Non-Additive Random Utility Functions*

Morgan McClellon†

May 30, 2015

Abstract

This paper 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 non-additive random utility functions: that is, 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. These results help to decipher the Block-Marschak inequalities, and are applied to study the relationship between random choice over finite sets and random choice over lotteries.

*I would like to thank Drew Fudenberg and Jerry Green for helpful discussions. Tomasz Strzalecki suggested the line of research which led to this project and provided excellent advice. All remaining errors are my own. This paper is based on a chapter of my dissertation at Harvard University.

†Contact mcclellon.morgan@gmail.com.

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.

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 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 $\rho\colon \mathcal{D} \times X \to [0,1]$ 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 1. 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):

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

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 3 shows that any BM inequality indexed by $(x,\{x\})$ is implied by other BM inequalities and the normalization equalities. $\diamondsuit$

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.

3 Capacity Representations

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

Definition 1. A *capacity* on a finite set S is a function $\nu\colon 2^S \to [0,1]$ 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 2. A capacity ν on $\mathcal{R}$ represents a RCR ρ 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 3 (Regularity)**.** A RCR ρ 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 ρ be a RCR on X. TFAE:*

1. *ρ satisfies regularity.*
2. *There exists a capacity ν on $\mathcal{R}$ that represents ρ.*
3. *There exists a superadditive capacity ν on $\mathcal{R}$ that represents ρ; i.e. ν 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 ρ 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 $\mathcal{A} \subseteq 2^{\mathcal{R}}$ 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 $\mathcal{A}$, define a set function $\hat{\nu}\colon \mathcal{A} \to [0, 1]$ 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 $\mathcal{A}$ since ρ 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

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

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

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

The set $\mathcal{A}$ 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 ρ is a probability measure on $\mathcal{R}$. Unfortunately, the collection $\mathcal{A}$ is not an algebra,[1] so traditional extension theorems cannot be used and the proof strategy above fails. A RUF μ must of course be additive, so the behavior of μ 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 McClellon (2015).

Traditional RUFs have been interpreted in two ways. First, a RUF μ can be taken to represent a distribution of preferences for a single agent; in this case ρ 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, μ could represent the distribution of individuals' preferences across a society, where each individual has deterministic preferences and ρ 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 this setting.

Under a third, novel interpretation ρ captures the predictions on an expert regarding the choice behavior of either an individual or a population.[2] In this case, μ 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)).

[1]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.

[2]I thank Tomasz Strzalecki for suggesting this interpretation.

4 Unraveling the BM Inequalities

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 4. A capacity ν 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.

Example 2. Let ρ 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 ρ is represented by a capacity ν. Let

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

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 ν cannot be supermodular. $\diamondsuit$

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

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

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 5. A capacity ν 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 be 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 ρ be a RCR on X and let $(D, x) \in \mathcal{D} \times X$ with $x \in D$. If $|D| = N - K$, the BM inequality indexed by (D, x) is a necessary condition for ρ to be represented by a K-monotonic capacity.*

Lemma 3 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 ρ can be represented by an $(N-2)$-monotonic capacity is equivalent to showing that all the BM inequalities hold, and thus that ρ can be represented by a standard RUF.

Corollary 1. *A RCR ρ can be represented by a RUF if and only 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 1, 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 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.

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 $\rho\colon \mathcal{D} \to \Delta(\Delta(X))$, 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 μ on the set of expected utility functions over $\Delta(X)$.[3] 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 $\rho^{D'}(p) \leq \rho^D(p)$.

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 λ.

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 ρ be a RCR on X. If ρ 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

[3]See Gul and Pesendorfer (2006) for the precise definition of a RUF.

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 ρ, and in turn uses this representation to extend ρ to $\Delta(X)$:

Proof. Let ρ be a RCR on X and let ν be a capacity representation for ρ, 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 ν 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. ■

References

Barberá, S. and Pattanaik, P. K. (1986). Falmagne and the Rationalizability of Stochastic Choices in Terms of Random Orderings. *Econometrica*, 54(3):707–715.

Block, H. D. and Marschak, J. (1960). Random Orderings and Stochastic Theories of Responses. In *Contributions to Probability and Statistics*. Stanford University Press.

Chambers, C. P. (2008). Proper Scoring Rules for General Decision Models. *Games and Economic Behavior*, 63.

Chateauneuf, A. and Jaffray, J.-Y. (1989). Some Characterizations of Lower Probabilities and Other Monotone Capacities Through the Use of Möbius Inversion. *Mathematical Social Sciences*, 17:263–283.

Dempster, A. P. (1967). Upper and Lower Probabilities Induced by a Multivalued Mapping. *Annals of Mathematical Statistics*.

Denneberg, D. (1994). *Non-Additive Measure and Integral*. Kluwer Academic Publishers.

Falmagne, J. C. (1978). A Representation Theorem for Finite Random Scale Systems. *Journal of Mathematical Psychology*, 18:52–72.

Fiorini, S. (2004). A short proof of a theorem of Falmagne. *Journal of Mathematical Psychology*, 48(1):80–82.

Fishburn, P. C. (1998). Stochastic Utility. In Barberá, S., Hammond, P. J., and Seidl, C., editors, *Handbook of Utility Theory*, pages 272–311.

Gul, F. and Pesendorfer, W. (2006). Random Expected Utility. *Econometrica*, 74(1):121–146.

Gul, F. and Pesendorfer, W. (2013). Random Utility Maximization with Indifference. *Working Paper*, pages 1–25.

McClellon, M. (2015). Unique Random Utility Representations. *Working Paper*.

McFadden, D. L. (2005). Revealed stochastic preference: a synthesis. *Economic Theory*, 26(2):245–264.

Shafer, G. (1976). *A Mathematical Theory of Evidence*. Princeton University Press.

A Proof of Proposition 1

To complete the proof of this proposition, it remains to show that (1) ⇒ (3). Let ρ 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 ρ:

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

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

Next, extend $\hat{\nu}$ to a set function ν 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 ν is monotonic with respect to set inclusion on $\mathcal{A}'$. However, it remains to be verified that ν 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 ν,

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

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 ν is superadditive.

B 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 2. *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 ν 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 $A_1, \dots, A_k$ in $\mathcal{A}$. Since $\hat{\nu}$ is strongly superadditive, ν is well-defined. Clearly ν is monotonic, and it is superadditive by the same logic as in the proof of Proposition 1 above. ■

C Proof of Proposition 3

Let $(D, x) \in \mathcal{D} \times X$ 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 $\{y_1, \dots, y_K\}$. For each $i = 1, \dots K$, define $D_i := D \cup \{y_i\}$ and $A_i := N(D_i, x)$. Suppose that ν is a capacity representation for ρ. Then

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

and

$$\sum_{\emptyset \neq I \subseteq 1, \dots, K} (-1)^{|I|+1} \nu\left(\bigcap_{k \in I} A_k\right) = \sum_{C: D \subsetneq C} (-1)^{|C \setminus D|+1} \rho^C(x).$$

If the BM inequality indexed by (D, x) is violated, it follows from these two equations that

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

showing that ν is not K-monotonic. Note that Proposition 2 is a special case of Proposition 3 with $K = 2$.

D Proof of Corollary 1

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

Lemma 3. *For any RCR ρ, if BM$(\rho, D, x) \geq 0$ for all pairs (D, x) with $|D| = 2$, then BM$(\rho, \{x\}, x) \geq 0$.*

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

$$\begin{aligned}\mathrm{BM}(\rho,\{x\},x) &:= \sum_{U:x\in U}(-1)^{|U|-1}\rho^U(x)\\ &= 1+\sum_{y\in S\setminus\{x\}}\mathrm{BM}(\rho,\{x,y\},y)+\sum_{i=1}^{N-1}(-1)^i\binom{N-1}{i}\\ &= \sum_{y\in S\setminus\{x\}}\mathrm{BM}(\rho,\{x,y\},y). \qquad \blacksquare\end{aligned}$$

Suppose that a RCR can be represented by a capacity ν 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.