), t_m x + t'_m x')) \\
&= \sum_{x'\in D'} \rho^{t_m D + t'_m D'}(t_m x + t'_m x')
\end{aligned}$$

Choose an open ball O such that $tD \cap O = \{z\}$. Choose m^* large enough so that for all $m \geq m^*$, $O \cap t_m D + t'_m D' = t_m\{x\} + t'_m D'$. Then, $\rho^{tD}(O) = \lim_{m\to\infty} \rho^{t_m D + t'_m D'}(O)$ follows from the last display equation above and proves mixture continuity of ρ. $\square$

Lemmas 4,5 and 7 establish that if ρ maximizes μ then ρ is mixture continuous, monotone, linear, and extreme.

For the converse, let ρ be a mixture continuous, monotone linear and extreme random choice rule. By Proposition 4, for any polyhedral cone K there exists (D, x) such that $K = N(D, x)$. We define $\mu : \mathcal{H} \to I\!R$ as follows:

$$\mu(\mathrm{ri}\, K) = \rho^D(x) \tag{3}$$

for D, x such that $K = N(D,x), K \in \mathcal{K}$. Lemma 6 ensures that μ is well-defined. Since ρ is extreme, $\rho^D(x) > 0$ implies $\operatorname{int} N(D,x) \neq \emptyset$ (Propositions 1(iii)). Hence, $\mu(\operatorname{ri} K) = 0$ for any polyhedral cone 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

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

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

Lemma 8: *If ρ is mixture continuous, monotone, linear and extreme then $\mu : \mathcal{H} \to I\!R$ 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 $i = 1, \dots, m$ with $\operatorname{ri} K_i, i = 1, \dots, n$ 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 $i = 0, \dots, m$. 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 $i \in \{0, \dots, m\}$ and let $\beta_k = (\beta_k^0, \dots, \beta_k^m)$ 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

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

Let $G(y) = G_{(1,\dots,1)}(y)$. 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, (5) 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. □

Next, we extend μ to $\mathcal{F}$. Equation (3) 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 $\bigcup_{j=1}^{k} H_i = F$, then $\sum_{i=1}^{m} \mu_i(H_i) = \sum_{i=1}^{m} \sum_{j=1}^{k} \mu(H_i \cap H_j') = \sum_{j=1}^{k} \mu_i(H_j')$.

Note that $\bigcup_{x \in E^*} \operatorname{int} N(E^*, x) \subset I\!R^n$. Hence, $\mu(\bigcup_{x \in E^*} \operatorname{int} N(E^*, x)) \geq \mu(I\!R^n)$. Since interiors of normal cones at distinct points are disjoint, we have $\mu(\bigcup_{x \in E^*} \operatorname{int} N(E^*, x)) = \sum_{x \in E^*} \mu(\operatorname{int} N(E^*, x)) = \rho^{E^*}(E^*) = 1$. Proving that μ is a finitely additive probability.

Next, we show that ρ maximizes μ. 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 of μ this holds for all D, x such that D has dimension n and hence $N(D, x) \in \mathcal{K}$. 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 $\bigcup_{y \in E^*} N(E^*, y) = I\!R^n$. 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

By Lemma 1, the only if part of the Theorem is equivalent to the following lemma:

Lemma 9: *If ρ maximizes the full-dimensional countably additive random utility 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 (9) is satisfied. Then, choose k so that for $i > k$ 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,

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

Since, ϵ is arbitrary, $\rho^D(G) \geq \limsup \rho^{D_i}(G)$ 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

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

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

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

Choose $\lambda \in (0,1)$ such that $\|x - y\| < \frac{1-\lambda}{\lambda}\|x' - y'\|$ for all $x, y \in D$ and $x', y' \in E^*$, $x' \neq y'$. 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

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

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

By Lemma 1, the if part of the Theorem is equivalent to the following lemma:

Lemma 10: *If the continuous random choice rule ρ maximizes the full-dimensional random utility μ 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 $\mathbb{R}^n \in \mathcal{H}$ and $\mu(\mathbb{R}^n) = 1$, 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), $K \in \mathcal{K}$ can be expressed as the disjoint union of sets $\operatorname{ri} A$ for $A \in F(K)$. Recall that each face of a polyhedron is a polyhedron. Note that $A \in F(K)$

and $A \neq K$ implies $A = H(u, \alpha) \cap K$ for some $u \neq o$. Hence, $A \neq K$ implies $\dim A < n$. By Corollary 1, ρ is extreme. Therefore, Propositions 1(iii), 5(i) and finite additivity 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 $\mu(\bigcup_{i=1}^{\infty} K_i) \leq \sum_{i=1}^{\infty} \mu(K_i)$

By Proposition 8 and the continuity of ρ for every $\epsilon > 0$ there are open sets O_i and cones $\tilde{K}_i$ such that (1) $\mu(K_i) \geq \mu(\tilde{K}_i) - 2^i \epsilon$ and (2) $K_i \cap S \subset O_i \subset \tilde{K}_i$ with $K_i = N(D_i, o), \tilde{K}_i = N(\tilde{D}_i, o)$. Therefore, $K_0 \cap S \subset \bigcup_{i=1}^{\infty} O_i$. 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. 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. $\square$

8.4 Proof of Corollary 2

Corollary 2 follows from Theorem 5 and Lemma 11 below.

Lemma 11: *A continuous random choice rule is undominated if and only if it is extreme.*

Proof: Note that $\operatorname{ext} D \subset \operatorname{bd} \operatorname{conv} D$. Hence, every extreme random choice rule 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 $1 < m \leq n$. To conclude the proof, we show that if $\rho^D(\operatorname{ext} D) = 1$ for all $D \in \mathcal{D}$ such that $\dim D = m$ then $\rho^D(\operatorname{ext} D) = 1$ for all $D \in \mathcal{D}$ such that $\dim D = m - 1$. Let $\dim D = m - 1$ and $x \in D \backslash \operatorname{ext} D$. Choose $y \in \operatorname{ext} D$ and $z \notin \operatorname{aff} D$. 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 $\operatorname{ext} D_k = (\operatorname{ext} D) \cup \{\frac{k-1}{k} y + \frac{1}{k} z\}$ for all k. Hence, there exists an open set O such that $x \in O$ and $O \cap \operatorname{ext} D_k = \emptyset$ for all k. By assumption, $\rho^{D_k}(O) = 0$ for all k. Then, by continuity $\rho^D(x) \leq \rho^D(O) = 0$. $\square$

9. Appendix

Proof of Proposition 1: (i) If $o \notin \operatorname{ext} N(A,x)$, then there exist $u \neq o$ such that $u, -u \in N(A,o)$. 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) Note that $N(A,x) = N(A - \{x\}, o) = N(\operatorname{pos}(A - \{x\}), o)$. Hence, $x \in \operatorname{ext} A$ iff $o \in \operatorname{ext}(A - \{x\})$ and $\operatorname{pos}(A - \{x\}) = N(L,o)$ for $L = N(\operatorname{pos}(A - \{x\}), o)$. Therefore, by part (i), $x \in \operatorname{ext} A$ 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. $\square$

Proof of Proposition 2: 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. $\square$

Proof of Proposition 3: Theorem 2.2.1(a) of Schneider (1993) proves the result for $m = 2$ which is equivalent to this proposition. $\square$

Proof of Proposition 4: 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. $\square$

Proof of Proposition 5: (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 $F(A)$.

That $\mathbb{R}^n \subset \bigcup_{K \in \mathcal{N}(A)} K$ is obvious. By part (i), $F^0(K)$ is a partition of K for each $K \in \mathcal{N}(A)$. Hence, $\mathbb{R}^n \subset \bigcup_{K \in \mathcal{N}^0(A)} K$. To complete the proof we need only that

$K, K' \in \mathcal{N}^0$ and $K \cap K' \neq \emptyset$ implies $K = K'$. Suppose, $\operatorname{ri} N(A,x) \cap \operatorname{ri} N(A,y) \neq \emptyset$. Then, for $u \in \operatorname{ri} N(A,x) \cap \operatorname{ri} N(A,y)$ Proposition 2(ii) yields $x, y \in \operatorname{ri} F(A,u)$. But then Proposition 2(i) establishes $N(A,x) = N(A,y)$ and therefore $\operatorname{ri} N(A,x) = \operatorname{ri} N(A,y)$. Hence, $\mathcal{N}^0$ is a partition. □

Proof of Proposition 6: (i) First, we show that $H, H' \in \mathcal{H}$ implies $H \cap H'$ in $\mathcal{H}$. Let $H = \operatorname{ri} K$ and $H' = \operatorname{ri} K'$ for $K, K' \in \mathcal{K}$ such that $o \in \operatorname{ext} K \cap \operatorname{ext} K'$ and hence $o \in \operatorname{ext}(K \cap K')$. If $H \cap H' = \emptyset$, we are done. Otherwise, by Proposition 1(v), $H \cap H' = \operatorname{ri}(K \cap K') \in \mathcal{H}$ 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 $\mathcal{F} \subset \{\cup_{i=1}^m H_i \mid H_i \in \mathcal{H} \text{ for } i = 1, \ldots, m\}$. Clearly, the set of all finite unions of elements of a semiring is a field. Hence, $\{\cup_{i=1}^m H_i \mid H_i \in \mathcal{H} \text{ for } i = 1, \ldots, m\}$ 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 $\mathcal{H}$ contains $\mathcal{K}$. Let K be a polyhedral cone. Then, by Proposition 4, there exists A, x such that $N(A,x) = K$. Since $\bigcup_{\operatorname{ext} B} N(B,x) = I\!R^n$, 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 $H = \text{ri}\, K$ for some $K \in \mathcal{K}$. Since $\mathcal{F}$ contains all polyhedral cones, $K \in \mathcal{F}$. By Proposition 5(i), $F^0(K)$ is a partition of K that measures each face of K. Hence,

$$K = \text{ri}\, K \cup (\bigcup_{F\in F(K), F\neq K} F)$$
$$\emptyset = \text{ri}\, K \cap (\bigcup_{F\in F(K), F\neq K} F)$$

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

Proof of Proposition 7: Since $K \in \mathcal{K}$ Proposition 1(i) implies $\dim \text{conv}\, D = n$. Let $y^* \in \text{int conv}\, D$ and let $\tilde{D}_j = \{x\} \cup (\frac{j}{j+1} D + \frac{1}{j+1}\{y^*\})$. Note that $y^* \in \text{int conv}\, \tilde{D}_j$. 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 $y \in D$. Since $u \cdot x \geq u \cdot y^*$ 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 $y^* + \epsilon' z \in \text{int conv}\, \tilde{D}_j$ 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 $u \cdot (\frac{j}{j+1} y + \frac{1}{j+1} y^*) \geq u \cdot x$, therefore $u\cdot y > u \cdot x$.

Let $R_j(y) := N(D_j, \frac{j}{j+1} y + \frac{1}{j+1} y^*) \cap S, y \in D$ and note that $R_j(y)$ is compact. By the argument above, $u\cdot y > u\cdot x$. 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 $u \in R_j(y), x_i \in B_{\epsilon_j}(x) \cap D_i, y_i \in B_{\epsilon_j}(y) \cap D_i$ 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 $x_i \in B_{\epsilon_j}(x) \cap D_i$. We conclude that $\bigcup_{y \in D_i \cap B_{\epsilon_j}(x)} N(D_i, y) \subset K_j$ for all $i > k_j$. $\square$

Proof of Proposition 8: By Proposition 4, there is $D \in \mathcal{D}$ such that $o \in D$ and $K = N(D, o)$. By Proposition 1(i), $\dim D = n$. Choose $y \in \operatorname{int} \operatorname{conv} D$ and let $D' = \{o\} \cup ((1 - \lambda) D + \lambda \{y\})$. Choose $\lambda > 0$ so that $d_h(D, D') < \epsilon$. 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 $0 > -\epsilon' + \epsilon \geq u \cdot x + (u' - u) \cdot x = u' \cdot x\ \forall x \in D' \backslash \{o\}, \forall u' \in \bigcup_{u \in K \cap S} B_\epsilon(u)$. Let $O := \bigcup_{u \in K \cap S} B_\epsilon(u)$. Clearly, $K \cap S \subset O \subset K'$ as desired. $\square$

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

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

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

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

7. Kalai, G, A. Rubinstein and R. Spiegler, "Comments on Rationalizing Choice Functions which Violate Rationality", mimeo October 2001.

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

9. McFadden, D. and M. Richter, "Revealed Stochastic Preference", mimeo, Department of Economics, MIT, 1970.

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

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

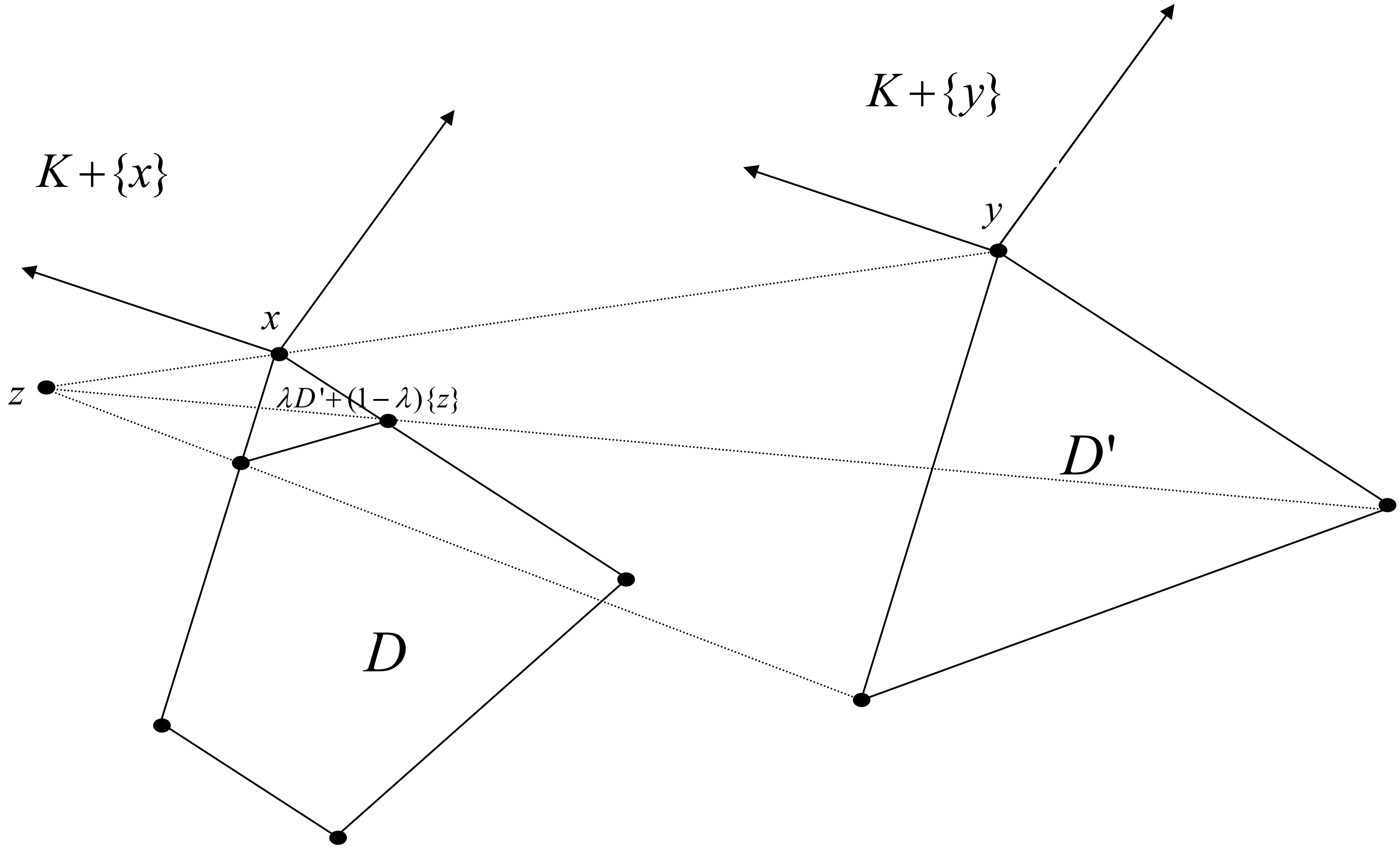
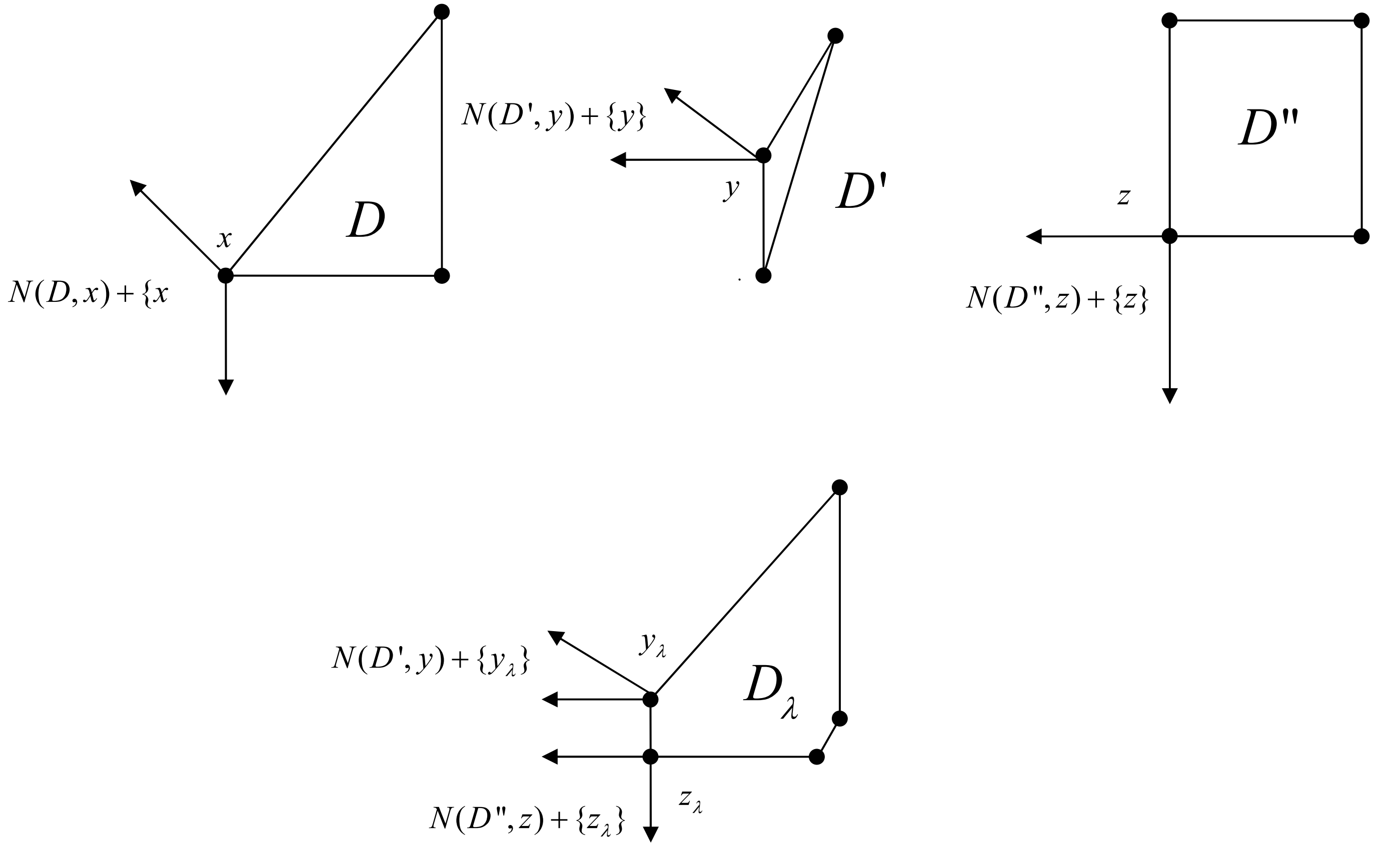


Figure 1



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

Figure 2