Essays on Random Choice

The Harvard community has made this article openly available. Please share how this access benefits you. Your story matters.

<table>
<thead>
<tr>
<th>Citation</th>
<th>McClellon, Morgan. 2015. Essays on Random Choice. Doctoral dissertation, Harvard University, Graduate School of Arts &amp; Sciences.</th>
</tr>
</thead>
<tbody>
<tr>
<td>Citable link</td>
<td><a href="http://nrs.harvard.edu/urn-3:HUL.InstRepos:17465146">http://nrs.harvard.edu/urn-3:HUL.InstRepos:17465146</a></td>
</tr>
<tr>
<td>Terms of Use</td>
<td>This article was downloaded from Harvard University’s DASH repository, and is made available under the terms and conditions applicable to Other Posted Material, as set forth at <a href="http://nrs.harvard.edu/urn-3:HUL.InstRepos:dash.current.terms-of-use#LAA">http://nrs.harvard.edu/urn-3:HUL.InstRepos:dash.current.terms-of-use#LAA</a></td>
</tr>
</tbody>
</table>
Essays on Random Choice

A dissertation presented
by
Morgan McClellon
to
The Department of Economics
in partial fulfillment of the requirements
for the degree of
Doctor of Philosophy
in the subject of
Economics

Harvard University
Cambridge, Massachusetts
May 2015
Abstract

Chapter 1 introduces and axiomatizes a new class of representations for incomplete preferences called confidence models, which describe decision makers who behave as if they have probabilistic uncertainty over their true preferences, and are only willing to express a binary preference if it is sufficiently likely to hold. Confidence models provide a natural way to connect incomplete preferences with stochastic choice; this connection is characterized by a simple condition that serves to identify the behavioral content of incomplete preferences. Chapter 2 studies random choice rules over finite sets that obey regularity but potentially fail to satisfy all of the Block-Marschak inequalities. Such random choice rules can be represented by capacities on the space of preferences. The higher-order Block-Marschak inequalities are shown to be related to the degree of monotonicity that can be achieved by a capacity representation. Finally, Chapter 3 shows that failures of uniqueness for random utility representations are widespread. Uniqueness can be restored by introducing a finite state space and considering random choice over Savage acts. A representation is characterized in which acts are chosen according to the probability that they are optimal in every state.
## Contents

Abstract ................................................................. iii  
Acknowledgments ..................................................... vi  

1 Confidence Models of Incomplete Preferences ............ 1  
  1.1 Introduction ...................................................... 1  
  1.2 Confidence Models on Finite Sets of Alternatives ....... 3  
     1.2.1 Basic Definitions ........................................... 3  
     1.2.2 Axiomatization ............................................... 4  
     1.2.3 Regular Confidence Preferences and Stochastic Choice ... 5  
     1.2.4 Confidence Models and Random Choice with Indifference ... 9  
  1.3 Confidence Models on Sets of Lotteries ............ 11  
     1.3.1 Setup and Axiomatization .................................. 11  
     1.3.2 Stochastic Choice Over Lotteries ....................... 13  
  1.4 Uniqueness ....................................................... 15  
  1.5 Related Literature .............................................. 17  
     1.5.1 Multi-Utility Models and Transitivity ................. 17  
     1.5.2 Stochastic Preference ...................................... 19  
     1.5.3 Incomplete Preferences on Anscombe-Aumann Acts .... 19  

2 Non-Additive Random Utility Functions ............ 21  
  2.1 Introduction ...................................................... 21  
  2.2 Background and Related Literature ...................... 22  
  2.3 Capacity Representations ...................................... 24  
  2.4 Unraveling the BM Inequalities ......................... 28  
     2.4.1 Supermodularity ........................................... 28  
     2.4.2 K-Monotonicity ........................................... 30  
  2.5 Connection to Random Choice over Lotteries .......... 32
Acknowledgments

I would like to thank Mira Frick, Drew Fudenberg, and especially Tomasz Strzalecki for excellent advice and support throughout the past five years.
To Susie McClellon
Chapter 1

Confidence Models of Incomplete Preferences\(^1\)

1.1 Introduction

This paper introduces and axiomatizes a new class of representations for incomplete preferences called confidence models.\(^2\) A decision maker (DM) with preferences represented by a confidence model has probabilistic uncertainty over her underlying preferences. When comparing two alternatives, the DM computes the probability that one is weakly preferred to the other, and expresses a weak preference if and only if this probability exceeds a fixed threshold. Intuitively, the DM only expresses binary preferences in which she holds sufficient confidence.

Confidence models can be formulated for incomplete preferences on a variety of

\(^1\)This chapter benefited from helpful conversations with Mira Frick, Drew Fudenberg, and Tomasz Strzalecki, as well as feedback from two anonymous referees. All remaining errors are my own. This chapter was partially funded by a Graduate Research Fellowship from the National Science Foundation.

\(^2\)It is worth mentioning that the term “confidence model” has been used in a different decision-theoretic context by Chateauneuf and Faro (2009). That paper studies complete preferences over Anscombe-Aumann acts and characterizes a representation that features a fuzzy set of priors.
domains. Here, we define and axiomatize confidence models for preferences over a finite set and for preferences over lotteries. We show that confidence models are remarkably general and nest the prominent existing models of incomplete preferences on both these domains. Potential extensions to additional domains are discussed in the conclusion.

The generality of confidence models allows them to encompass behavior outside the scope of existing models for incomplete preferences. In particular, preferences represented by confidence models need not be transitive. When interpreting incomplete preferences in the context of political science, failures of transitivity arise naturally due to Condorcet cycles. Confidence preferences can accommodate such phenomena and can be interpreted as a model of supermajority aggregation rules. The potential lack of transitivity distinguishes confidence models from the popular multi-utility models of incomplete preferences axiomatized by Evren and Ok (2008) and Dubra et al. (2004).

Confidence models suggest a natural way to connect incomplete preferences and stochastic choice. Specifically, an incomplete preference and a random choice rule are defined to be consistent if they can be represented using a common probability measure on the space of complete preference relations. We show that this notion of consistency has a simple characterization in terms of a joint condition on the incomplete preferences and the stochastic choice data. The tight connection between confidence preferences and random choice provides a novel way to understand the behavioral content of incomplete preferences.

In a recent experiment, Agranov and Ortoleva (2015) found that subjects chose stochastically even when the same menu was presented to them multiple times in a row. Afterwards, subjects were asked to explain their choices, and many indicated that they chose stochastically as a result of being unable to select an optimal alternative. To be sure, this is anecdotal evidence, but it certainly suggests that stochastic choice may arise at the individual level as a result of incomplete preferences.
Section 1.2 studies confidence preferences on a finite set of alternatives, while Section 1.3 studies preferences on lotteries over a finite set of alternatives. Limitations on the uniqueness of confidence representations are discussed in Section 1.4. Section 1.5 discusses related literature, and in particular formalizes the relationship between confidence models and multi-utility models.

1.2 Confidence Models on Finite Sets of Alternatives

1.2.1 Basic Definitions

Let $X$ be a set of alternatives with $|X| = N$ and $2 \leq N < \infty$. Let $W$ denote the set of all complete and transitive relations on $X$. A binary relation $R$ is is antisymmetric if $xRy$ and $yRx$ implies $x = y$; let $\mathcal{R} \subseteq W$ denote the set of all complete, transitive, antisymmetric relations on $X$. A traditional agent with complete preferences is described by a preference relation $W \in W$; if the agent is allowed to express only strict preferences between distinct objects, she is described by a preference relation $R \in \mathcal{R}$.

We are interested in a DM with a potentially-incomplete preference relation $\succsim$ on $X$, who behaves as if she is unsure of her “true” preferences and is willing to express only binary preferences in which she is sufficiently confident. Formally, such preferences are represented as follows:

**Definition 1.** A confidence model for a preference $\succsim$ is a pair $(\mu, \alpha)$ where $\mu \in \Delta(W)$ and $\alpha \in (1/2, 1]$ such that for any $x, y \in X$ we have

$$x \succsim y \text{ if and only if } \mu(\{W \mid xWy\}) \geq \alpha.$$ 

A regular confidence model for $\succsim$ is a confidence model $(\mu, \alpha)$ where supp$(\mu) \subseteq \mathcal{R}$.
In a confidence model, the measure $\mu$ captures the DM’s uncertainty over her preferences, and the parameter $\alpha$ captures the degree of confidence the DM must have before expressing a given binary preference. Specifically, when comparing two alternatives $x$ and $y$, the DM behaves as if she is computing the $\mu$-probability that $x$ is preferred to $y$ and expressing the preference $x \succeq y$ if and only if this probability exceeds the threshold $\alpha$.

A DM with preferences captured by a regular confidence model behaves as if she is certain that her preferences will not contain any nontrivial indifference. The distinction between confidence models and regular confidence models will be important in Section 1.2.3 when we connect confidence preferences to random choice.

1.2.2 Axiomatization

The following proposition shows that confidence models are extremely general: they can represent any reflexive preference. The subclass of regular confidence models is characterized by antisymmetry.

Corollary 1. Let $\succsim$ be a binary relation on $X$.

1. $\succsim$ is reflexive if and only if it can be represented by a confidence model.

2. $\succsim$ is reflexive and antisymmetric if and only if it can be represented by a regular confidence model.

This result is a corollary of McGarvey’s Theorem, a classic result in social choice theory established by McGarvey (1953).\(^3\) McGarvey’s Theorem states that any complete relation can arise as the outcome of majority rule applied a collection of complete, transitive and antisymmetric preferences.

\(^3\)I thank an anonymous referee for pointing out this connection.
Corollary 1 can be established directly from McGarvey’s Theorem with straightforward arguments. The appendix provides a novel proof for Corollary 1 that is not based on McGarvey’s Theorem, but instead relies on the connection between confidence preferences and stochastic choice that is formalized in Lemma 1. An outline of this proof may be found after the statement of Lemma 1 in Section 1.2.3. The (non)uniqueness of confidence representations is discussed in Section 1.4.

The following example illustrates how confidence preferences may fail to be transitive when the measure \( \mu \) supports a Condorcet cycle.

**Example 1.** Let \( X = \{x, y, z\} \). Consider the following relations from \( \mathcal{R} \):

\[
R_1: \quad xR_1yR_1z;
R_2: \quad zR_2xR_2y;
R_3: \quad yR_3zR_3x.
\]

Let \( \mu \) assign probability 1/3 to each relation \( R_1 \), \( R_2 \), and \( R_3 \). Let \( \alpha = .6 \). If \( \succeq \) is represented by \((\mu, \alpha)\) then it is straightforward to check that \( x \succeq y \) and \( y \succeq z \), but \( \neg (x \succeq z) \), in violation of transitivity. \( \Diamond \)

Consistent with the connection to McGarvey’s theorem, confidence preferences can be given a social-choice interpretation. In this view, \( \succeq \) represents the preferences of “Society,” which are formed by aggregating the (complete and transitive) preferences of individuals. If \( \succeq \) has a confidence representation, we can interpret \( \mu \) as the distribution of preferences in the population, and \( \alpha \) as the threshold for a supermajority voting rule.

### 1.2.3 Regular Confidence Preferences and Stochastic Choice

Confidence models provide a natural way to connect incomplete preferences with random choice data. Before formalizing this connection, we briefly review the essential
concepts from the literature on stochastic choice.

Let \( X \) again be a finite set of alternatives with \(|X| = N\), and let \( \mathcal{D} \) denote the set of all nonempty subsets of \( X \). A random choice rule (RCR) is a function \( \rho: \mathcal{D} \times X \to [0, 1] \), where \( \rho^D(x) := \rho(D, x) \) is interpreted as the observed probability that \( x \) is chosen from \( D \). For this to make sense, we impose that for all \( D \in \mathcal{D} \)

\[
\sum_{x \in D} \rho^D(x) = 1.
\]

In the context of random choice, a probability measure \( \mu \in \Delta(\mathcal{R}) \) is called a random utility function (RUF). A RUF is said to represent the RCR \( \rho \) if

\[
\rho^D(x) = \mu(\{R \in \mathcal{R} \mid x \text{ is } R\text{-optimal in } D\})
\]

for all \((D, x) \in \mathcal{D} \times X\). A classic result established by Falmagne (1978) and rediscovered by Barberá and Pattanaik (1986) shows that a RCR \( \rho \) can be represented by some RUF if and only if the following inequalities hold:

\[
\sum_{C: D \subseteq C} (-1)^{|C \setminus D|} \rho^C(x) \geq 0 \quad \forall (D, x) \in \mathcal{D} \times X.
\]

These inequalities are known as the Block-Marschak (BM) inequalities;\(^4\) they are written here in the form of Fiorini (2004).

In the language of stochastic choice, regular confidence models represent a reflexive, antisymmetric preference relation using two components: a RUF and a threshold. This suggests a natural connection between such preferences and stochastic choice.

**Definition 2.** A reflexive, antisymmetric preference \( \succeq \) is consistent with a RCR \( \rho \) satisfying the BM inequalities if there exists a RUF \( \mu \in \Delta(\mathcal{R}) \) such that: (1) \((\mu, \alpha)\) represents \( \succeq \) for some \( \alpha \in (1/2, 1]\); and (2) \( \mu \) represents \( \rho \).

\(^4\)In honor of Block and Marschak (1960), which established that they are necessary for a RCR to be represented by a RUF.
In words, the preference $\succeq$ is consistent with stochastic choice data if we can find representations for each that share a common RUF. Although we have defined consistency in terms of the representations for $\succeq$ and $\rho$, there is a natural condition that allows an observer to check directly whether an incomplete preference is consistent with a given set of stochastic choice data.

**Definition 3** (Binary Compatibility). A preference $\succeq$ and a RCR $\rho$ are compatible on binary sets if for all $x \succeq y$ and $\neg(x' \succeq y')$, we have $\rho^{(x,y)}(x) > \rho^{(x',y')}(x')$.

Binary compatibility requires that if $x$ is weakly preferred to $y$ while $x'$ is not weakly preferred to $y'$ (either because $y' \succ x'$ or because $x'$ and $y'$ are not comparable) then $x$ must be chosen with a higher frequency from the binary menu $\{x, y\}$ than $x'$ is chosen from the menu $\{x', y\}$.

The following lemma shows that binary compatibility characterizes consistency between preferences and stochastic choice.

**Lemma 1.** Let $\succeq$ be a reflexive, antisymmetric preference and let $\rho$ be a RCR satisfying the BM inequalities. Then $\succeq$ and $\rho$ are consistent if and only if they are compatible on binary sets.

**Proof.** Assume that $\succeq$ and $\rho$ are compatible on binary sets. Let $\mu$ be a RUF representing $\rho$. Let $S = \{(x, y) \mid x \succeq y\}$ and define

$$\alpha := \min_S \rho^{(x,y)}(x).$$

Because $\succeq$ is reflexive, $S$ is nonempty and $\alpha \leq 1$. Together, the antisymmetry of $\succeq$ and compatibility of $\succeq$ and $\rho$ on binary sets further imply that $\alpha > 1/2$. Recalling that $\mu$ represents $\rho$, binary compatibility and the definition of $\alpha$ imply that $\mu\{R \in \mathcal{R} \mid x R y\} \geq \alpha$ if and only if $x \succeq y$, showing that $(\mu, \alpha)$ is regular confidence model for
\(\succeq\). This representation features the same RUF used to represent \(\rho\), which establishes consistency between \(\succeq\) and \(\rho\).

Conversely, assume \(\succeq\) and \(\rho\) are consistent—that is, there exists a \(\mu \in \Delta(\mathcal{R})\) such that \(\mu\) represents \(\rho\) and \((\mu, \alpha)\) represents \(\succeq\) for some \(\alpha \in (1/2, 1]\). Let \(x \succeq y\) and \(\neg(x' \succeq y')\). Then

\[
\rho^{\{x,y\}}(x) = \mu(\{R \in \mathcal{R} | xRy\}) \geq \alpha \geq \mu(\{R \in \mathcal{R} | x'Ry'\}) = \rho^{\{x',y'\}}(x'),
\]

showing that \(\succeq\) and \(\rho\) are compatible on binary sets. In the preceding equation, the equalities hold because \(\mu\) represents \(\rho\) and the inequalities hold because \((\mu, \alpha)\) represents \(\succeq\).

Lemma 1 provides a tractable characterization of consistency between reflexive, antisymmetric preferences and stochastic choice data. In addition, it provides a useful technical tool that plays a key role in the proof of Theorem 1. Indeed, the proof of Theorem 1(B) takes a reflexive, antisymmetric preference \(\succeq\) and constructs a hypothetical RCR that satisfies the BM inequalities. This RCR is constructed to be compatible with \(\succeq\) on binary sets, allowing Lemma 1 to deliver a confidence model for \(\succeq\).

The proof of Theorem 1(A) also uses Lemma 1 to construct a measure \(\mu \in \Delta(\mathcal{R})\). However, the hypothetical RCR used to generate this measure is not necessarily compatible with \(\succeq\) on binary sets, and the resulting RUF must be modified before it can be used in a confidence representation of \(\succeq\). Specifically, \(\mu\) is mixed with several different measures on \(\mathcal{W}\) to form a new measure \(\mu' \in \Delta(\mathcal{W})\). This measure is then combined with an appropriate threshold to generate a confidence model for \(\succeq\). Details may be found in the appendix.
1.2.4 Confidence Models and Random Choice with Indifference

Antisymmetry plays a central role in the connection between regular confidence models of incomplete preferences and random choice. Formalizing a similar connection without antisymmetry requires a model of random choice with indifference, which has recently been provided by Gul and Pesendorfer (2013). These authors define a (generalized) RCR to be a function $\rho: \hat{D} \rightarrow [0, 1]$ where

$$\hat{D} := \{(A, B) \in D \times D \mid A \subseteq B\}$$

and require that $\rho(B, B) = 1$ for all $B \in D$. The value $\rho(A, B)$ is interpreted as the minimal probability that an element from $A$ is selected from the choice set $B$. A measure $\mu \in \Delta(W)$ is said to represent $\rho$ if $\rho(A, B)$ is equal to the $\mu$-probability that $A$ contains all the $W$-maximal elements in $B$.

Gul and Pesendorfer (2013) provide conditions on $\rho$ (hereafter referred to as the GP-inequalities) that are necessary and sufficient for $\rho$ to be represented by a measure $\mu \in \Delta(W)$. Using this result, we may define consistency between a reflexive preference and a generalized RCR as follows:

**Definition 4.** A reflexive preference $\succeq$ is consistent with a generalized RCR $\rho$ satisfying the GP-inequalities if there exists a measure $\mu \in \Delta(W)$ such that: (1) $(\mu, \alpha)$ represents $\succeq$ for some $\alpha \in (1/2, 1]$; and (2) $\mu$ represents $\rho$.

Characterizing consistency is somewhat more delicate in the presence of non-trivial indifference. For instance, consider a DM who is completely indifferent among all alternatives. In this case, any generalized RCR would vacuously satisfy binary compatibility with the DM’s preference relation. But clearly many generalized RCRs are not consistent with a completely indifferent preference in the sense of Definition 4.
The following definition strengthens binary compatibility to accommodate generalized RCRs. It is equivalent to binary compatibility for traditional, additive RCRs.

**Definition 5 (Generalized Binary Compatibility).** A preference \( \succeq \) and a generalized RCR \( \rho \) satisfy *generalized binary compatibility* if for all \( x \succeq y \) and \( \neg(x' \succeq y') \), we have

\[
\rho(\{y\}, \{x, y\}) < \rho(\{y'\}, \{x', y'\})
\]

and in addition \( \rho(\{y\}, \{x, y\}) < 1/2 \) for all \( x \succeq y \).

Lemma 2 shows that generalized binary compatibility characterizes consistency between a reflexive preference and a generalized RCR. The proof is a slight modification of the proof of Lemma 1 and may be found in the appendix.

**Lemma 2.** Let \( \succeq \) be reflexive and let \( \rho \) be a generalized RCR satisfying the GP-inequalities. Then \( \succeq \) and \( \rho \) are consistent if and only if generalized binary compatibility holds.

The definition of generalized binary compatibility explicitly requires that certain choice probabilities be less than 1/2. This corresponds to the requirement that the threshold \( \alpha \) in a confidence representation be strictly larger than 1/2. Restricting the allowable thresholds to \((1/2, 1]\) is natural in the case of antisymmetric preferences; indeed, it guarantees that antisymmetry is a necessary condition for representation by a regular confidence model. Nonetheless, it is certainly feasible to imagine relaxing this requirement when defining confidence models. If arbitrary thresholds were allowed, preferences with non-trivial indifference would become consistent with a wider variety of generalized RCRs. The definition of generalized binary compatibility could also be loosened to merely require that \( \rho(\{y\}, \{x, y\}) < \rho(\{y'\}, \{x', y'\}) \) for all \( x \succeq y \) and \( \neg(x' \succeq y') \). Axiomatically, nothing would be gained: any preference represented by
$(\mu, \alpha')$ for $\alpha' < 1/2$ is still reflexive, and thus has a confidence representation satisfying the original definition.

If $\alpha \leq 1/2$, regular confidence models could support non-trivial indifference. However, any regular confidence model with non-trivial indifference would necessarily be a complete preference. So allowing $\alpha$ to take arbitrary values does not eliminate the need for non-regular RUFs.

### 1.3 Confidence Models on Sets of Lotteries

#### 1.3.1 Setup and Axiomatization

The intuition underlying confidence models—a DM is unsure of her preferences, and expresses only those she believes will hold with sufficiently high probability—does not depend on the structure of the domain $X$. In principle, then, confidence models can be defined on a variety of domains. In this section we study confidence models for incomplete preferences over lotteries. Several technical issues arise when characterizing confidence models, but the strong connection to stochastic choice remains.

Let $X$ again be a finite set with $|X| = N$. The primitive for this section is a preference relation $\succeq$ defined on a set of $\mathcal{L} \subseteq \Delta(X)$. Endow the domain $\mathcal{L}$ with the standard topology it inherits as a subset of $\mathbb{R}^N$.

The set of twice normalized expected utility functions on $\Delta(X)$ is

$$\mathcal{U} := \left\{ u \in \mathbb{R}^n | \sum_i u_i = 0, \sum_i u_i^2 = 1 \right\}.$$ 

Endow $\mathcal{U}$ with the standard Euclidean topology, and let $\Delta(\mathcal{U})$ be the set of finitely-additive measures on the Borel $\sigma$-algebra of $\mathcal{U}$.

Confidence models for incomplete preferences over lotteries are defined as follows.
Definition 6. A confidence model for a preference $\succsim$ over $\mathcal{L}$ is a pair $(\mu, \alpha)$ where $\mu \in \Delta(\mathcal{U})$ and $\alpha \in (1/2, 1]$, such that for all $p, q \in \mathcal{L}$

$$p \succsim q \iff \mu(\{u \in \mathcal{U} \mid u \cdot p \geq u \cdot q\}) \geq \alpha.$$  

Confidence models describe DMs who are unsure of their underlying expected utility preferences, and express a preference for $p$ over $q$ if and only if their subjective measure $\mu$ assigns a sufficiently high probability to the expected utility of $p$ exceeding the expected utility of $q$.

In Section 1.2, we defined regular confidence models, which featured measures that put zero probability on the DM being indifferent between two alternatives. Regular confidence models for preferences over lotteries are defined analogously: they use measures on $\mathcal{U}$ for which ties are a zero-probability event.

Definition 7. A measure $\mu \in \Delta(\mathcal{U})$ is regular if any finite collection of lotteries has a unique maximal element with $\mu$-probability 1.5 A confidence model $(\mu, \alpha)$ featuring a regular $\mu$ is a called a regular confidence model.

The main theorem of this section characterizes confidence models when $\mathcal{L}$ is finite.

Theorem 1. Let $\mathcal{L} \subseteq \Delta(\mathcal{X})$ be finite. A preference $\succsim$ on $\mathcal{L}$ can be represented by a confidence model if and only if it is reflexive and satisfies Independence, namely: for any $p, q \in \mathcal{L}$, $r \in \Delta(\mathcal{X})$ and $\lambda \in (0, 1)$,

$$p \succsim q \Rightarrow \lambda p + (1 - \lambda) r \succsim \lambda q + (1 - \lambda) r$$

whenever both mixtures in the implication are in $\mathcal{L}$.

Compared to Theorem 1, this result adds only the standard Independence axiom from expected utility theory to characterize confidence models on finite sets of lotteries.

\footnote{For a complete formal definition, see Gul and Pesendorfer (2006).}
Clearly this axiom is necessary: every \( u \in \mathcal{U} \) represents a complete preference over lotteries that satisfies Independence. Because the domain \( \mathcal{L} \) is finite, there is no need for a continuity axiom.

The case of a finite \( \mathcal{L} \) holds some interest, as experiments necessarily can elicit only a finite number of preferences. From a theoretical viewpoint, the other natural case to consider is \( \mathcal{L} = \Delta(X) \), the standard domain for axiomatizing preferences over lotteries. I conjecture, but have been unable to prove, that in this case regular confidence models are characterized by reflexivity, antisymmetry, and Independence. Surprisingly, the standard continuity axiom does not necessarily hold unless the measure \( \mu \) is countably additive.\(^6\) In addition, merely dropping antisymmetry will no longer suffice to distinguish confidence models from regular confidence models. See Appendix A.2 for details on these technical issues.

### 1.3.2 Stochastic Choice Over Lotteries

Section 1.2.3 showed that regular confidence models on finite sets of alternatives have a natural connection to stochastic choice. A similar connection exists for regular confidence models on lotteries. We begin by redefining notation to match the new domain. Throughout this section, we consider preferences on the domain \( \mathcal{L} = \Delta(X) \).

Let \( \mathcal{D} \) denote the set of all nonempty, finite subsets of \( \Delta(X) \). Let \( \Delta(\Delta(X)) \) denote the set of all simple probability measures on \( \Delta(X) \). A random choice rule (RCR) is function \( \rho: \mathcal{D} \to \Delta(\Delta(X)) \), where we let \( \rho^D(p) \) denote the probability that \( p \) is chosen from \( D \), and require that \( \rho^D(D) := \sum_{p \in D} \rho^D(p) = 1 \) for all \( D \).

Gul and Pesendorfer (2006) introduce the following (paraphrased) axioms for a RCR:

\(^6\)I thank an anonymous referee for calling attention to this issue and correcting an error in a previous version of this paper.
GP1 Mixture continuity: $\rho^{\lambda D + (1-\lambda)D'}$ is continuous in $\lambda$.

GP2 Monotonicity: if $p \in D \subset D'$, then $\rho^{D'}(p) \leq \rho^D(p)$.

GP3 Linearity: $\rho^{\lambda D + (1-\lambda)q} (\lambda p + (1-\lambda)\{q\}) = \rho^D(p)$.

GP4 Extremeness: $\rho^D(\text{ext}(D)) = 1$.

In the context of random choice over lotteries, a measure $\mu \in \Delta(\mathcal{U})$ is called a random utility function (RUF). An RCR $\rho$ is maximized by a RUF $\mu$ if $\rho^D(p)$ is equal to the $\mu$-probability that $p$ is optimal in $D$. Gul and Pesendorfer (2006) prove that an RCR satisfies GP1–4 if and only if it is maximized by a regular measure $\nu$, which is defined on a set of expected utility functions that are normalized differently than the expected utilities in $\mathcal{U}$. These differences are addressed by Ahn and Sarver (2013), who prove that an RCR $\rho$ satisfies Axioms GP1–4 if and only if it is maximized by a unique regular $\mu \in \Delta(\mathcal{U})$.

We are now ready to introduce the analogue of Definition 2 for preferences over lotteries. The term ”regular confidence preference” refers a preference that has a regular confidence representation.

**Definition 8.** A regular confidence preference $\succeq$ on $\mathcal{L} = \Delta(X)$ is consistent with a RCR $\rho$ satisfying GP1–4 if there exists an $\alpha \in (1/2, 1]$ such that the regular confidence model $(\mu, \alpha)$ represents $\succeq$, where $\mu$ is the unique, regular RUF that maximizes $\rho$.

The following lemma shows that consistency is once again characterized by binary compatibility.

**Definition 9 (Binary Compatibility).** A RCR $\rho$ is compatible on binary sets with a preference $\succeq$ if for all pairs of lotteries $(p, p')$ and $(q, q')$ with $p \succeq p'$ and $q \succeq q'$, we have $\rho(p, p')(p) > \rho(q, q')(q)$.
Lemma 3. A RCR $\rho$ satisfying GP1–4 is compatible on binary sets with a regular confidence preference $\succsim$ that satisfies the axioms of Theorem 1 if and only if $\rho$ and $\succsim$ are consistent.

The proof (similar in spirit to that of Lemma 1) may be found in the appendix. If an analyst is confronted with a DM who reports an incomplete preference relation and is observed to choose stochastically, Lemma 3 allows the analyst to check whether the reported preferences and the observed data are consistent. In this sense, Lemma 3 identifies the behavioral content of incomplete preferences and shows that it rests in the frequency of choice from binary menus.

Lemma 3 is relevant for selecting a “canonical” confidence model to represent a given preference, as discussed in the next section.

1.4 Uniqueness

Confidence models feature two parameters and hoping for both to be uniquely identified from preferences is clearly overoptimistic. A more reasonable expectation is for “conditional uniqueness”; that is, for $\alpha$ to be unique given $\mu$ and vice-versa. Unfortunately, Example 2 demonstrates that for the domains studied in this paper, even conditional uniqueness is too much to ask.

Example 2. Consider a reflexive, antisymmetric preference $\succsim$ on $X = \{x, y\}$. The set $\mathcal{R}$ consists of two relations: $R_1$ strictly prefers good $x$ while $R_2$ strictly prefers good $y$. Suppose a DM’s preferences are represented by the (regular) confidence model $(\mu, \alpha)$ where

$$\mu(R_1) = 3/4, \mu(R_2) = 1/4, \text{ and } \alpha = 2/3.$$ 

A brief inspection reveals that this DM’s preferences are actually identical to a standard DM with preference relation $R_1$. Moreover, fixing $\mu$ this will still be true for any
\( \alpha \in (1/2, 3/4] \). Alternatively, if we fix \( \alpha = 2/3 \) then preferences are the same for any \( \mu(R_1) \in [2/3, 1] \).

This example can easily be modified to illustrate nonuniqueness for preferences over lotteries. After normalizing, there are two strict expected utility functions on \( \Delta(X) \): \( u_1 \) strictly prefers the lottery \( \delta_x \) and \( u_2 \) strictly prefers the lottery \( \delta_y \). So, if we define \( \mu \in \Delta(\mathcal{U}) \) by \( \mu(u_1) = 3/4 \) and \( \mu(u_2) = 1/4 \) then the argument given above shows that the preferences on \( \Delta(X) \) represented by \( (\mu, 2/3) \) have many different confidence representations. \( \diamond \)

Facing this lack of uniqueness, one might hope to identify a “canonical” representation for a given confidence preference. In light of the tight connection between confidence preferences and random choice, a natural way to approach this problem is by gathering stochastic choice data. Specifically, suppose we have preference \( \succeq \) on \( \Delta(X) \) with a regular confidence representation, and we observe the DM make stochastic choices captured by an RCR \( \rho \). If \( \succeq \) and \( \rho \) are found to be compatible on binary sets, then the proof of Lemma 3 provides a confidence model \( (\mu, \alpha) \) that is an excellent candidate for a canonical representation of \( \succeq \). The measure \( \mu \) is the unique RUF that maximizes the DM’s observed choice data, and the threshold \( \alpha \) is easily seen to be the maximal threshold that can represent \( \succeq \) in conjunction with \( \mu \).

For preferences over finite sets, no such canonical representation can be identified: it is well known that a RCR \( \rho \) on \( X \) can be represented by more than one RUF \( \mu \in \Delta(\mathcal{R}) \) (see Fishburn (1998) for an example and Chapter 3 for extensive discussion of this issue). Nonetheless, the threshold \( \alpha \) constructed in the proof of Lemma 1 is still the maximal threshold that can represent \( \succeq \) in conjunction with any \( \mu \) that represents \( \rho \).
1.5 Related Literature

1.5.1 Multi-Utility Models and Transitivity

For finite $X$, Theorem 1(A) shows that confidence models nest any reflexive preference. This includes, of course, the “traditional” model for an economic DM: a single preference relation $W \in \mathcal{W}$ (take $\mu = \delta_W$ and any $\alpha > 1/2$). It also includes the multi-utility model axiomatized by Evren and Ok (2008). Those authors show that any reflexive and transitive preference can be represented by a set of weak orders $\mathcal{M} \subseteq \mathcal{W}$:

$$x \succsim y \iff xWy \text{ for all } W \in \mathcal{M}.$$

The interpretation of this model is that the DM considers a set of preference orderings to be possible, and requires unanimous agreement among these preferences before she is willing to express a preference. Unlike in confidence models, the DM does not express any probability judgments among the preferences in $\mathcal{M}$.

Like Multi-Utility models, Confidence models with $\alpha = 1$ require unanimous agreement before any preference is expressed. A simple corollary of Evren and Ok (2008)’s theorem shows that the existence of such a representation is characterized by transitivity.

**Corollary 2.** A preference $\succsim$ on a finite set $X$ is reflexive and transitive if and only if it can be represented by a confidence model $(\mu, \alpha)$ with $\alpha = 1$.

**Proof.** Necessity is obvious. For sufficiency, let $\mathcal{M}$ be a finite set of weak orderings for a multi-utility representation of $\succsim$ and let $\mu$ be the uniform measure on $\mathcal{M}$. Clearly, $(\mu, 1)$ is a confidence model for $\succsim$. 

In the context of preferences over lotteries, Dubra et al. (2004) show that any reflexive and transitive preference satisfying continuity and independence has an expected
multi-utility representation, in which

\[ p \succeq q \iff u \cdot p \geq u \cdot q \text{ for all } u \in \mathcal{M} \subseteq \mathcal{U} \]

where \( \mathcal{M} \) is closed and convex. The interpretation is the same as that for the finite case, and once again we can relate such preferences to confidence models with a threshold of one. A confidence preference on \( \Delta(X) \) represented by a confidence model with a threshold \( \alpha = 1 \) is transitive. In addition,

**Corollary 3.** Any preference \( \succeq \) on \( \Delta(X) \) satisfying the axioms of Dubra et al. (2004) has a confidence representation \( (\mu, \alpha) \) where \( \mu \in \Delta(U) \) is countably additive and \( \alpha = 1 \).

*Proof.* As before, let \( \mathcal{M} \subseteq \mathcal{U} \) be an expected multi-utility representation for \( \succeq \), and let \( m \) be a strictly-positive, countably-additive measure on \( \mathcal{M} \).\(^7\) If \( p \succeq q \), then \( u \cdot p \geq u \cdot q \) for all \( u \in \mathcal{M} \), and \( m(\{u \in \mathcal{M} \mid u \cdot p \geq u \cdot q\}) = 1 \). On the other hand, if \( \neg(p \succeq q) \), then there exists \( u \in \mathcal{M} \) such that \( u \cdot p < u \cdot q \), which implies that there exists a relatively open set \( U \subseteq \mathcal{M} \) such that \( u \cdot p < u \cdot q \) for all \( u \in U \). This implies that \( m(\{u \in \mathcal{M} \mid u \cdot p \geq u \cdot q\}) < 1 \), showing that \( (m, 1) \) is a confidence model for \( \succeq \). \( \blacksquare \)

Although the results in Section 1.3 did not provide a characterization of confidence models on the domain \( \mathcal{L} = \Delta(X) \), the preceding corollary establishes that such models do in fact nest expected multi-utility preferences, and clearly they nest traditional EU preferences.

\(^7\)Since \( \mathcal{M} \) is closed and hence compact, such a measure exists; consult Corollary 2.8 of Herbert and Lacey (1968).
1.5.2 Stochastic Preference

There is a long tradition in the literature on stochastic choice\(^8\) of using a random choice rule \(\rho\) to define a stochastic preference \(\succeq_\rho\) by

\[
x \succeq_\rho y \iff \rho^{(x,y)}(x) \geq \rho^{(x,y)}(y).
\]

To connect this idea with confidence models, let \(X\) be a finite set and \(\rho\) a RCR on \(X\) satisfying the BM-inequalities. Clearly, the preference \(\succeq_\rho\) is complete, so by Theorem 1(A) it has a confidence representation. If elements of binary menus are never chosen with probability \(1/2\) according to \(\rho\)—that is, if \(\rho^{(x,y)}(x) \neq 1/2\) for all \(x, y \in X\)—then \(\succeq_\rho\) is antisymmetric and by Theorem 1(B) has a regular confidence representation. In the latter case, it is clear that \(\succeq_\rho\) and \(\rho\) are compatible on binary sets, so any \(\mu \in \Delta(\mathcal{R})\) that represents \(\rho\) can be paired with an appropriate \(\alpha \in (1/2, 1]\) to form a regular confidence representation for \(\succeq_\rho\). However, if there are elements \(x\) and \(y\) such that \(\rho^{(x,y)}(x) = 1/2\), then \(\succeq_\rho\) is not antisymmetric. Furthermore, \(\succeq_\rho\) and \(\rho\) will fail to satisfy generalized binary compatibility, and a measure \(\mu\) representing \(\rho\) cannot form part of a confidence representation for \(\succeq_\rho\). This situation arises due to the requirement for confidence models that \(\alpha > 1/2\), which does not prevent confidence models from nesting complete preferences, but does constrain the stochastic choice data that is consistent with such preferences.

1.5.3 Incomplete Preferences on Anscombe-Aumann Acts

There are certainly domains other than finite sets or lotteries over finite sets on which confidence models might prove interesting. On infinite sets without a lottery structure, one could study confidence models featuring measures on the set of continuous utility

\(^8\)See Fishburn (1998) for an overview of this literature.
functions. But perhaps the most interesting domain to consider is the set of Anscombe-Aumann acts on a finite state space. The classic model for incomplete preferences on this domain is that of Bewley (2002), in which a DM considers a set $\mathcal{M}$ of probability measures on the state space and expresses a preference $f \succeq g$ iff $E_p[u \circ f] \geq E_p[u \circ g]$ for every measure $p \in \mathcal{M}$ and a fixed expected utility function $u$. Confidence models could generalize this representation by allowing for probabilistic uncertainty over $\mathcal{M}$, and perhaps over $u$ as well.\(^9\)

Confidence models on this domain could potentially be used to address interesting questions on the arrival of information and the updating of incomplete preferences. In addition, confidence models could connect incomplete preferences over acts to random choice over acts, as studied recently by Lu (2013).

Minardi and Savochkin (2013) work with a “graded preference relation” $\mu$ where $\mu(f, g)$ is interpreted as the DM’s confidence that $f$ is better than $g$, and introduce a rule for connecting this primitive to deterministic choice. Comparing and contrasting this approach with the link between incomplete preferences and stochastic choice highlighted in this paper may prove to be an interesting exercise.

---

\(^9\)Ok et al. (2012) and others modify the Bewley model to accommodate a set of utility functions.
Chapter 2

Non-Additive Random Utility Functions

2.1 Introduction

A random choice rule (RCR) on a finite set can be represented by a random utility function (RUF) if and only if its associated Block-Marshack polynomials are all nonnegative. These inequalities can be quite complex: as the number of available alternatives grows both the set of Block-Marshack inequalities and the individual polynomials themselves expand. Consequently, for all but the smallest sets of alternatives the complete collection of Block-Marschak inequalities is difficult to interpret.

The first goal of this paper is to “unravel” the Block-Marshack inequalities by progressively relaxing them to study new representations that are more general than random utility. Specifically, the representations will feature set functions on the space of preferences that are capacities rather than probability measures.

1 This chapter benefited from helpful discussions with Drew Fudenberg and Jerry Green. Tomasz Strzalecki suggested the line of research which led to this project and provided excellent advice. All remaining errors are my own.
One the simplest conditions to impose on a RCR is regularity, which demands that adding alternatives to a choice set cannot increase the frequency with which existing alternatives are chosen. Proposition 1 shows that a RCR satisfies regularity if and only if it can be represented by a superadditive capacity on the space of preferences. The higher-level Block-Marschak inequalities are then related to the degree of monotonicity that a capacity representation of a RCR can satisfy.

These results promote a deeper understanding of the Block-Marschak inequalities. Technically, the Block-Marschak inequalities are shown to interact in somewhat-subtle ways with the normalization equalities, which require that choice frequencies from any set sum to one. A surprising consequence of the analysis is that the full set of Block-Marschak inequalities is necessary (as well as sufficient) for a RCR to be represented by a belief function. Therefore, there is no observable difference between representations featuring traditional probability measures and those featuring belief functions.

Secondly, this paper examines the connection between random choice over finite sets and random choice over lotteries. In the latter case, the necessary and sufficient conditions for random utility maximization are—at least superficially—much simpler than the Block-Marschak inequalities. Proposition 4 provides a formal sense in which the mixture continuity axiom substitutes for the higher-order Block-Marschak inequalities for random choice over lotteries.

The next section discusses related literature and provides the necessary background on random choice. Proofs omitted from the main text may be found in the appendix.

2.2 Background and Related Literature

Let $X$ be a finite set of alternatives with $|X| = N$, and let $\mathcal{D} := 2^X \setminus \{\emptyset\}$ denote the collection of all nonempty subsets of $X$. A random choice rule (RCR) is a function
\( \rho : \mathcal{D} \times X \to [0, 1] \) with the property that for all \( D \in \mathcal{D} \),

\[
\sum_{x \in D} \rho^D(x) = 1.
\]

Here \( \rho^D(x) := \rho(D, x) \) is interpreted as the probability of choosing \( x \) from the set \( D \).

The above equations are called the *normalization equalities*, and they guarantee that this interpretation makes sense by requiring that choice frequencies for every menu sum to one.

Let \( \mathcal{R} \) denote the set of complete, transitive, and antisymmetric relations on \( X \). A *random utility function (RUF)* is a probability measure \( \mu \in \Delta(\mathcal{R}) \). A RUF is said to *represent* a RCR if for all \( (D, x) \in \mathcal{D} \times X \),

\[
\rho^D(x) = \mu(\{ R \in \mathcal{R} \mid x \text{ is } R\text{-optimal in } D \}).
\]

Block and Marschak (1960) introduced the following collection of polynomial inequalities, now known as the *Block-Marschak (BM) inequalities*:

\[
\sum_{C : D \subseteq C} (-1)^{|C \setminus D|} \rho^C(x) \geq 0 \quad \text{for all } (D, x) \in \mathcal{D} \times X \text{ with } x \in D
\]

and demonstrated that these inequalities are necessary for a RCR to be represented by some RUF. Falmagne (1978) proved that the BM inequalities are also sufficient for the existence of a random utility representation. This result was rediscovered and popularized among economists by Barberá and Pattanaik (1986).

The following example partially enumerates the BM inequalities for a small set of alternatives.

**Example 3.** Let \( X = \{x, y, z\} \). The BM inequalities are indexed by pairs \( (D, x) \) where \( x \in D \). With \( x \) as the selected alternative, there are four BM inequalities
(similar inequalities are required for $y$ and $z$ as the selected alternatives):

$$
\rho^X(x) \geq 0 \quad \text{(Take } D = X) \\
\rho^{x,y}(x) - \rho^X(x) \geq 0 \quad \text{(Take } D = \{x, y\}) \\
\rho^{x,z}(x) - \rho^X(x) \geq 0 \quad \text{(Take } D = \{x, z\}) \\
1 - \rho^{x,y}(x) - \rho^{x,z}(x) + \rho^X(x) \geq 0 \quad \text{(Take } D = \{x\})
$$

Note that first inequality (with $D = X$) holds trivially, since RCRs were defined to take values in $[0, 1]$. The second and third inequalities are particular instances of regularity. For example, the second inequality demands that adding $z$ to $\{x, y\}$ cannot increase the probability of selecting $x$. It turns out that the final inequality is redundant: Lemma 11 shows that any BM inequality indexed by $(x, \{x\})$ is implied by other BM inequalities and the normalization equalities. ◊

The BM inequalities above are written in the form of Fiorini (2004). In that paper, Fiorini provides a novel proof that the BM inequalities are sufficient for a random utility representation, utilizing the concept of Möbius inversion and tools from the theory of network flows. Möbius inversion also plays a critical role in Gul and Pesendorfer (2013), which characterizes random utility models that permit indifference.

Fishburn (1998) and more recently McFadden (2005) provide excellent overviews on the theory of random utility maximization.

### 2.3 Capacity Representations

This section studies representations of RCRs that feature capacities on the space of preferences.

**Definition 10.** A **capacity** on a finite set $S$ is a function $\nu : 2^S \to [0, 1]$ that satisfies
\( \nu(\emptyset) = 0, \nu(S) = 1, \) and \( \nu(E) \leq \nu(F) \) whenever \( E \subseteq F. \)

Capacities are normalized set functions that are monotonic with respect to set inclusion, but need not be additive.

**Definition 11.** A capacity \( \nu \) on \( \mathcal{R} \) represents a RCR \( \rho \) if

\[
\rho^D(x) = \nu(\{R \in \mathcal{R} \mid x \text{ is } R\text{-optimal in } D\})
\]

for all \((D, x) \in \mathcal{D} \times X.\)

The goal of this section is to characterize the existence of a capacity representation for a RCR. An immediate necessary condition is that the RCR satisfy regularity, which demands that adding elements to a choice set cannot increase the frequency with which existing elements are selected. Formally,

**Definition 12 (Regularity).** A RCR \( \rho \) satisfies regularity if \( \rho^D(x) \geq \rho^E(x) \) whenever \( D \subseteq E. \)

Proposition 1 shows that regularity is also sufficient for the existence of a capacity representation. Moreover, no additional conditions are required to ensure representation by a superadditive capacity. Intuitively, this occurs because the normalization equalities impose some additive structure on any RCR.

**Proposition 1.** Let \( \rho \) be a RCR on \( X. \) TFAE:

1. \( \rho \) satisfies regularity.

2. There exists a capacity \( \nu \) on \( \mathcal{R} \) that represents \( \rho. \)

3. There exists a superadditive capacity \( \nu \) on \( \mathcal{R} \) that represents \( \rho; \) i.e. \( \nu \) satisfies

\[
\nu(A \cup B) \geq \nu(A) + \nu(B)
\]

whenever \( A, B \subseteq \mathcal{R} \) are disjoint.
Clearly (3) \(\Rightarrow\) (2) \(\Rightarrow\) (1). So it suffices to show that (1) \(\Rightarrow\) (3); the proof of this fact may be found in the appendix, along with a general extension theorem for superadditive capacities. Demonstrating that (1) \(\Rightarrow\) (2) illustrates the central idea.

Proof of (1) \(\Rightarrow\) (2). Given is an RCR \(\rho\) satisfying regularity. For any pair \((D, x) \in \mathcal{D} \times X\), define the set of relations

\[ N(D, x) := \{ R \in \mathcal{R} \mid x \text{ is } R\text{-optimal in } D \}. \]

Next, let \(A \subseteq 2^{\mathcal{R}}\) be all subsets \(A\) of \(\mathcal{R}\) that can be written in the form \(A = N(D, x)\) for some \((D, x) \in \mathcal{D} \times X\). On \(A\), define a set function \(\hat{\nu} : A \to [0, 1]\) by

\[ \hat{\nu}(A) = \rho^D(x) \]

where \((D, x)\) is the unique pair in \(\mathcal{D} \times X\) that defines the set \(A\). Observe that \(\hat{\nu}\) is monotonic on \(A\) since \(\rho\) satisfies regularity.

Set \(\hat{\nu}(\emptyset) = 0\) and \(\hat{\nu}(\mathcal{R}) = 1\). All that remains is to extend \(\hat{\nu}\) to a capacity on \(\mathcal{R}\). A natural approach is to define

\[ \nu(E) := \max\{\hat{\nu}(A) \mid A \subseteq E, A \in A \cup \{\emptyset\} \cup \{\mathcal{R}\}\} \]

for each \(E \subseteq \mathcal{R}\). As shown in Proposition 2.4 of (Denneberg, 1994), this extension procedure creates a capacity that represents \(\rho\).

Loosely, the set \(A\) contains subsets of \(\mathcal{R}\) whose \(\nu\)-value is directly constrained by the data. So the question becomes “when can one create a set function on \(\mathcal{R}\) that respects \(\rho\) on \(A\) and also satisfies some desirable properties both on and off of \(A\)?” For the case of capacities, the desired property is monotonicity, and the answer to the preceding question is “whenever \(\rho\) is regular.”

The set \(A\) is a strict subset of \(2^{\mathcal{R}}\), and it is easy to see that capacity representations
of a RCR are not unique. The construction above identified the minimal capacity that can represent a regular random choice rule. To find the maximal capacity, set

$$\nu(E) := \inf \{ \hat{\nu}(A) \mid A \supseteq E, A \in \mathcal{A} \cup \{\emptyset\} \cup \{\mathcal{R}\} \}$$

in the final step of the above proof (see the discussion on page 21 of (Denneberg, 1994)).

In the case of traditional RUFs, the set function representing $\rho$ is a probability measure on $\mathcal{R}$. Unfortunately, the collection $\mathcal{A}$ is not an algebra,\textsuperscript{2} so traditional extension theorems cannot be used and the proof strategy above fails. A RUF $\mu$ must of course be additive, so the behavior of $\mu$ is constrained even off the set $\mathcal{A}$. Nevertheless, whenever $N \geq 4$ random utility representations are not unique. This issue is discussed extensively in Chapter 3.

Traditional RUFs have been interpreted in two ways. First, a RUF $\mu$ can be taken to represent a distribution of preferences for a single agent; in this case $\rho$ arises from repeated samples of an individual’s choice behavior at different points in time. Proposition 1 shows that an individual who is observed to choose stochastically and satisfies regularity can be modeled as if she has random preferences captured by a capacity.

Alternatively, $\mu$ could represent the distribution of individuals’ preferences across a society; in this case each individual has deterministic preferences and $\rho$ summarizes choice behavior across the population. This interpretation implicitly assumes that an additive distribution of preferences exists, so it is unclear how to apply Proposition 1 in a manner consistent with this interpretation.

Under a third, novel interpretation $\rho$ captures the predictions on an expert re-

\textsuperscript{2}Although $\mathcal{A}$ is not an algebra, it has an interesting structure: $\mathcal{A}$ is a collection of partitions of $\mathcal{R}$. This structure could potentially be exploited to deliver new characterizations for random utility models.
garding the choice behavior of either an individual or a population.³ In this case, \( \mu \) represents the expert’s uncertainty regarding the true distribution of preferences. Under Proposition 1, if an expert’s predictions satisfy regularity then they can modeled as arising from non-additive beliefs about the likelihood of particular preferences being realized. Formalizing this idea by constructing scoring rules for stochastic predictions is the subject of on-going work, complicated by the fact that proper scoring rules for capacities do not exist (as shown by Chambers (2008)).

### 2.4 Unraveling the BM Inequalities

#### 2.4.1 Supermodularity

Proposition 1 shows that regularity alone suffices to guarantee the existence of a superadditive capacity representation for a RCR. The remaining higher-order BM inequalities must therefore facilitate the move from superadditivity to full additivity. First, consider strengthening superadditivity to supermodularity.

**Definition 13.** A capacity \( \nu \) on a set \( S \) satisfies supermodularity (also called 2-monotonicity or convexity) if

\[
\nu(A \cup B) + \nu(A \cap B) \geq \nu(A) + \nu(B)
\]

for all \( A, B \subseteq S \).

In general, the capacities constructed in the proof of Proposition 1 are not supermodular. The following example illustrates how the BM inequalities interact with supermodularity.

³I thank Tomasz Strzalecki for suggesting this interpretation.
Example 4. Let $\rho$ be a RCR on $X = \{x, y, z, w\}$. The BM inequality indexed by $(\{x, y\}, x)$ states that

$$
\rho^{\{x,y\}}(x) - \rho^{\{x,y,z\}}(x) - \rho^{\{x,y,w\}}(x) + \rho^X(x) \geq 0.
$$

The supplement to Gul and Pesendorfer (2006) interprets this inequality as demanding that the effect on the choice frequency of $x$ from adding $z$ to the set $\{x, y\}$ is at least as great as the effect from adding $z$ to the larger set $\{x, y, w\}$.

Suppose $\rho$ is represented by a capacity $\nu$. Let

$$
E := N(\{x, y\}, x), \quad A := N(\{x, y, z\}, x), \quad \text{and} \quad B := N(\{x, y, w\}, x).
$$

If the above BM inequality is violated, then

$$
\nu(A \cup B) + \nu(A \cap B) \leq \nu(E) + \nu(A \cap B) < \nu(A) + \nu(B),
$$

which shows that $\nu$ cannot be supermodular. ◇

The following proposition identifies the reason behind the choice of $(\{x, y\}, x)$ as the indexing pair for the BM inequality in the previous example: namely, in that example the set $\{x, y\}$ has two fewer elements than the complete set of alternatives.

**Proposition 2.** Let $\rho$ be a RCR on $X$ and let $(D, x) \in \mathcal{D} \times X$ with $x \in D$. If $|D| = N - 2$, the BM inequality indexed by $(D, x)$ is a necessary condition for $\rho$ to be represented by a supermodular capacity.

I conjecture but have been unable to prove that the following condition (which generalizes the BM inequalities indexed by sets with $N - 2$ elements) is both necessary and sufficient for a regular RCR to be represented by a supermodular capacity:

$$
\rho^D(x) - \rho^{D \cup B}(x) - \rho^{D \cup C}(x) + \rho^{D \cup C \cup B}(x) \geq 0 \text{ for all } B, C, D \subseteq X.
$$
Chateauneuf and Jaffray (1989) provide a characterization of supermodular capacities in terms of Möbius inversion that could prove useful in establishing this characterization.

Before moving on, note that the difference between supermodular and submodular representations is flimsier than one might expect given that the BM inequalities related to supermodularity are only required to hold in one direction.

**Lemma 4.** Let $\mathcal{A}$ be defined as in the proof of Proposition 1. A RCR $\rho$ can be represented by a supermodular capacity $\nu$ that satisfies $\nu(A) = 1 - \nu(A^c)$ for all $A \in \mathcal{A}$ iff it can be represented by a submodular capacity $\nu'$ that satisfies $\nu'(A) = 1 - \nu'(A^c)$ for all $A \in \mathcal{A}$.

**Proof.** Let $\nu$ be a supermodular representation for $\rho$ satisfying the condition of the proposition. By Proposition 2.3 of Denneberg (1994), the set function $\eta$ defined by $\eta(E) = 1 - \nu(E)$ for all $E \in 2^R$ is a submodular capacity. Since $\nu$ and $\eta$ agree on $\mathcal{A}$, $\eta$ also represents $\rho$. The same construction works in the opposite direction after replacing $\nu$ with $\nu'$.

### 2.4.2 $K$-Monotonicity

The BM inequalities indexed by $(D, x)$ become increasingly cumbersome as the cardinality of $D$ decreases. To understand how these inequalities affect capacity representations of a RCR, consider the following generalization of supermodularity based upon the inclusion-exclusion principle:

**Definition 14.** A capacity $\nu$ is called $K$-monotonic (or monotone of order $K$) if

$$\nu \left( \bigcup_{k=1}^{K} A_k \right) \geq \sum_{\emptyset \neq I \subseteq \{1, \ldots, K\}} (-1)^{|I|+1} \nu \left( \bigcap_{k \in I} A_k \right)$$

for any $(A_1, \ldots, A_K)$ with $A_k \subseteq S$ for all $k$. 
This definition is best illustrated by considering the case of three sets, $A_1$, $A_2$, and $A_3$. The inclusion-exclusion principle demands that

$$|A_1 \cup A_2 \cup A_3| = |A_1| + |A_2| + |A_3| - |A_1 \cap A_2| - |A_1 \cap A_3| - |A_2 \cap A_3| + |A_1 \cap A_2 \cap A_3|.$$  

Loosely, a 3-monotonic capacity needs to respect this equality by assigning $A_1 \cup A_2 \cup A_3$ “enough” measure relative to its underlying component sets.

The following result shows that violations of higher-order BM inequalities limit the order of monotonicity that a capacity representation of a RCR can satisfy.

**Proposition 3.** Let $\rho$ be a RCR on $X$ and let $(D, x) \in D \times X$ with $x \in D$. If $|D| = N - K$, the BM inequality indexed by $(D, x)$ is a necessary condition for $\rho$ to be represented by a $K$-monotonic capacity.

Lemma 11 in the Appendix demonstrates that the BM inequalities indexed by $(\{x\}, x)$ are irrelevant in that they are implied by the remaining BM inequalities and the normalization equalities. Therefore, in light of the preceding results, establishing that a RCR $\rho$ can be represented by an $(N - 2)$-monotonic capacity is equivalent to showing that all the BM inequalities hold, and thus that $\rho$ can be represented by a standard RUF.

**Corollary 4.** A RCR $\rho$ can be represented by a RUF if and only if it can be represented by a capacity of order $N - 2$.

Capacities that are $K$-monotonic for any $K$ are called *belief functions*. Introduced by Dempster (1967) and Shafer (1976), belief functions are used in statistics and other areas as a method for aggregating evidence to produce a “degree of belief” in some uncertain event. Belief functions need not be additive and thus strictly generalize probability measures.
By Corollary 4, a RCR can be represented by a belief function if and only if it can be represented by a RUF. So demanding that a capacity representation for a RCR satisfy additivity as opposed to merely $K$-monotonicity for all $K$ has no empirical content. Intuitively, this surprising fact illustrates the somewhat subtle interaction between the BM inequalities and the normalization equalities. By themselves, the normalization equalities add a significant amount of additivity to a RCR. The set $\mathcal{A}$ constructed in Section 2.3 consists of several partitions of the set $\mathcal{R}$, and the normalization equalities ensure that the set function $\hat{\nu}$ forms a probability measure when restricted to each of these partitions individually. This structure serves implicitly to negate the difference between representations based on probability measures and those based on belief functions.

2.5 Connection to Random Choice over Lotteries

Continue to let $X$ represent a finite set of alternatives. Gul and Pesendorfer (2006) study random choice over $\Delta(X)$, the set of lotteries on $X$. Let $\mathcal{D}$ denote the set of all nonempty, finite subsets of $\Delta(X)$ and let $\Delta(\Delta(X))$ denote the set of all simple probability measures on $\Delta(X)$. A random choice rule (RCR) is function $\rho: \mathcal{D} \to \Delta(\Delta(X))$, where $\rho^D(p)$ denotes the probability that a lottery $p$ is chosen from $D$. The normalization equalities require that $\rho^D(D) := \sum_{p \in D} \rho^D(p) = 1$ for all $D$.

In the context of random choice over lotteries, a RUF is a (finitely-additive) measure $\mu$ on the set of expected utility functions over $\Delta(X)$.\(^4\) Gul and Pesendorfer (2006) prove that the following list of (paraphrased) axioms are necessary and sufficient for the existence of a random expected utility representation:

**GP1** Regularity: if $p \in D \subset D'$, then $\rho^{D'}(p) \leq \rho^{D}(p)$.

\(^4\)See Gul and Pesendorfer (2006) for the precise definition of a RUF.
GP2 Linearity: $\rho^{\lambda D + (1-\lambda)\{q\}}(\lambda p + (1-\lambda)\{q\}) = \rho^D(p)$.

GP3 Mixture continuity: $\rho^{\lambda D + (1-\lambda)D'}$ is continuous in $\lambda$.

GP4 Extremeness: $\rho^D(\text{ext}(D)) = 1$.

In the supplement to their paper, Gul and Pesendorfer show that any RCR on $X$ that satisfies the BM inequalities can be extended to a RCR on $\Delta(X)$ in such a way that choice over degenerate lotteries agrees with the original RCR and all of GP1–4 are satisfied.

GP1–4 appear to be simple axioms. Regularity is interpreted exactly as it is for random choice on finite sets. Linearity and mixture continuity are stochastic versions of the corresponding axioms from the standard theory of expected utility. And extremeness is a technical axiom that intuitively holds because any expected utility preference has linear indifference curves.

Nevertheless, the preceding result provides a sense in which GP1–4 encompass the BM inequalities, which are much harder to interpret. While it is not surprising that more elegant results can be obtained when using $\Delta(X)$ instead of $X$ as a domain, a question still arises: is it possible to identify which axiom(s) from GP1–4 are serving to replace the higher-order BM inequalities? The following proposition provides an answer.

**Proposition 4.** Let $\rho$ be a RCR on $X$. If $\rho$ satisfies regularity, then it can be extended to a RCR on $\Delta(X)$ in such a way that

1. Choice over degenerate lotteries agrees with the original RCR; and

2. GP1, GP2, and GP4 are satisfied.

In words, a RCR on $X$ that satisfies only regularity—and potentially none of the higher-order BM inequalities—can be extended to a RCR on $\Delta(X)$ that satisfies all of
the Gul and Pesendorfer (2006) axioms except for mixture continuity. In this sense, mixture continuity serves to replace all the power of the BM inequalities beyond the simple property of regularity.

The proof of Proposition 4 invokes Proposition 1 to obtain a capacity representation for $\rho$, and in turn uses this representation to extend $\rho$ to $\Delta(X)$:

**Proof.** Let $\rho$ be a RCR on $X$ and let $\nu$ be a capacity representation for $\rho$, which exists by Proposition 1. For each $R \in \mathcal{R}$, let $u_R$ be any expected utility function that agrees with $R$ when restricted to degenerate lotteries. Define a RCR $\hat{\rho}$ on $\Delta(X)$ by

$$
\rho^D(p) := \nu(\{ R \in \mathcal{R} \mid p \text{ is } u_R\text{-dominant in } D \})
$$

if $p$ is an extreme point of $D$ and $\rho^D(p) = 0$ otherwise, with ties over extreme lotteries broken uniformly. By construction $\hat{\rho}$ satisfies extremeness. It satisfies regularity because $\nu$ is monotonic. Finally, for any $R \in \mathcal{R}$, any $p, q \in \Delta(X)$, and any $D \subseteq \Delta(X)$, $p$ is $u_R$-optimal in $D$ iff $\lambda p + (1 - \lambda)q$ is $u_R$-optimal in $\lambda D + (1 - \lambda)q$. Therefore $\hat{\rho}$ is linear.

$\blacksquare$
Chapter 3

Unique Random Utility Representations\textsuperscript{1}

3.1 Introduction

A random choice rule (RCR) over a finite set of alternatives can be represented by a random utility function (RUF) if and only if the RCR satisfies a set of polynomial inequalities known as the Block-Marschak inequalities. However, when there are at least four alternatives RUFs are not unique. In fact, there are RUFs with disjoint support that nonetheless represent the same RCR. Section 3.2 studies the extent of this nonuniqueness problem in detail.

The obvious strategy for improving identification is to enrich the domain. Gul and Pesendorfer (2006) consider random choice over lotteries and provide a unique distribution over expected utility functions using simple axioms. Their representation delivers a unique distribution over preferences, but requires an infinite number of menus.

\textsuperscript{1}This chapter benefited from helpful discussions with Mira Frick, Drew Fudenberg, Ben Golub, and Tomasz Strzalecki. All remaining errors are my own.
and restricts attention to expected utility preferences. These features are undesirable for an analyst who does not want to commit to expected utility and only wishes to identify a measure over ordinal preferences.

This paper maintains a finite set of alternatives but introduces a finite state space and studies random choice over Savage acts. Choice data is restricted to complete menus, which are shown to be rich enough for ordinal preferences to completely determine choice. In particular, any complete menu $D$ contains enough acts so that for any ordinal preference $R$, there exists an act $f \in D$ that delivers a weakly optimal consequence in every state. A representation is developed and axiomatized in which the probability that an act is chosen from a complete menu is the probability that it is dominant within the menu. The measure over ordinal preferences used in this representation is unique whenever the state space is large enough. Thus, this paper provides unique random utility models with a finite number of menus and while making only minimal implicit assumptions on attitudes towards risk or uncertainty.

All proofs omitted from the text and a review of the related literature may be found in the Appendix.

3.2 The Problem of Nonuniqueness

Let $X$ be a finite set of alternatives with $|X| = N > 2$. A random choice rule (RCR) over alternatives is a function $P: 2^X \setminus \{\emptyset\} \to [0,1]$ that satisfies the normalization equalities $\sum_{x \in D} P^D(x) = 1$ for all nonempty sets $D \subseteq X$. Here $P^D(x) := P(D, x)$ is interpreted as the probability that alternative $x$ is selected from the menu $D$.

Let $\mathcal{R}$ denote the set of complete, transitive, and antisymmetric binary relations on $X$. A random utility function (RUF) is a measure $\mu \in \Delta(\mathcal{R})$. A RCR is said to be
represented by a RUF \( \mu \) if

\[
P^D(x) = \mu(\{R \in \mathcal{R} \mid x \text{ is } R\text{-optimal in } D\})
\]

for all alternative-menu pairs \((x, D)\) with \(x \in D\). It is well known that a RCR is represented by some RUF if and only if the Block-Marschak inequalities

\[
\sum_{C: D \subseteq C} (-1)^{|C \setminus D|} P_C(x) \geq 0
\]

hold for all alternative-menu pairs \((x, D)\) with \(x \in D\).²

Any given RUF represents exactly one RCR. When \(N = 3\), the converse holds as well—a RCR can be represented by at most one RUF (this follows from Proposition 6 below). However, when \(N = 4\) the following classic example from Fishburn (1998) demonstrates that uniqueness of RUF representations can fail.

**Example 5.** Let \(X = \{x, y, z, w\}\). Consider the following RUFs:

\[
\mu_1 = \begin{cases} 
xywz & \text{with probability } 1/2; \\
yxz\bar{w} & \text{with probability } 1/2 
\end{cases}
\quad \text{and} \quad
\mu_2 = \begin{cases} 
xyzw & \text{with probability } 1/2; \\
yxz\bar{w} & \text{with probability } 1/2 
\end{cases}
\]

These RUFs are not only distinct but in fact have disjoint support. Nevertheless, it may be verified by inspection that they represent exactly the same RCR over alternatives. 

Failures of uniqueness such as those demonstrated in Example 5 complicate the interpretation of a RUF representation and severely limit their applications. Intuitively, the two RUFs above capture different correlations between binary preferences. Distinguishing between these two models could matter for a social planner or mechanism designer who wants to offer agents lotteries or menus of alternatives, but can only

²The inequalities were introduced by Block and Marschak (1960) and are written here in the form of Fiorini (2004). The characterization was established by Falmagne (1978) and rediscovered by Barberá and Pattanaik (1986).
observe random choice over alternatives.

Failures of uniqueness might be safely dismissed as pathological examples if they were rare, but Proposition 5 provides a precise sense in which failures of uniqueness are the rule rather than the exception.

**Proposition 5.** Assume that \( N \geq 4 \). Recall that \( \Delta(\mathcal{R}) \) denotes the set of all RUFs and let \( \mathcal{P} \) denote the set of all RCRs on \( X \) that can be represented by some RUF.

1. For any \( \mu \in \text{int}(\Delta(\mathcal{R})) \), there exists \( \mu' \in \Delta(\mathcal{R}) \) such \( \mu \neq \mu' \) but nonetheless \( \mu \) and \( \mu' \) represent the same RCR.

2. Under the natural topology on \( \mathcal{P} \), the set of RCRs that can be represented by two distinct RUFs is dense in \( \mathcal{P} \).

The heart of Proposition 5 is generalizing Example 5 to show that for any RUF \( \mu \) with full support one can construct a different RUF \( \mu' \) that represents the same RCR as \( \mu \).

Random utility representations can be viewed as the solution to a system of linear inequalities (see Theorem 3.3 in McFadden (2005)). Although not pursued here, it is likely that nonuniqueness results similar to Proposition 5 could be obtained by studying the dimensionality of these systems.

The next two sections solve the pervasive uniqueness problem for RCRs by introducing states and considering random choice over Savage acts.

### 3.3 The Model

Let \( S = \{s_1, \ldots, s_M\} \) be a collection of \( M \) states of nature. It will be shown in Section 3.5 that \( N - 2 \) is the minimal number of states needed to identify a unique RUM. An act is a function \( f : S \rightarrow X \) and the set of all acts is denoted \( \mathcal{F} \). For a menu
of acts $D \subseteq \mathcal{F}$, define $D(s_i) := \bigcup_{f \in D} \{f(s_i)\}$ to be the set of alternatives that could be obtained in the $i$’th state.

**Definition 15.** A menu $D \subseteq \mathcal{F}$ is called complete if it contains all acts $g$ satisfying $g(s_i) \in D(s_i)$ for all $i = 1, \ldots, M$.

For example, let $X = \{x, y\}$ and $S = \{s_1, s_2\}$. Denote acts by $f = (f(s_1), f(s_2))$. The menu $D = \{(x, y), (y, x)\}$ is not complete because neither $(x, x)$ nor $(y, y)$ are in $D$. Given any menu $C \subseteq \mathcal{F}$, the complete menu generated by $C$ is defined to be the intersection of all complete menus containing $C$.

Let $\mathcal{D}$ denote the set of all nonempty complete menus. A random choice rule over complete menus of acts is a function $\rho : \mathcal{D} \times \mathcal{F} \to [0, 1]$ that satisfies $\sum_{f \in D} \rho^D(f) = 1$ for all $D \in \mathcal{D}$, where $\rho^D(f) := \rho(D, f)$ is interpreted as the probability of choosing the act $f$ from the menu $D$.

Complete menus are rich enough to render almost irrelevant any considerations of risk or uncertainty. To be more precise, let $\mathcal{R}$ denote the set of complete, transitive, and antisymmetric binary relations on $X$. Then the following result holds:

**Lemma 5.** Let $D \in \mathcal{D}$ be a complete nonempty menu. For any $R \in \mathcal{R}$, there exists a unique act $f \in D$ such that $f(s) R g(s)$ for all $s \in S$ and $g \in D$. Such an act is said to be $R$-dominant in $D$.

**Proof.** For each $i = 1, \ldots, M$ let $x_i$ denote the unique $R$-maximal element of $D(s_i)$. The act $f$ defined by $f(s_i) = x_i$ is in $D$ because $D$ is complete and is $R$-dominant in $D$ by construction. 

A DM’s preferences $\succeq$ over $\mathcal{F}$ are state-wise monotonic if

$$f(s) \succeq g(s) \text{ for all } s \in S \Rightarrow f \succeq g.$$
When confronted with a complete menu $D$, any DM with state-wise monotonic preferences over acts and ordinal preferences given by $R$ will select the $R$-dominant act from $D$, regardless of her beliefs (probabilistic or otherwise). This is the precise sense in which attitudes towards risk or uncertainty are irrelevant for complete menus.

Exactly as in Section 3.2, a probability measure $\mu \in \Delta(\mathcal{R})$ is called a random utility function (RUF). In light of the preceding lemma, it makes sense to study the following representations.

**Definition 16.** A RCR over complete menus of acts $\rho$ is represented by a RUF $\mu$ if for all $D \in \mathcal{D}$

$$\rho^D(f) = \mu\{R \in \mathcal{R} \mid f \text{ is } R\text{-dominant in } D\}.$$

A RCR choice rule $\rho$ that is represented by a RUF $\mu$ can be interpreted as reflecting a situation where ordinal preferences may vary but state-wise monotonicity is always respected. As usual, $\mu\{R\}$ can be interpreted either as the probability that a single individual has ordinal preference $R$ at a specific moment in time, or as the proportion of individuals in a society with fixed ordinal preference $R$.

In accordance with the discussion following Lemma 5, note that beliefs (probabilistic or otherwise) over the state space $S$ play no role in this type of representation. Unlike the approach in Gul and Pesendorfer (2006), which restricts attention to expected utility preferences, the only assumption encoded by this representation is that the underlying (implicit) preferences over acts satisfy state-wise monotonicity. Intuitively, this rules out state-dependent preferences. Monotonicity is satisfied by the Savage model of subjective expected utility as well as many models of ambiguity aversion.

To relate RCRs over complete menus of acts to RCRs over alternatives, let $C \subset \mathcal{F}$ be a menu of constant acts and let $D$ be the complete menu generated by $C$. Fix a RUF $\mu$. Let $\rho$ be the RCR over complete menus of acts that is represented by $\mu$ and let $P$ be the RCR over alternatives that is represented by $\mu$. Then it is straightforward
to establish that $\rho^D(\bar{x}) = P^C(x)$ for all $\bar{x} \in C$, where $\bar{x}$ denotes the constant act that always delivers $x$ and notation has been abused by treating $C$ simultaneously as a subset of $\mathcal{F}$ and $X$.

The core idea of this paper is that RCRs over complete menus of acts have more power to distinguish between RUF representations than RCRs over alternatives. The following example illustrates this idea by reconsidering Example 5.

**Example 6.** Let $X = \{x, y, z, w\}$ and $S = \{s_1, s_2\}$. Acts are written as ordered pairs where $(x, y)$ represents the act that returns $x$ in state $s_1$ and $y$ in state $s_2$. As in Example 5, consider the following RUFs:

$$
\mu_1 = \begin{cases} 
    xzw \text{ with probability } 1/2; & \text{and} \\
    yzxw \text{ with probability } 1/2
\end{cases} \quad \text{and} \quad
\mu_2 = \begin{cases} 
    xyzw \text{ with probability } 1/2; & \text{and} \\
    yxzw \text{ with probability } 1/2
\end{cases}
$$

Let $\rho_1$ and $\rho_2$ be the RCRs over complete menus of acts represented by $\mu_1$ and $\mu_2$ respectively. Let $P$ be the RCR over alternatives that is represented by both $\mu_1$ and $\mu_2$. Given the above discussion about the relationship between RCRs over complete menus of acts and RCRs over alternatives, Example 5 demonstrates that $\rho_1^D(f) = \rho_2^D(f)$ whenever $D$ is the complete menu generated by a menu of constant acts. However, consider the complete menu

$$
D := \{f_1, f_2, f_3, f_4\} := \{(x,z), (x,w), (y,z), (y,w)\}.
$$

The RCR $\rho_1$ selects $f_2$ and $f_3$ with equal probability from $D$ while $\rho_2$ selects $f_1$ and $f_4$ with equal probability. Therefore, $\mu_1$ and $\mu_2$ do not represent the same RCR over complete menus of acts.

To understand why the menu $D$ suffices to distinguish $\mu_1$ and $\mu_2$, consider the event $E \subseteq \mathcal{R}$ defined by $E := \{R \in \mathcal{R} \mid xRy \text{ and } zRw\}$. RCRs over alternatives are incapable of identifying the probability of this event. But for the RCRs $\rho_1$ and $\rho_2$ over
complete menus of acts, it must be that \( \mu_1(E) = \rho_1^D(f_1) \) and \( \mu_2(E) = \rho_2^D(f_1) \). ⊓⊔

### 3.4 Characterizing a Unique RUF Representation

Built into the definition of \( \rho \) are the so-called normalization equalities: \( \sum_{f \in D} \rho^D(f) = 1 \) for all menus \( D \in \mathcal{D} \). If \( D = \{f_1, \ldots, f_k\} \), the normalization equality for \( D \) may be viewed as a constraint on the act-menu pairs \((f_1, D), \ldots, (f_k, D)\). In particular, because \( D \) is complete it is clear that for each \( R \in \mathcal{R} \) there exists a unique act-menu pair \((f_i, D)\) from this list for which \( f_i \) is \( R \)-optimal in \( D \). Definitions 18 and 19 each utilize the concept of cycles across act-menu pairs to expand upon the normalization equalities.

**Definition 17.** A cycle across \( k \geq 2 \) act-menu pairs \((f_1, D_1), \ldots, (f_k, D_k)\) is a collection of \( m \geq 2 \) alternatives \( x_1 \succ x_2 \succ \cdots \succ x_m \succ x_1 \), where \( x_\ell \succ x_{\ell+1} \) is defined as follows:

1. \( x_\ell \neq x_{\ell+1} \); and
2. \( x_\ell = f_i(s) \) and \( x_{\ell+1} \in D_i(s) \) for some \( s \in S \) and \( i \in \{1, \ldots, k\} \).

**Definition 18.** A collection of \( k \geq 2 \) act-menu pairs \((f_1, D_1), \ldots, (f_k, D_k)\) is called disjoint if there is a cycle across the act-menu pairs \((f_i, D_i)\) and \((f_j, D_j)\) for any \( i \neq j \).

Notice that (unlike in the normalization equalities) there is no requirement that \( D_i = D_j \) for \( i \neq j \). Disjoint collections of act-menu pairs are best understood through the following simple lemma:

**Lemma 6.** A collection of act-menu pairs \((f_1, D_1), \ldots, (f_k, D_k)\) is disjoint if and only if the sets \( \{N(f_i, D_i)\}_{i=1}^k \) defined by

\[
N(f_i, D_i) := \{R \in \mathcal{R} \mid f_i \text{ is } R\text{-dominant in } D_i\}
\]

are themselves disjoint in \( \mathcal{R} \).
Proof of the “only if” direction. Let \((f_1, D_1), \ldots, (f_k, D_k)\) be disjoint. By definition, there is a cycle across any distinct pairs \((f_i, D_i)\) and \((f_j, D_j)\). Upon examination, this cycle is also a preference cycle for any \(R \in N(f_i, D_i) \cap N(f_j, D_j)\). Therefore, no such \(R\) can exist and \(N(f_i, D_i) \cap N(f_j, D_j) = \emptyset\). 

The proof of the converse direction is provided in Appendix C.1.1. The next definition identifies a subclass of disjoint-menu act pairs.

Definition 19. A collection of \(k \geq 2\) act-menu pairs \((f_1, D_1), \ldots, (f_k, D_k)\) is partial if

1. It is disjoint; and

2. There is a cycle across any \(k\) act-menu pairs \((f'_1, D_1), \ldots, (f'_k, D_k)\) for which \(f'_i \in (D_i \setminus f_i)\) for all \(i = 1, \ldots, k\) and \(\{(f_i, D_i)\}_{i=1}^k \cap \{(f'_i, D_i)\}_{i=1}^k = \emptyset\).

As the name suggests, partial lists of act-menu pairs serve to partition \(\mathcal{R}\).

Lemma 7. A collection of act-menu pairs \((f_1, D_1), \ldots, (f_k, D_k)\) is partial if and only if the sets \(\{N(f_i, D_i)\}_{i=1}^k\) defined by

\[
N(f_i, D_i) := \{R \in \mathcal{R} \mid f_i \text{ is } R\text{-dominant in } D_i\}
\]

partition \(\mathcal{R}\)—that is, \(N(f_i, D_i) \cap N(f_j, D_j) = \emptyset\) for all \(i \neq j\) and \(\bigcup_{i=1}^k N(f_i, D_i) = \mathcal{R}\) (empty sets are permitted in the partition).

Proof of the “only if” direction. Let \((f_1, D_1), \ldots, (f_k, D_k)\) be partial. For any \(i \neq j\), \(N(f_i, D_i) \cap N(f_j, D_j) = \emptyset\) by Lemma 6. Next, fix a relation \(R \in \mathcal{R}\). From Lemma 5, each menu \(D_i\) contains an \(R\)-dominant act \(f' \in D_i\). However, if \(f'_i \neq f_i\) for all \(i = 1, \ldots, k\), then there is a cycle across the act-menu pairs \((f'_1, D_1), \ldots, (f'_k, D_k)\). By inspection, this cycle is also a preference cycle for the relation \(R\). It follows that for some \(i\) the act \(f_i\) must be \(R\)-dominant in \(D_i\), and therefore that \(\bigcup_{i=1}^k N(f_i, D_i) = \mathcal{R}\). 

43
The converse direction is again established in Appendix C.1.1 and is useful in the proof of Theorem 2.

The following property can be viewed as an extension of the normalization equalities to all partitional collections of act-menu pairs.\textsuperscript{3}

**Definition 20** (Partition Consistency (PC)). A RCR $\rho$ over complete menus of acts satisfies *partition consistency (PC)* if

$$\sum_{i=1}^{k} \rho^{D_i}(f_i) = 1$$

whenever $(f_1, D_1), \ldots, (f_k, D_k)$ are partitional.

Partition consistency turns out to be sufficient for characterizing the existence of a unique RUF representation when there are same number of states as there are consequences.

**Theorem 2.** Let $|X| = N = |S|$ and let $\rho$ be a RCR on $\mathcal{D}$. There exists a unique $\mu \in \Delta(\mathcal{R})$ that represents $\rho$ if and only if $\rho$ satisfies PC.

The proof of Theorem 2 begins by defining a collection of sets

$$\mathcal{A} \subseteq 2^{\mathcal{R}} := \{A \in 2^{\mathcal{R}} \mid A = N(f, D) \text{ for some } (f, D) \in \mathcal{F} \times \mathcal{D}\}$$

and showing that $\mathcal{A}$ is a semi-algebra. A set function $\nu$ is then defined on $\mathcal{A}$ by $\nu(A) := \rho^D(f)$, where $(f, D)$ is the act-menu pair that defines $A$. After checking that $\nu$ is well-defined, the generalized normalization equalities are used to show that $\sum_{i=1}^{n} \nu(A_n) = \nu(A)$ whenever $A_1, \ldots, A_n$ are disjoint sets in $\mathcal{A}$ with union $A \in \mathcal{A}$. A standard extension theorem then delivers a unique measure $\mu$ that extends $\nu$ to $2^{\mathcal{R}}$.

\textsuperscript{3}Technically, it does not nest the normalization equalities because the degenerate pair $(f, \{f\})$ is not partitional.
3.5 How Many States Suffice for Uniqueness?

Let $|X| = N$ and $|S| = M$. Suppose that a RCR $\rho$ over complete menus of acts can be represented by a RUF $\mu$. Theorem 2 shows that $\mu$ is unique if $M = N$.

It is immediate upon inspection that the unique characterization result in Theorem 2 will continue to hold whenever $M > N$. In fact, the result also holds when $M = N - 1$.

**Corollary 5.** When $M \geq N - 1$, it remains true that a RCR $\rho$ over complete menus of acts has a unique RUF representation $\mu$ if and only if

$$\sum_{i=1}^{k} \rho^{D_i}(f_i) = 1$$

whenever $(f_1, D_1), \ldots, (f_k, D_k)$ are partitional.

The only non-trivial part of this corollary is demonstrating that the set $\mathcal{A}$ defined in the proof of Theorem 2 remains a semi-algebra when $M = N - 1$. This is proven in the appendix.

When $M = N - 2$, $\mathcal{A}$ is no longer an algebra. This can be seen by directly examining the case of $N = 3$ and $M = 1$, which is equivalent to a RCR over three alternatives. Let $X = \{x, y, z\}$. Then

$$\mathcal{A} = \left\{ \emptyset, R, \{xyz, xzy\}, \{yxz, yzx\}, \{zxy, zyx\}, \{xyz, yxz\}, \{zyx, zxy\} \right\}.$$

In addition, PC no longer suffices to guarantee the existence of a random utility representation. When $N = 3$ and $M = 1$, Block and Marschak (1960) show that regularity is both necessary and sufficient for a RCR to be represented by a RUF. However, the only collections of act-menu pairs that are partitional are the normalization equalities.
So PC does not imply regularity when \( M = N - 2 \).

That said, if a RUF does represent an RCR when \( M = N - 2 \), it is unique. In particular, since a RCR over complete menus of acts is equivalent to a RCR over alternatives when \( N = 3 \) and \( K = 1 \), this shows that RUF representations for RCRs over alternatives are unique when there are only three alternatives.

**Proposition 6.** Let \( M = N - 2 \). If \( \mu \) and \( \nu \) both represent the same RCR \( \rho \), then \( \mu = \nu \).

The proof of Proposition 6 provides an explicit procedure that can elicit from \( \rho \) the precise probability that a RUF \( \mu \) must assign to any single relation \( R \in \mathcal{R} \). To understand how this procedure works, return to the setting of Example 6 where \( X = \{x, y, z, w\} \) and \( S = \{s_1, s_2\} \). Consider the relation \( R = xyzw \). To find \( \mu(\{R\}) \), first construct the complete menu \( D_R \) that satisfies

\[
D_R(s_1) = X, \\
D_R(s_2) = \{y, z, w\}.
\]

Then form the complete menu \( D'_R \) where \( D'_R(s_1) = D_R(s_1) \) and \( D'_R(s_2) = \{z, w\} \). Let \( f_R := (x, z) \). The probability \( \mu(\{R\}) \) is now defined to be

\[
\mu(\{R\}) := \rho^{D'_R}(f_R) - \rho^{D_R}(f_R).
\]

Intuitively, this works because

\[
N(D'_R, f_R) = \{xyzw, xzyw, xzwy\} \quad \text{and} \quad N(D_R, f_R) = \{xzyw, xzwy\}.
\]

It remains true in the general procedure that only \( D_R, D'_R \), and \( f_R \) (appropriately defined) are used to construct \( \mu(\{R\}) \). Therefore, the data provided by \( \rho \) is sufficiently rich that eliciting the implicit probability assigned to a single ordinal preference requires
only two data points.

I conjecture, but have been unable to prove, that this elicitation procedure can
be used to show that PC and regularity together are necessary and sufficient for the
existence of a unique random-utility representation when $M = N - 2$.

Finally, when $M < N - 2$ Example 5 (which illustrates the case of $N = 4$ and
$K = 1$) shows that uniqueness can fail. In fact when $M = N - 3$ uniqueness can fail
for any $N$:

**Proposition 7.** Let $M = N - 3$. There exist RUFs $\mu_1$ and $\mu_2$ with disjoint support
that represent the same RCR over complete menus of acts.

The proof is by induction. It is instructive to consider the following example:

**Example 7.** Let $X = \{x_1, \ldots, x_5\}$ and $M = 2$. Consider the RUFs

\[
\mu_1 := \begin{cases} 
  x_1 x_2 x_5 x_3 x_4 & \text{with prob. } 1/4 \\
  x_2 x_1 x_5 x_4 x_3 & \text{with prob. } 1/4 \\
  x_1 x_2 x_4 x_3 x_5 & \text{with prob. } 1/4 \\
  x_2 x_1 x_3 x_4 x_5 & \text{with prob. } 1/4 \\
\end{cases}
\quad \text{and} \quad
\mu_2 := \begin{cases} 
  x_1 x_2 x_5 x_4 x_3 & \text{with prob. } 1/4 \\
  x_2 x_1 x_5 x_3 x_4 & \text{with prob. } 1/4 \\
  x_1 x_2 x_3 x_4 x_5 & \text{with prob. } 1/4 \\
  x_2 x_1 x_4 x_3 x_5 & \text{with prob. } 1/4. \\
\end{cases}
\]

One can verify that these RUFs represent the same RCR over complete menus of acts.
The change from Example 5 to this example is a special case of the move from $N$ to
$N + 1$ in the induction argument for Proposition 7. \Diamond

Together, the results in this section establish that $N - 2$ is a tight lower bound
on the number of states needed to guarantee that RUF representations for RCRs over
complete menus of acts are unique.
3.6 Discussion and Extensions

3.6.1 Enriching the Dataset

Up to this point, random choice rules have been defined only over complete menus of acts. Both the data and the representation are silent about what happens off these menus. If an observer happens to have data for all menus of acts, she may of course choose to ignore incomplete menus and apply Theorem 2 to obtain a RUF.

When incomplete menus are available, fixing a distribution of ordinal preferences will no longer completely determine stochastic choice—but it will impose some restrictions. For a simple example, let \( X = \{x, y\} \) and \( S = \{s_1, s_2\} \). Let \( D_1 = \{(x, y), (y, x), (y, y)\} \) and \( D_2 = \{(x, x), (x, y), (y, x), (y, y)\} \). Suppose an observer knows that \( x \) is preferred to \( y \) with probability 2/3. For her data to be consistent with this representation, certainly \((x, x)\) should be chosen from the (complete) menu \( D_2 \) with probability 2/3. In addition, \((y, y)\) should be chosen from the (incomplete) menu \( D_1 \) with probability 1/3. But it is not clear how frequently \((x, y)\) should be chosen from \( D_1 \). This type of reasoning motivates the following definition.

**Definition 21.** A RCR \( \rho \) defined on \( 2^F \setminus \{\emptyset\} \) is constrained by a RUF \( \mu \in \Delta(\mathcal{R}) \) if \( \rho^D(f) \) is at least

\[
\mu(\{R \in \mathcal{R} \mid f \text{ is } R\text{-dominant in } D\})
\]

and at most

\[
1 - \mu(\{R \in \mathcal{R} \mid f \text{ is } R\text{-dominantated in } D\})
\]

for all \( f \in \mathcal{F} \) and \( D \in 2^F \setminus \{\emptyset\} \).

If \( \rho \) is constrained by a RUF, the bounds in Definition 21 will precisely determine choice over a complete menu (therefore, a RCR can be constrained by at most one RUF). For a menu such \( \{(x, y), (y, x)\} \), they will not constrain choice at all. This is
the best one can hope for given only a distribution on ordinal preferences.

The following simple result is a partial step towards characterizing when a RCR is constrained by an RUF. It shows that imposing PC on complete menus and regularity across all menus suffices to establish the lower bound in Definition 21.

**Proposition 8.** Let \( \rho \) be a RCR on \( 2^F \setminus \{ \emptyset \} \). If \( \rho \) satisfies PC on complete menus and regularity on all menus, then

\[
\rho^D(f) \geq \mu(\{ R \in \mathcal{R} \mid f \text{ is } R\text{-dominant in } D \})
\]

for all \((f, D) \in \mathcal{F} \times (2^F \setminus \{ \emptyset \})\).

**Proof.** Fix an act-menu pair \((f, D) \in \mathcal{F} \times (2^F \setminus \{ \emptyset \})\). Let \( D' \) be the complete menu generated by \( D \). Notice that \( N(f, D) = N(f, D') \). Let \( \mu \) be the measure that represents \( \rho \) when it restricted to complete menus (which exists by Theorem 2 and the assumption that \( \rho \) satisfies PC on complete menus). By regularity,

\[
\rho^D(f) \geq \rho^{D'}(f) = \mu(N(f, D')) = \mu(N(f, D)).
\]

#### 3.6.2 The Role of the State Space

The goal of this paper is to axiomatize a unique random utility representation for stochastic choice on a finite data set. The previous sections achieved this goal by introducing a state space and providing a representing in which acts are chosen according the probability they are dominant in a complete menu. When evaluating this probability, subjective beliefs over the likelihood of various states occurring are irrelevant. Nonetheless, if an analyst has stochastic choice data for all menus (rather than just complete menus), it would be natural to study a representation featuring a joint measure over preferences and beliefs. This exercise immediately encounters several problems: for instance, the joint measure will not be unique if the state space remains
finite. A relevant paper is Lu (2013), which considers stochastic choice over menus of Anscombe-Aumann acts and axiomatizes a representation in which (cardinal) utility is fixed but beliefs over the state space are random.

Stepping back, in this paper $S$ is an arbitrary finite set and need not represent “states.” It is the Cartesian structure of acts that facilitates a representation in which DMs choose dominant options. The set $S$ might instead represent time periods or some other structure in which DMs are presented with a bundle of choices organized by coordinates. This freedom expands the types of data that could conceivably be used to estimate unique random utility functions.
References


Appendix A

Appendix to Chapter 1

A.1 Proofs

A.1.1 Proof of Corollary 1

We begin by proving Theorem 1(B). To show necessity, let $\succsim$ be represented by $(\mu, \alpha)$ for $\mu \in \Delta(R)$ and $\alpha \in (1/2, 1]$. Obviously, $\succsim$ is reflexive: every $R \in \mathcal{R}$ is reflexive. Every $R \in \mathcal{R}$ is also antisymmetric, so if $\mu(\{R \mid xRy\}) \geq \alpha$ then $\mu(\{R \mid yRx\}) \leq 1 - \alpha < 1/2$, showing that $\succsim$ is antisymmetric.

For sufficiency, let $\succsim$ be reflexive and antisymmetric. Let $\varepsilon = [2(\binom{N}{N-2})^{-1}$ and construct a hypothetical RCR $\rho$ on $X$ as follows. On binary sets, define

$$\rho^{\{x,y\}}(x) = \begin{cases} 
1/2 + \varepsilon & \text{if } x \succsim y; \\
1/2 - \varepsilon & \text{if } y \succsim x; \\
1/2 & \text{otherwise.}
\end{cases}$$

For all other sets $D \in \mathcal{D}$, let $\rho^D(x) = 1/|D|$ for all $x \in D$. Recall that the BM
inequalities demand

\[ BM(\rho, D, x) := \sum_{C : D \subseteq C} (-1)^{|C \setminus D|} \rho^C(x) \geq 0 \quad \forall (D, x) \in D \times X. \]

The hypothetical \( \rho \) constructed above satisfies these inequalities. To see that, note for \(|D| \geq 3\), \( \rho^D(x) = m(\{R : xRy\}) \), where \( m \) is the uniform measure on \( \mathcal{R} \). Thus, letting \( \rho_m \) denote the RCR that maximizes the RUF \( m \), we have \( BM(\rho, D, x) = BM(\rho_m, D, x) \geq 0 \) whenever \(|D| \geq 3\). When \(|D| = 2\),

\[
BM(\rho, D, x) \geq 1/2 - \varepsilon + \sum_{C : D \subseteq C} (-1)^{|C \setminus D|} \rho^C(x) \\
= 1/2 - \varepsilon + \sum_{i=1}^{N-2} \frac{(-1)^i}{i+2} \binom{N-2}{i} \\
= \left[ 2 \left( \frac{N}{N-2} \right) \right]^{-1} - \varepsilon \\
= 0.
\]

It remains to check \( BM(\rho, D, x) \geq 0 \) for \(|D| = 1\). However, this turns out to be a consequence of the fact that \( BM(\rho, D, x) \geq 0 \) for \(|D| = 2\), as shown by the following lemma from Chapter 3.

**Lemma 8.** Let \( \rho \) be an RCR on a set \( X \) with \(|X| = N\). If \( BM(\rho, D, x) \geq 0 \) for all pairs \((D, x)\) with \(|D| = 2\), then \( BM(\rho, \{x\}, x) \geq 0 \).

Applying this result, it follows that the hypothetical \( \rho \) constructed using \( \succeq \) satisfies the BM-inequalities. Furthermore, it is obvious from the construction of \( \rho \) that \( \succeq \) and \( \rho \) are compatible on binary sets. Therefore, by Lemma 1, \( \succeq \) and \( \rho \) are consistent, which by Definition 2 entails the existence of a regular confidence model representing \( \succeq \).

We now turn to proving Theorem 1(A). Necessity is obvious. Given a reflexive preference \( \succeq \), construct the following hypothetical RCR: for \(|D| \neq 2\), let \( \rho^D(x) = 1/|D| \).
for all $x \in D$. For binary sets, once again define $\varepsilon = [2(N^2)]^{-1}$ and set

$$
\rho^{\{x,y\}}(x) = \begin{cases} 
1/2 + \varepsilon & \text{if } x \succcurlyeq y \text{ and } \neg(y \succcurlyeq x); \\
1/2 - \varepsilon & \text{if } y \succcurlyeq x \text{ and } \neg(x \succcurlyeq y); \\
1/2 & \text{otherwise.}
\end{cases}
$$

The same arguments used in the proof of Theorem 1(A) show that $\rho$ satisfies the BM-inequalities, and is therefore represented by a RUF $\mu \in \Delta(R)$.

Write $x \sim y$ if $x \succcurlyeq y$ and $y \succcurlyeq x$. Let $I$ denote the collection of binary sets $\{x, y\}$ with $x \sim y$ and let $M$ denote the number of indifferent pairs in $I$. For each such pair, define a set $S_{x,y} \subseteq W$ consisting of all the weak orders $W$ for which (1) $xWy$ and $yWx$ and (2) for all binary sets $\{x',y'\} \neq \{x,y\}$, either $\neg(x'Wy')$ or $\neg(y'Wx')$. In words, $S_{x,y}$ consists of all weak orders for which $x$ and $y$ are the only distinct alternatives that are indifferent to each other.

For each $\{x, y\} \in I$, let $\nu_{x,y}$ denote the uniform measure on $S_{x,y}$. Choose your favorite $\lambda \in (0, 1)$ and define a measure $\mu' \in \Delta(W)$ by

$$
\mu' = (1 - \lambda)\mu + \sum_{\{x,y\} \in I} \frac{\lambda}{M} \nu_{x,y}.
$$

Observe that if $x'$ and $y'$ are unranked according to $\succcurlyeq$, then

$$
\mu(\{W \mid x'Wy'\}) = \nu_{x,y}(\{W \mid x'Wy'\}) = 1/2
$$

for any $\{x, y\} \in I$, showing that $\mu'(\{W \mid x'Wy'\}) = 1/2$. If $x' \succcurlyeq y'$ but $\neg(y' \succcurlyeq x')$, then

$$
\mu(\{W \mid x'Wy'\}) = 1/2 + \varepsilon \text{ while } \nu_{x,y}(\{W \mid x'Wy'\}) = 1/2
$$

again for any $\{x, y\} \in I$. Therefore, $\mu'(\{W \mid x'Wy'\}) = \lambda(1/2 + \varepsilon) + (1 - \lambda)(1/2)$. Similarly, $\mu'(\{W \mid y'Wx'\}) = \lambda(1/2 - \varepsilon) + (1 - \lambda)(1/2)$. 55
Finally, if \( x' \sim y' \) then we have

\[
\mu(\{W \mid x'Wy'\}) = \nu_{x,y}(\{W \mid x'Wy'\}) = 1/2
\]

provided that \( \{x, y\} \neq \{x', y'\} \), and \( \nu_{x', y'}(\{W \mid x'Wy'\}) = 1 \). Therefore, \( \mu'(\{W \mid x'Wy'\}) = (1 - \lambda/M)(1/2) + \lambda/M \). If we set

\[
\alpha := \min \{\lambda(1/2 + \varepsilon) + (1 - \lambda)(1/2), (1 - \lambda/M)(1/2) + \lambda/M\} > 1/2
\]

then the preceding analysis shows that \((\mu', \alpha)\) is a confidence model for \( \succeq \).

### A.1.2 Proof of Lemma 2

Assume that a generalized RCR \( \rho \) satisfying the GP-inequalities and a reflexive preference \( \succeq \) jointly satisfy generalized binary compatibility. Let \( S = \{(x, y) \mid x \succeq y\} \) and define

\[
\alpha := \min_{S} 1 - \rho(\{y\}, \{x, y\}).
\]

Because \( \succeq \) is reflexive, \( S \) is nonempty. By the definition of generalized binary compatibility, \( \alpha > 1/2 \). Finally, letting \( \mu \in \Delta(W) \) represent \( \rho \), we have

\[
\mu(\{W \mid xWy\}) = 1 - \rho(\{y\}, \{x, y\}) \geq \alpha
\]

if and only if \( x \succeq y \).

The converse follows immediately from the definitions of the respective representations.

### A.1.3 Proof of Lemma 3

Let \( \succeq \) be a regular confidence preference on \( \Delta(X) \). We begin by establishing a simple necessary condition for \( \succeq \) to have a regular confidence representation:
Claim: If \( \neg(p \succ q) \), then \( \exists \lambda \in (0, 1) \) such that \( \neg(\lambda p + (1 - \lambda)q \succ q) \).

Proof. Let \((\mu, \alpha)\) be the regular confidence representation for \( \succ \). Let \( \rho \) be the unique RCR on \( \Delta(X) \) that maximizes \( \mu \). Then we have \( \rho^{[p\succ q]}(p) < \alpha \). Let \( p\alpha q := \lambda p + (1 - \lambda)q \).

By mixture continuity of \( \rho \), there exits \( \lambda \in (0, 1) \) such that \( \rho^{[p\alpha q]}(p\alpha q) < \alpha \). The conclusion follows.

Let \( W := \{(p, q) \in \Delta(X)^2 \mid p \succ q\} \) and define \( \alpha = \inf\{\rho^{[p\succ q]}(p) \mid (p, q) \in W\} \).

From this definition and binary compatibility, we have

\[
\rho^{[p\succ q]}(p) \geq \alpha \geq \rho^{[p'\succ q']}(p')
\]

for all pairs of lotteries \((p, q)\) and \((p', q')\) with \( p \succ q \) and \( \neg(p' \succ q') \). We claim that in fact the latter inequality holds strictly. If \( \alpha \) is attained on \( W \) (e.g., in the trivial case of \( \alpha = 1 \)) this follows directly from binary compatibility. In general, additional arguments are needed. So, let \( \alpha < 1 \) and assume for a contradiction that there exists a pair of lotteries \((p', q')\) with \( \neg(p' \succ q') \) but \( \rho^{[p'\succ q']}\}(p') = \alpha \). By the claim above, there exists some \( \lambda \in (0, 1) \) such that \( \neg(\lambda p' + (1 - \lambda)q' \succ q') \). By linearity of \( \rho \), we have

\[
\rho^{\lambda(p'\succ q')} + (1 - \lambda)(q') = \lambda \rho^{[p'\succ q']}(p') + (1 - \lambda)\rho^{[q']}(q')
\]

\[
= \lambda \cdot \alpha + (1 - \lambda) \cdot 1 =: \beta > \alpha.
\]

By the definition of \( \alpha \), however, there exists some pair \((\hat{p}, \hat{q})\) with \( \hat{p} \succ \hat{q} \) and \( \rho^{[\hat{p}\succ \hat{q}]}(\hat{p}) < \beta \), which contradicts binary compatibility. Therefore, Equation (A.1) becomes:

\[
\rho^{[p\succ q]}(p) \geq \alpha > \rho^{[p'\succ q']}(p').
\]

(A.2)

Since \( \rho \) satisfies GP1–4 there exists a unique \( \nu \) that maximizes \( \rho \). By definition, this means that for any lotteries \( p \) and \( q \) we have \( \nu(\{u \mid u \cdot p \geq u \cdot q\}) = \rho^{[p\succ q]}(p) \). By Equation (A.2), \( \rho^{[p\succ q]}(p) \geq \alpha \) if and only if \( p \succ q \) and thus \((\nu, \alpha)\) is a confidence model.
for $\gtrsim$.

### A.1.4 Proof of Theorem 1

Necessity of the axioms is obvious: $\mu(\{u \mid u \cdot p \geq u \cdot p\}) = 1$ so $p \gtrsim p$ for any $p$. For Independence, let $p' = \lambda p + (1 - \lambda)q$ and $q' = \lambda q + (1 - \lambda)r$. Clearly $\mu(\{u \mid u \cdot p' \geq u \cdot q'\}) = \mu(\{u \mid u \cdot p \geq u \cdot q\})$, which implies Independence.

The proof of sufficiency rests on the following lemma.

**Lemma 9.** Let $p \neq q \in \mathcal{L}$ and fix $\alpha \in (1/2, 1)$. There exists a finitely-additive probability measure $\mu_{pq} \in \Delta(\mathcal{U})$ that satisfies

$$
\mu_{pq}(\{u \in \mathcal{U} \mid u \cdot p' \geq u \cdot q'\}) = \alpha \text{ if } p' = \lambda p + (1 - \lambda)w \text{ and } q' = \lambda q + (1 - \lambda)w
$$

for $\lambda \in (0, 1]$ and $w \in \Delta(x)$, and $\mu_{pq}(\{u \in \mathcal{U} \mid u \cdot p' \geq u \cdot q'\}) = 1/2$ for any other $p' \neq q'$.

**Proof.** Let $\mathcal{A} \subseteq 2^\mathcal{U}$ consist of the empty set, $\mathcal{U}$, and in addition all subsets of $\mathcal{U}$ that can be written in the form $A_{p', q'} := \{u \in \mathcal{U} \mid u \cdot p' \geq u \cdot q'\}$ for some $(p', q') \in \mathcal{L}^2$ with $p' \neq q'$. Define a set function $\nu: \mathcal{A} \rightarrow [0, 1]$ as follows. Let $\nu(\mathcal{U}) = 1$ and $\nu(\emptyset) = 0$. Let $\nu(A_{p', q'}) = \alpha$ if $p' = \lambda p + (1 - \lambda)w$ and $q' = \lambda q + (1 - \lambda)w$ for some $\lambda \in (0, 1]$ and $w \in \Delta(x)$. For any other pair $p' \neq q'$, let $\nu(A_{p', q'}) = 1/2$.

Let $\{A_1, \ldots, A_n\}$ and $\{B_1, \ldots, B_m\}$ be finite collections of sets in $\mathcal{A}$. No nontrivial sets in $\mathcal{A}$ overlap. The set $A_{p, q}$ must be covered by $\mathcal{U}$ or at least two sets with measure $1/2$. Therefore,

$$
\sum_{i=1}^n 1_{A_i} \leq \sum_{j=1}^m 1_{B_j} \quad \Rightarrow \quad \sum_{i=1}^n \nu(A_i) \leq \sum_{j=1}^m \nu(B_j).
$$

The Borel $\sigma$-algebra on $S_{r,s}$ contains $\mathcal{A}$, so the desired measure $\mu_{pq}$ exists by Theorem 3.2.10 in Rao and Rao (1983).
With this result in hand, let \( \succsim \) be a preference on \( \mathcal{L} \) that is reflexive and satisfies Independence. Let \( \mathcal{P} \subseteq \mathcal{L}^2 \) consist of all pair of lotteries \( (p, q) \in \mathcal{L}^2 \) with \( p \succsim q \) but \( p \neq q \). For each such pair, construct a measure \( \mu_{pq} \) from the lemma above using a common value \( \alpha' \) in every construction.

Let \( M = |\mathcal{P}| \). Define a measure \( \mu \in \Delta(\mathcal{U}) \) by

\[
\mu = \sum_{(p,q) \in \mathcal{P}} (1/M)\mu_{pq}
\]

and let \( \alpha = (1/M)\alpha' + ((M - 1)/M)(1/2) \). By Independence, if \( (p', q') \notin \mathcal{P} \), then \( \mu_{pq}(\{u | u \cdot p' \geq u \cdot q'\}) = 1/2 \) for all \( (p, q) \in \mathcal{P} \). On the other hand, if \( (p', q') \in \mathcal{P} \) then \( \mu_{p'q'}(\{u | u \cdot p' \geq u \cdot q'\}) = \alpha' \), which implies that \( \mu(\{u | u \cdot p' \geq u \cdot q'\}) \) is at least \( \alpha \). It follows that \((\mu, \alpha)\) is a confidence representation for \( \succsim \).

A.2 Confidence Models and Continuity

Let \( \mathcal{L} = \Delta(X) \) and suppose \( \succsim \) is represented by a regular confidence model \((\mu, \alpha)\). In the proof of Lemma 3, this is shown to imply that \( \succsim \) satisfies a mild continuity property. However, consider the following standard continuity axiom:

**Continuity** For any \( p \in \mathcal{L} \), the sets \( \{q \in \mathcal{L} | q \succsim p\} \) and \( \{q \in \mathcal{L} | p \succsim q\} \) are closed.

In Section 1.3, \( \mu \) is required only to be finitely additive. If \( \mu \) fails to be countably additive, \( \succsim \) can fail to satisfy continuity. This is somewhat counterintuitive; it occurs even though each \( u \) in the support of \( \mu \) represents an EU preference that certainly does satisfy continuity. I thank an anonymous referee for suggesting the following example: let \( \mu \) be the finitely-additive measure constructed in Example S1 of the supplement to Gul and Pesendorfer (2006). Loosely, this measure supports a single utility function \( u \) with a uniform tie-breaking rule. If \( p \) is indifferent to \( q \) under \( u \), then a sequence of
lotteries $p_n \rightarrow p$ can be found with each $p_n$ strictly preferred to $q$. This will violate continuity when a confidence model is constructed by pairing $\mu$ with any $\alpha \in (1/2, 1]$.

If $\mu$ is required to be countably additive, the RCR $\rho$ represented by $\mu$ is continuous in the Hausdorff topology on $\Delta(X)$ and this problem disappears. As stated in the text, I conjecture that reflexivity, antisymmetry, Independence and Continuity are necessary and sufficient for a preference to have a regular confidence representation featuring a countably additive measure. Conceivably, this could be proved by adapting the proof of Theorem 1 and constructing a hypothetical RCR that is compatible on binary sets with the preference, and then applying Lemma 3. The hypothetical RCR would need to be continuous (rather than just mixture continuous) according the definition in Gul and Pesendorfer (2006). Alternatively or in addition, one could possibly weaken continuity on $\succeq$ to a form of mixture continuity, construct a hypothetical mixture-continuous RCR, and obtain a finitely-additive confidence representation. However, constructing hypothetical RCRs over lotteries is substantially more complex than constructing them over finite domains.

The strategies described in the previous paragraph would generate regular confidence models. In the context of preferences on $\Delta(X)$, dropping regularity is more complicated than merely removing the antisymmetry axiom. Loosely, the reason is that for a confidence preference with representation $(\mu, \alpha)$ to have non-trivial indifference, the measure $\mu$ must assign positive probability to a set that does not have full dimension in $\mathcal{U}$. This cannot occur for an arbitrary collection of lower-dimensional sets, and the restrictions may vary for the case of finitely- versus countably-additive measures.
Appendix B

Appendix to Chapter 2

B.1 Proof of Proposition 1

To complete the proof of this proposition, it remains to show that (1) $\Rightarrow$ (3). Let $\rho$ be a RCR satisfying regularity. As in the main text, let $\mathcal{A} \subseteq 2^\mathcal{R}$ denote the collection of sets whose measure is “identified” by $\rho$:

$$\mathcal{A} := \{A \subseteq \mathcal{R} \mid A = N(D, x) \text{ for some } (D, x) \in \mathcal{D} \times X\}.$$

and again define a monotonic set function $\hat{\nu}: \mathcal{A} \rightarrow [0, 1]$ by $\hat{\nu}(A) = \rho^D(x)$.

Next, extend $\hat{\nu}$ to a set function $\nu$ on $2^\mathcal{R}$ using the formula

$$\nu(E) := \max \sum_{i=1}^k \hat{\nu}(A_i)$$

where the maximum is taken over all collections of disjoint sets $A_1, \ldots, A_k$ in $\mathcal{A}$ that satisfy $\bigcup_{i=1}^k A_i \subseteq E$. (Set $\nu(\emptyset) = 0$.)

It is straightforward to verify that $\nu$ is monotonic with respect to set inclusion on $\mathcal{A}'$. However, it remains to be verified that $\nu$ agrees with $\hat{\nu}$ on $\mathcal{A}$ and that $\nu(\mathcal{R}) = 1$. The first claim holds because based on the structure of $\mathcal{A}$, the only collection of disjoint
sets in \( \mathcal{A} \) whose union is contained in a set \( B \in \mathcal{A} \) is the degenerate collection \( \{B\} \).

For the second claim, note that \( \nu(\mathcal{R}) \geq 1 \) by the normalization equalities. Therefore, it suffices to show that \( \sum_{i=1}^{k} \nu(A_i) \leq 1 \) for any collection \( A_1, \ldots, A_k \) of disjoint sets in \( \mathcal{A} \). Let \( N(D_i, x_i) \) define each set \( A_i \). Then

\[
\sum_{i=1}^{k} \nu(A_i) = \sum_{i=1}^{k} \rho^{D_i}(x_i).
\]

In order for \( \{N(D_i, x_i)\}_{i=1}^{k} \) to be pairwise disjoint, \( x_1, \ldots, x_k \) must be distinct elements of \( X \) and each set \( D_i \) must contain all of these elements. By regularity and the normalization equality for \( \{x_1, \ldots, x_k\} \),

\[
\sum_{i=1}^{k} \rho^{D_i}(x_i) \leq \sum_{i=1}^{k} \rho^{(x_1, \ldots, x_k)}(x_i) = 1.
\]

Finally, let \( E \) and \( F \subseteq \mathcal{R} \) be disjoint. By the definition of \( \nu \),

\[
\nu(E) = \sum_{i=1}^{I} \hat{\nu}(A_i) \quad \text{and} \quad \nu(F) = \sum_{j=1}^{J} \hat{\nu}(B_j)
\]

for some disjoint collections of sets \( \{A_1, \ldots, A_I\} \) and \( \{B_1, \ldots, B_J\} \) in \( \mathcal{A} \). Because \( E \cap F = \emptyset \), the collection \( \{A_1, \ldots, A_I, B_1, \ldots, B_J\} \) is disjoint. Furthermore, the union of this set is contained in \( E \cup F \). It follows that \( \nu \) is superadditive.

### B.2 An Extension Theorem for Superadditive Capacities

Let \( S \) be an arbitrary set and let \( \mathcal{A} \subseteq 2^S \) with \( \emptyset, S \in \mathcal{A} \). The following simple extension theorem is based on the proof of Proposition 1.

**Lemma 10.** Let \( \hat{\nu} \) be a set function mapping \( \mathcal{A} \) to \([0, 1]\) satisfying \( \hat{\nu}(\emptyset) = 0 \) and
\( \hat{\nu}(S) = 1 \). Suppose that \( \hat{\nu} \) is monotonic and strongly superadditive on \( \mathcal{A} \); that is

\[
\hat{\nu}(A) \geq \sum_{i=1}^{k} \hat{\nu}(A_i)
\]

whenever \( A_1, \ldots, A_k \) are disjoint and contained in \( A \). Then \( \hat{\nu} \) can be extended to a superadditive capacity \( \nu \) on \( 2^S \).

**Proof.** Define

\[
\nu(E) := \sup \sum_{i=1}^{k} \hat{\nu}(A_i)
\]

where the supremum is taken over all collections of disjoint sets \( A_1, \ldots, A_k \) in \( \mathcal{A} \). Since \( \hat{\nu} \) is strongly superadditive, \( \nu \) is well-defined. Clearly \( \nu \) is monotonic, and it is superadditive by the same logic as in the proof of Proposition 1 above. \( \blacksquare \)

### B.3 Proof of Proposition 3

Let \((D, x) \in D \times X\) with \( x \in D \) and \(|D| = N - K \). The BM inequality indexed by \((D, x)\) states that

\[
\sum_{C : D \subseteq C} (-1)^{|C \setminus D|} \rho^C(x) \geq 0.
\]

Denote \( X \setminus D \) as \( \{y_1, \ldots, y_K\} \). For each \( i = 1, \ldots, K \), define \( D_i := D \cup \{y_i\} \) and \( A_i := N(D_i, x) \). Suppose that \( \nu \) is a capacity representation for \( \rho \). Then

\[
\nu \left( \bigcup_{i=1}^{K} A_i \right) \leq \nu(N(D, x)) = \rho^D(x),
\]

and

\[
\sum_{\emptyset \neq I \subseteq 1, \ldots, K} (-1)^{|I|+1} \nu \left( \bigcap_{k \in I} A_k \right) = \sum_{C : D \subseteq C} (-1)^{|C \setminus D|+1} \rho^C(x).
\]
If the BM inequality indexed by \((D, x)\) is violated, it follows from these two equations that
\[
\nu\left(\bigcup_{i=1}^{K} A_i\right) < \sum_{\emptyset \neq I \subseteq \{1, \ldots, K\}} (-1)^{|I|+1} \nu\left(\bigcap_{k \in I} A_k\right)
\]
showing that \(\nu\) is not \(K\)-monotonic. Note that Proposition 2 is a special case of Proposition 3 with \(K = 2\).

### B.4 Proof of Corollary 4

Let \(\text{BM}(\rho, D, x)\) denote the BM inequality associated with \(\rho\) indexed by \((D, x) \in \mathcal{D} \times X\). The following lemma demonstrates that the BM inequalities indexed by \((\{x\}, x)\) are redundant.

**Lemma 11.** For any RCR \(\rho\), if \(\text{BM}(\rho, D, x) \geq 0\) for all pairs \((D, x)\) with \(|D| = 2\), then \(\text{BM}(\rho, \{x\}, x) \geq 0\).

**Proof.** Using the normalization equalities, the definition of the BM inequalities, and a fact about binomial sums

\[
\text{BM}(\rho, \{x\}, x) := \sum_{U: x \in U} (-1)^{|U|-1} \rho^U(x)
\]

\[
= 1 + \sum_{y \in S \setminus \{x\}} \text{BM}(\rho, \{x, y\}, y) + \sum_{i=1}^{N-1} (-1)^i \binom{N-1}{i} \]

\[
= \sum_{y \in S \setminus \{x\}} \text{BM}(\rho, \{x, y\}, y).
\]

\[\blacksquare\]

Suppose that a RCR can be represented by a capacity \(\nu\) that is monotonic of order \(N - 2\). By Proposition 3, this implies that all of the BM inequalities indexed by pairs \((D, x)\) with \(|D| \geq 2\) must hold. In light of the preceding lemma, in fact all the BM inequalities must hold, and the RCR can be represented by a RUF. The reverse direction is trivial since a probability measure is \(K\)-monotonic for any \(K\).
Appendix C

Appendix to Chapter 3

C.1 Proofs

C.1.1 Proofs for Lemmas 6 and 7

The “if” directions of Lemmas 6 and 7 rely on the following simple preference-extension argument.

**Lemma 12.** Let $\succeq$ be a binary relation on $X$. If $\succeq$ is reflexive and acyclic then there exists a linear partial order (i.e. a complete, antisymmetric, transitive relation) $\succeq'$ that extends $\succeq$—that is, $x \succeq y$ implies $x \succeq' y$.

**Proof.** By induction on $N = |X|$. The case $N = 1$ is trivial, so assume the result holds for $N - 1$ and let $\succeq$ be a reflexive, acyclic relation on $X$ with $|X| = N$. Since $X$ is finite, there exists an element $x \in X$ such that $\neg(x \succeq y)$ for all $y \neq x$ (otherwise, $\succeq$ would have a cycle). Let $X' = X \setminus \{x\}$ and let $\succeq' = \succeq |_{X\setminus\{x\}}$. By the IH, there exists a linear partial order $\succeq' = y_1y_2\ldots y_{N-1}$ on $X'$ that extends $\succeq'$. Then $\succeq = y_1y_2\ldots y_{N-1}x$ is a linear partial order on $X$ that extends $\succeq$. ■

To prove the “if” direction for Lemma 7, let $(f_1, D_1), \ldots, (f_k, D_k)$ be a collection of
act-menu pairs for which \( N(f_i, D_i) \cap N(f_j, D_j) = \emptyset \) for all \( i \neq j \) and \( \bigcup_{i=1}^{k} N(f_i, D_i) = \mathcal{R} \).

To show this collection is partitional in the sense of Definition 19, I first need to show there is a cycle on any two distinct act-menu pairs \((f_i, D_i)\) and \((f_j, D_j)\). Define a relation \( \succsim \) on \( X \) by \( x \succsim y \) iff \( x = f_{\ell}(s) \) and \( y \in D_{\ell}(s) \) for some \( \ell \in \{i, j\} \) and \( s \in S \).

Clearly \( \succsim \) is reflexive. If \( \succsim \) were acyclic, then by the preceding extension lemma there would exist a relation \( R \in N(f_i, D_i) \cap N(f_j, D_j) \), contradicting the assumption of this lemma. Therefore, \( \succsim \) must contain a preference cycle, which by construction is also a cycle on the act menu pairs \((f_i, D_i)\) and \((f_j, D_j)\). Note that this argument establishes the “if” direction for Lemma 6.

Next, I need to show there is a cycle across any \( k \) act-menu pairs \((f'_1, D_1), \ldots, (f'_k, D_k)\) with \( f'_i \in (D_i \setminus f_i) \) for all \( i = 1, \ldots, k \). Consider such a collection and define a relation \( \succsim \) by \( x \succsim y \) iff \( x = f'_{i}(s) \) and \( y \in D_{i}(s) \) for some \( i \in \{1, \ldots, k\} \) and \( s \in S \). Again, \( \succsim \) is reflexive. If it were acyclic, it could extended by the preceding lemma to a relation \( R \in \bigcap_{i=1}^{k} N(f'_i, D_i) \). This would contradict the statement that \( \bigcup_{i=1}^{k} N(f_i, D_i) = \mathcal{R} \), Therefore, \( \succsim \) must contain a preference cycle, and by construction this preference cycle is also a cycle on \((f'_1, D_1), \ldots, (f'_k, D_k)\). Together with the previous paragraph, this establishes the “if” direction for Lemma 7.

**C.1.2 Proof of Proposition 5**

**Part 1.** Let \( X = \{x_1, \ldots, x_N\} \) and let \( \mu \in \Delta(\mathcal{R}) \) have full support. Let \( a_1, \ldots, a_4 > 0 \) denote the probability that \( \mu \) assigns to the relations

\[
\begin{align*}
x_5, x_6, \ldots, x_n, x_1, x_2, x_3, x_4 \\
x_5, x_6, \ldots, x_n, x_2, x_1, x_4, x_3 \\
x_5, x_6, \ldots, x_n, x_1, x_2, x_4, x_3 \\
x_5, x_6, \ldots, x_n, x_2, x_1, x_3, x_4 \\
x_5, x_6, \ldots, x_n, x_2, x_1, x_3, x_4
\end{align*}
\]
respectively. Take \( \varepsilon < \min\{a_1, \ldots, a_4\} \). Construct a new measure \( \mu' \) by assigning probabilities
\[
a'_1 = a_1 - \varepsilon, \quad a'_2 = a_2 - \varepsilon, \quad a'_3 = a_3 + \varepsilon, \quad a'_4 = a_4 + \varepsilon
\]
to the four relations above and leaving the remaining probabilities from \( \mu \) unchanged. It straightforward to verify that \( \mu' \) represents exactly the same RCR as \( \mu \).

**Part 2.** Let \( \mathcal{P}^* \) denote the set of all RCRs on \( X \), and endow \( \mathcal{P}^* \) with the sup metric
\[
d(P, Q) := \sup_{(D,x)} P^D(x) - Q^D(x).
\]
Note that convergence under this metric is equivalent to pointwise convergence in menu-item pairs \((D, x)\) because \( X \) is finite. Endow \( \mathcal{P} \), the set of all RCRs that can be represented by some RUF, with the topology it inherits as a subset of \( \mathcal{P}^* \).

Take \( P \in \mathcal{P} \) and let \( P \) be represented by \( \mu \in \Delta(\mathcal{R}) \). By Part 1, there exists a sequence of measures \( \mu_n \to \mu \) with the property that for each \( \mu_n \), there exists a measure \( \mu'_n \neq \mu_n \) such that \( \mu_n \) and \( \mu'_n \) represent the same RCR \( P_n \). Clearly, \( P_n^D(x) \to P^D(x) \) for any pair \((D, x)\), which establishes that the set
\[
\{P \in \mathcal{P} \mid P \text{ can be represented by two distinct RUFs}\}
\]
is dense in \( \mathcal{P} \).

**C.1.3 Proof of Theorem 2**

Necessity is immediate in light of Lemma 7. For sufficiency, let \( \rho \) be a RCR over complete menus of acts that satisfies PC. Define a collection of sets \( \mathcal{A} \subseteq 2^\mathcal{R} \) by
\[
\mathcal{A} := \{A \subseteq \mathcal{R} \mid A = N(f, D) \text{ for some } (f, D) \in \mathcal{F} \times \mathcal{D}\}.
\]

**Lemma 13.** The collection \( \mathcal{A} \) is a semialgebra according the definition in Durrett.
(2010); that is,

1. If $A, B \in \mathcal{A}$ then $A \cap B \in \mathcal{A}$; and

2. If $A \in \mathcal{A}$ then $A^c$ is a finite disjoint union of sets in $\mathcal{A}$.

Proof. For Part 1, let $A = N(f, D)$ and $B = N(g, E)$. Let $x_1, x_2, \ldots, x_N$ be an enumeration of the $N$ elements in $X$ and let $h = (x_1, x_2, \ldots, x_N)$ be the corresponding act. Define the menu $D'$ by

$$D'(s_i) = \{x_i\} \cup \left[ \bigcup_{j \mid f(s_j) = x_i} D(s_j) \right]$$

and similarly define the menu $E'$. Note that $N(f, D) = N(h, D')$ and $N(g, E) = N(h, E')$. Now, let $C$ be the menu defined by $C(s_i) = D'(s_i) \cup E'(s_i)$. Then

$$N(f, D) \cap N(g, E) = N(h, D') \cap N(h, E') = N(h, C),$$

showing that $A \cap B \in \mathcal{A}$.

Part 2 is immediate since $N(f, D)^c = \bigcup_{f' \in (D \setminus \{f\})} N(f', D)$.

Define a set function $\nu: \mathcal{A} \to [0, 1]$ by setting $\nu(A) := \rho^D(f)$ where $A = N(f, D)$. The following corollary to Lemma 7 establishes that $\nu$ is well defined.

**Lemma 14.** Suppose that $N(f, D) = N(g, E)$ for two act-menu pairs $(f, D)$ and $(g, E)$. If $\rho$ satisfies PC, then $\rho^D(f) = \rho^E(g)$.

Proof. If $D$ and $E$ are singleton menus the claim follows trivially from the normalization equalities. Otherwise, since $N(f, D) = N(g, E)$ the collection $\{N(g, E), (N(f', D))_{f' \in (D \setminus f)}\}$ is pairwise disjoint and covers $\mathcal{R}$. By Lemma 7, it follows that this collection is partitional. PC then demands that $\rho^E(g) + \sum_{f' \in (D \setminus f)} \rho^D(f') = 1$, which implies that $\rho^D(f) = \rho^E(g)$.□

68
Note that \( \emptyset \in \mathcal{A} \): for instance, \( N((x, y), \{(x, x), (x, y), (y, x), (y, y)\}) = \emptyset \). In addition, \( \nu(\emptyset) = 0 \) by PC. By Theorem 1.1.4 in Durrett (2010), if \( \nu \) is additive on \( \mathcal{A} \) then there exists a unique measure \( \mu \) that extends \( \nu \) to the algebra generated \( \mathcal{A} \). Using a straightforward extension of elicitation approach in Section C.1.5 below, it is easy to see that \( \mathcal{A} \) contains all the singletons of \( \mathcal{R} \), so the algebra generated by \( \mathcal{A} \) is \( 2^\mathcal{R} \). Therefore, Theorem 2 holds once \( \nu \) is shown to be additive on \( \mathcal{A} \).

Let \( A \in \mathcal{A} \) be a disjoint union of sets \( \{A_i\}_{i=1}^n \in \mathcal{A} \). Since \( \mathcal{A} \) is a semialgebra, \( A^c \) is a disjoint union of sets \( \{B_j\}_{j=1}^m \in \mathcal{A} \). Each set \( A_i \) is defined by an act-menu pair \((f_i, D_i)\) through \( \nu(A_i) = \rho^{D_i}(f_i) \). Similarly, \( A \) is defined by an act-menu pair \((f, D)\) and each set \( B_j \) is defined by an act-menu pair \((g_j, E_j)\). The collection \( \{A_i\}_{i=1}^n \cup \{B_j\}_{j=1}^m \) partitions \( \mathcal{R} \), so by Lemma 7 the act-menu pairs \((f_i, D_i)_{i=1}^n, (g_j, E_j)_{j=1}^m \) are partitional. By the same reasoning, the collection \((f, D), (g_j, E_j)_{j=1}^m \) is also partitional. PC then implies that \( \rho^D(f) = \sum_{i=1}^n \rho^{D_i}(f_i) \) and thus that \( \nu(A) = \sum_{i=1}^n \nu(A_i) \).

### C.1.4 Proof of Corollary 5

Let \( X = \{x_1, \ldots, x_N\} \) and \( S = \{s_1, \ldots, s_{N-1}\} \). The proof of Theorem 2 follows without modification once it is shown that the collection \( \mathcal{A} \) constructed in that proof remains closed under intersection. To see this, again take two sets \( A = N(f, D) \) and \( B = N(g, E) \) from \( \mathcal{A} \). If \( |\text{range}(f) \cup \text{range}(g)| \leq N - 1 \), then the argument from Theorem 2 can be easily modified to show that \( A \cap B \in \mathcal{A} \).

So, suppose that \( \text{range}(f) \cup \text{range}(g) = X \). WLOG, assume that \( D(s_i) \) contains \( f(s_i) \) and some other element of \( x \) for all \( i \), and likewise for \( E(s_i) \) (otherwise, relabel consequences to get back to the case of the preceding paragraph). But this implies there is a cycle across \((f, D)\) and \((g, E)\). By Lemma 6, \( A \cap B = \emptyset \) and \( \emptyset \in \mathcal{A} \) (as shown in the proof of Theorem 2). Therefore, \( \mathcal{A} \) is closed under intersection.
C.1.5 Proof of Proposition 6

I show that any RUF representing $\rho$ must assign a precise probability to any relation $R \in \mathcal{R}$. Specifically, fix a generic relation $R = y_1y_2 \ldots y_N$, where each $y_i \in X$. Form the complete menu $D_R$ that satisfies

\[
D_R(s_1) = X \\
D_R(s_2) = X \setminus \{y_1\} \\
D_R(s_3) = X \setminus \{y_1, y_2\} \\
\vdots \\
D_R(s_{N-2}) = X \setminus \{y_1, y_2, \ldots, y_{N-3}\}.
\]

Then form the complete menu $D'_R$ where $D'_R(s_i) = D_R(s_i)$ for all $i \neq N - 2$, and $D'_R(s_{N-2}) = D_R(s_{N-2}) \setminus \{y_{N-2}\}$.

Let $f_R \in D_R \cap D'_R$ be defined by $f_R(s_i) = y_i$ for $i = 1, \ldots, N - 3$ and $f_R(s_{N-2}) = y_{N-1}$. It is clear from inspection of these menus and acts that if $\mu$ represents $\rho$, it must hold that

\[
\mu(\{R\}) = \rho^{D'R}(f_R) - \rho^{D'R}(f_R).
\]

This establishes that any two RUFs representing the same RCR must be identical.

C.1.6 Proof of Proposition 7

I prove the following claim, which is stronger than the desired result.

Claim. Let $X = \{x_1, x_2, \ldots, x_N\}$ and $S = \{s_1, s_2, \ldots, s_M\}$ with $M = N - 3$. For every $N > 3$, there exist two RUFs $\mu_1$ and $\mu_2$ with disjoint support that represent the same RCR, and in addition satisfy the following property: $x_iRx_j$ for all $i \leq 2$, $j > 2$ and $R \in \text{supp}(\mu_1) \cup \text{supp}(\mu_2)$. 
The proof is by induction on $N = |X|$. Fishburn’s example covers the case of $N = 4$, so assume that the claim holds for $N - 1$ with $N \geq 5$. Consider $X' = X \setminus \{x_N\}$ and $S' = S \setminus \{s_M\}$. By the IH, there exist two RUFs $\mu'_1$ and $\mu'_2$ on $X'$ with disjoint support that both rank $x_1$ and $x_2$ above all other elements and represent the same RCR when the state space is $S'$.

Form a RUF $\mu_1$ on $X$ that puts equal weight on the following relations. First, take each relation in the support of $\mu'_1$, and add in $x_N$ ranked third (that is, below $x_1$ and $x_2$). Then, take each relation in the support of $\mu'_2$ and add in $x_N$ ranked last. Then, form a RUF $\mu_2$ on $X$ that puts equal weight on the following relations. First, take each relation in the support of $\mu'_1$, and add in $x_N$ ranked last. Then, take each relation in the support of $\mu'_2$ and add in $x_N$ ranked third (that is, below $x_1$ and $x_2$). Note that the RUFs $\mu_1$ and $\mu_2$ have disjoint support.

Consider any act-menu pair $(f, D)$. I need to show that $\rho_1^D(f) = \rho_2^D(f)$, where $\rho_1$ and $\rho_2$ are the RCRs represented by $\mu_1$ and $\mu_2$ respectively. It is WLOG to assume that each outcome in the range of $f$ is distinct. To see why, suppose that $x_1 = x_2$. Form a new act $f' = (w, x_2, \ldots, x_{N-3})$ where $w \in X \setminus \{x_1, \ldots, x_{N-3}\}$, and a new menu $D'$ where $D'(s_1) = \{w\}$, $D'(s_2) = D(s_1) \cup D(s_2)$, and $D'(s_i) = D(s_i)$ for all $i \in \{3, \ldots, N-3\}$. Clearly $N(f, D) = N(g, D')$.

So, assume each element in the range of $f$ is distinct. If $x_N$ is not in the range of $f$, then $\rho_1^D(f) = \rho_2^D(f)$ because the marginal distributions of $\mu_1$ and $\mu_2$ on $X'$ are identical. If $x_N$ is in the range of $f$, then $\rho_1^D(f) = \rho_2^D(f)$ by the IH and the fact that $x_N$ is ranked either third (behind $x_1$ and $x_2$) or last with equal probability in both $\mu_1$ and $\mu_2$. 

71
C.2 Related Literature

For overviews of the literature on random choice and the BM inequalities, consult McFadden (2005) and Fishburn (1998). Section 9 of the latter work discusses early models of stochastic choice on the domains of lotteries and Savage acts, none of which address the issue of uniqueness considered here.

As discussed in the introduction, Gul and Pesendorfer (2006) axiomatize random expected utility functions for random choice over lotteries. Their representation is unique but restricts underlying preferences to those satisfying expected utility. The supplement to Gul and Pesendorfer (2006) shows that any RCR over alternatives that satisfies the BM inequalities can be extended to a RCR over lotteries that satisfies their axioms. This procedure works through the RUF and could equally well be applied to the unique RUFs characterized here.

Clark (1996) axiomatizes random utility representations for generic, potentially-infinite sets of alternatives. His main axiom is based on DeFinetti’s Coherency axiom. While powerful, it is imposed directly on sets of the form $N(x, D)$ and thus not purely on the primitives of the model. The resulting random utility representations are not unique.

Finally, Lu (2013) studies random choice over Anscombe-Aumann acts. In his representation, stochastic choice is generated by uncertainty over probabilistic beliefs on the state space and the utility function on (lotteries over) alternatives is fixed. Therefore, despite similarities in the domain Lu’s result is not directly comparable to the results in this paper.